International Journal of High Speed Electronics and Systems, Vol. 11, No. 1 (2001) 159-215
© World Scientific Publishing Company

Scaling of InGaAs/InAlAs HBTs for High Speed Mixed-Signal and mm-Wave ICs.

M.J.W. Rodwell, M. Urteaga, Y. Betser [*], T. Mathew, P. Krishnan, D. Scott, S. Jaganathan
D. Mensa, J. Guthrie[†], R. Pullela [‡], Q. Lee[§], B. Agarwal [¶], U. Bhattacharya [‖], S. Long
Department of Electrical and Computer Engineering, University of California
Santa Barbara, CA 93106, U.S.A.

and

S.C. Martin, R. P. Smith[**]
NASA Jet Propulsion Labs, California Institute of Technology
Pasadena, CA , USA

High bandwidths are obtained with heterojunction bipolar transistors by thinning the base and collector layers, increasing emitter current density, decreasing emitter contact resistivity, and reducing the emitter and collector junction widths. In mesa HBTs, minimum dimensions required for the base contact impose a minimum width for the collector junction, frustrating device scaling. Narrow collector junctions can be obtained by using substrate transfer processes, or –if contact resistivity is greatly reduced – by reducing the width of the base Ohmic contacts in a mesa structure. HBTs with submicron collector junctions exhibit extremely high f_{max} and high gains in mm-wave ICs. Logic gate delays are primarily set by depletion-layer charging times, and neither f_τ nor f_{max} is indicative of logic speed. For high speed logic, epitaxial layers must be thinned, emitter and collector junction widths reduced, current density increased, and emitter parasitic resistance decreased. Transferred-substrate HBTs have obtained 21 dB unilateral power gain at 100 GHz. If extrapolated at -20 dB/decade, the power gain cutoff frequency f_{max} is 1.1 THz. Transferred-substrate HBTs have obtained 295 GHz f_τ. Demonstrated ICs include lumped and distributed amplifiers with bandwidths to 85 GHz, 66 GHz master-slave flip-flops, and 18 GHz clock rate $\Delta - \Sigma$ ADCs.

1. Introduction

Research in wide bandwidth heterojunction bipolar transistors (HBTs) [1] [2] is driven by applications in high-frequency communications and radar. In optical fiber communications, integrated circuits for 40 Gb/s transmission are now in development [5], [6]. Emergence of 160 Gb/s transmission equipment in the near future must rely on a timely and substantial improvement in the bandwidth of semiconductor elec-

[*] Y. Betser is now with Anadigics Corp., Israel.
[†] J. Guthrie is now with Nortel Networks, Inc.
[‡] D. Mensa, R. Pullela, and S. Jaganathan are now with Gtran, Inc
[§] Q. Lee is now with Lucent Technologies
[¶] B. Agarwal is now with Conexant Corp.
[‖] U. Bhattacharya is now with Intel Corp.
[**] R. P. Smith is now with Cree Research, Inc.

tronics. 160 Gb/s fiber transmission will require amplifiers with flat gain and linear phase over a $\sim$ DC-110 GHz bandwidth and master-slave latches [3] (used in decision circuits, multiplexers, and phase-lock loops) operable at 80 GHz or 160 GHz clock frequency.

A second set of driving applications are wideband, high-resolution analog-digital converters, digital-analog converters, and direct digital frequency synthesizers [8]. Increased bandwidths of these mixed-signal ICs will increase the bandwidth and frequency agility of military radar and communications systems [4]. In ADCs and DACs, very high resolution is obtained using oversampling techniques [7,9], with clock frequencies $\sim$ 100 $\times$ the signal bandwidths. In high resolution ADCs, to avoid metastability errors in latched comparators driven by small input signals, the circuit time constants must be much smaller than the periods of the clock signals employed. Similar design constraints apply to high-resolution DACs. High resolution ADCs and DACs consequently require transistor bandwidths $10^2 : 1$ to $10^4 : 1$ larger than the signal frequencies involved. Transistors with several hundred GHz f_τ and f_{max} would enable high-resolution microwave mixed-signal ICs.

A third driving application is in monolithic millimeter-wave integrated circuits (MIMICs). In microwave and millimeter-wave receivers, the low-noise RF preamplifier, several stages of amplification, and frequency conversion (a mixer), are typically implemented as small-scale monolithic circuits. Similar MIMICs are used in the transmitter. The operating frequency is set by the application, but progressive improvements in transistor bandwidths permit the evolution of radar and communications ICs to progressively higher frequencies. A transistor with a 1 THz power-gain cutoff frequency would provide useful gain over the full 30-300 GHz millimeter-wave band. This would permit e.g. digital radio links with millimeter-wave carrier frequencies and 1–10 Gb/s channel capacities. Until recently, III-V high-electron-mobility field-effect-transistors (HEMTs) have shown f_{max} superior to that of HBTs, and have dominated in MIMICs. With recent work on scaling of HBTs to submicron dimensions [43], HBT power-gain cutoff frequencies now exceed those of HEMTs, and HBTs can compete for application in MIMICs.

In high-speed digital and mixed-signal applications, III-V HBTs must compete with their silicon counterparts. The primary advantage of III-V HBTs is superior bandwidth, and the primary disadvantage the relative immaturity of the technology, with consequently higher cost and lower scales of integration. There are several factors contributing to the superior bandwidth of III-V HBTs. For HBTs grown on GaAs or InP substrates, available lattice-matched materials allow use of an emitter whose bandgap energy is much larger than that of the base [1]. This allows the base doping to be increased to the limits of incorporation in growth, $\sim 10^{20}/\mathrm{cm}^3$, and results in very low base sheet resistance. 600 Ω/square sheet resistance and 0.15 ps base transit time is readily obtained in a Be-doped InGaAs base of 400 Å thickness. In contrast, constraints of allowable lattice mismatch in Si/SiGe HBTs limit the allowable Ge:Si alloy ratio. The emitter-base bandgap energy difference is then much smaller than in III-V HBTs, and base dopings are consequently lower. 4-8

kΩ/square base sheet resistivity is typical of SiGe HBTs [13]. High electron velocities are a second significant advantage of III-V HBTs. In InAlAs/InGaAs HBTs with 0.2–0.3 μm collector thickness, effective collector electron velocities exceed 4×10^7 cm/s, approximately 4:1 higher than observed in Si. This high electron velocity results in high current-gain cutoff frequencies.

With the exception of transferred-substrate HBTs (discussed subsequently), best reported results of InP-based HBTs include 225 GHz f_τ and 300 GHz f_{max} [35], [14]. Si/SiGe HBTs [10], [11] have obtained 156 GHz f_τ. Thus, despite the advantages of III-V HBTs provided by superior materials properties, Si bipolar junction transistors (BJTs) and Si/SiGe HBTs remain highly competitive. The high bandwidths of Si/SiGe HBTs arise in part from aggressive submicron scaling. In devices with 0.14 μm emitter-base junction widths, 92 GHz f_τ and 108 GHz f_{max} have been reported [12]. Self-aligned polysilicon contacts reduce both the parasitic collector-base capacitance and the base resistance. In marked contrast to the aggressive submicron scaling and aggressive parasitic reduction employed in Si/SiGe HBTs, III-V HBTs are typically fabricated with 1–2 μm emitter junction widths and 3–5 μm collector-base junction widths. This is remarkable in an era when commodity microprocessors are available with tens of millions of transistors at 0.13 μm gate lengths. Deep submicron scaling will improve the bandwidth of III-V heterojunction bipolar transistors, and is critical to their continued success.

To obtain improved HBT bandwidths by scaling, transit times are reduced by decreasing the thicknesses of the base and collector epitaxial layers. Important RC charging times are reduced by laterally scaling the base and collector junction widths. Most significant among several limits to HBT submicron scaling is the extrinsic (parasitic) collector-base junction lying under the base Ohmic contacts. The required minimum size for the base Ohmic contacts places a lower limit on the size of the collector-base junction, preventing submicron junction scaling. We have developed a substrate transfer process which allows fabrication of HBTs with submicron emitter-base and collector-base junctions lying on opposing sides of the base epitaxial layer. With this device, f_{max} increases rapidly with scaling. With transferred-substrate HBTs, 1.1 THz *extrapolated* power-gain cutoff frequencies and 295 GHz current-gain cutoff frequencies have been obtained. Further improvements in f_τ requires further epitaxial scaling, together with increased operating current density and greatly improved emitter parasitic resistance.

2. HBT scaling

In HBTs, thinning the base and collector epitaxial layers reduces the carrier transit times but increases the base resistance and the collector-base capacitance. These can be subsequently reduced by reducing the lithographically-defined widths of the emitter-base and collector-base junctions. To simultaneously obtain both high f_τ and high f_{max}, device epitaxial and lithographic dimensions must be concurrently scaled. Below we examine the limits to HBT scaling.

Figure 1 shows a simplified cross-section of a mesa HBT. To form the transis-

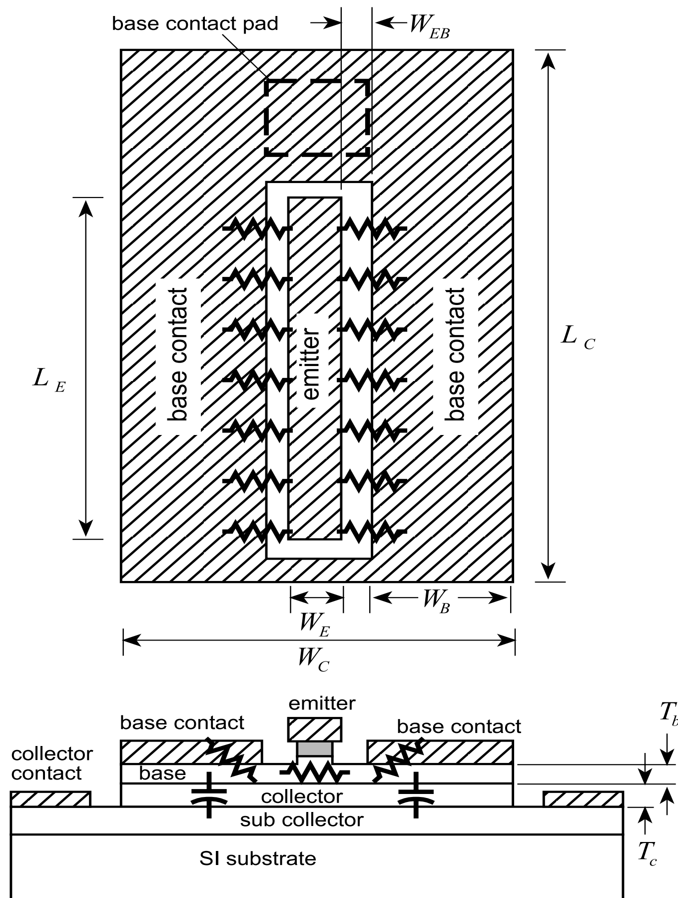


Figure 1: Plan and cross-section of a typical mesa HBT. The emitter-base junction has width W_e, length L_e and area $A_e = L_eW_e$, while the collector-base junction has width W_c, length L_c and area $A_c = L_cW_c$

tor, the emitter, base, and collector layers first grown by molecular-beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD) on a semi-insulating substrate. The HBT junctions are formed by a series of patterned etches, and contacts formed by depositing metal. This results in a device structure where the collector-base junction must lie under the full area of the base Ohmic contacts. There is also a parasitic collector-base junction lying under the area of the base contact pad. In this device structure, the collector-base junction must be substantially larger than the emitter dimensions. At the sides of the emitter stripe, the base Ohmic contact must be at least one Ohmic contract transfer length L_{contact} in order to obtain low contact resistance. In an InGaAs-base HBT with 400 Å base thickness and $5 \times 10^{19}/\text{cm}^3$ doping, $L_{\text{contact}} \simeq 0.4$ μm. Lithographic alignment tolerances between emitter and collector also constrain the minimum collector-base junction dimensions. Dependent upon the process minimum feature size and the length of the emitter stripe, the base contact pad area can contribute as much as 50 % of the total collector-base capacitance.

2.1. *Factors determining* f_τ

Before examining scaling for high cutoff frequencies, relevant HBT parameters must first be calculated. The current-gain cutoff frequency is

$$\frac{1}{2\pi f_\tau} = \tau_b + \tau_c + \frac{kT}{qI_c}\,(C_{je} + C_{cb}) + (R_{ex} + R_c)C_{cb}, \tag{1}$$

where R_{ex} and R_c are the parasitic emitter and collector resistances, C_{cb} is the collector junction capacitance, and I_c the collector current.

First examine the base transit time τ_b. If a linear grading of the base semiconductor bandgap energy with position is used to reduce τ_b, then [15]

$$\begin{aligned}\tau_b &= \frac{T_b^2}{D_n}\left(\frac{kT}{\Delta E}\right) - \frac{T_b^2}{D_n}\left(\frac{kT}{\Delta E}\right)^2\left(1 - e^{-\Delta E/kT}\right) \\ &+ \frac{T_b}{v_{exit}}\left(\frac{kT}{\Delta E}\right)\left(1 - e^{-\Delta E/kT}\right)\,,\end{aligned} \tag{2}$$

where ΔE is the grading in the base bandgap energy and T_b the base thickness. The base exit velocity v_{exit} is of the order of $(kT/m^*)^{1/2}$ for an ungraded base [15], and is somewhat larger with base bandgap grading. D_n is the base minority carrier diffusivity and m^* the electron effective mass. Equation 2 is derived from the drift-diffusion relationship, and is accurate only if the predicted τ_b is large in comparison with the momentum relaxation time $\tau_m = D_n m^*/kT$ [16]. Using the parameters of an InGaAs base at $5 \times 10^{19}/\text{cm}^3$ doping ($D_n = 40$ cm^2/sec, $v_{exit} \sim 3 \times 10^7$ cm/s, τ_m=35 fs), we note that 52 meV bandgap grading is sufficient to reduce τ_b by $\sim$ 2:1. For a thick base layer or a large v_{exit} , $\tau_b \propto T_b^2$; with InGaAs base layers below $\sim$400 Å thickness, the exit velocity term in eqn. 2 adds a significant correction.

The collector transit time τ_c is the mean delay of the collector displacement current, and is given by [17], [18]

$$\tau_c = \int_0^{T_c} \frac{(1 - x/T_c)}{v(x)} dx \equiv \frac{T_c}{2v_{\text{eff}}}, \tag{3}$$

where $v(x)$ is the position-dependent electron velocity in the collector drift region and v_{eff} an effective electron velocity. τ_c is most strongly dependent upon the electron velocity in the proximity of the base, and becomes progressively less sensitive to the electron velocity as the electron passes through the collector [18]. At low collector-base bias voltages, electrons must traverse a significant fraction of the collector drift region before acquiring sufficient kinetic energy (0.55 eV for InGaAs [19], 0.6eV for InP [20]) to undergo Γ-L scattering [17], [18], and $v(x)$ is fortuitously highest near the base. In thin InGaAs or InP layers, $v_{\text{eff}} = 3\text{–}5 \times 10^7$ cm/s. For scaling analysis, we will take $\tau_c \propto T_c$.

In InAlAs/InGaAs HBTs with $T_b \cong 400$ Å and $T_c \cong 0.2\ \mu$m, $f_\tau \simeq 250$ GHz, and the RC charging terms in eqn. 1 comprise 35% of the total forward delay. These terms must be considered in detail.

First consider the charging time $[kT/qI_c]C_{cb}$. This term has a major impact upon digital circuit delay (section 3.1) and is reduced by increasing the collector current density to limits set by collector space-charge screening (the Kirk effect [21]). If the collector doping N_d is chosen so as to obtain a fully-depleted collector at zero bias current and the applied V_{cb}, we must have

$$V_{cb} + \phi = qN_d T_c^2/2\epsilon\ , \tag{4}$$

while base pushout occurs at a current density J_{max} satisfying

$$V_{cb} + \phi = (J_{max}/v_{sat} - qN_d)\, T_c^2/2\epsilon\ , \tag{5}$$

hence the maximum collector current before base pushout is

$$I_{c,max} = A_e(V_{cb} + \phi) 4\epsilon v_{sat}/T_c^2 \propto A_e/T_c^2, \tag{6}$$

where v_{sat} is an (assumed) uniform electron velocity within the collector. With undoped collectors, $I_{c,max}$ is 2:1 smaller than in eqn. 6. The collector capacitance is $C_{cb} = \epsilon A_c/T_c$. With the HBT biased at $I_{c,max} \propto 1/T_c^2$, $(kT/qI_C)C_{cb} \propto T_c\,(A_c/A_e)$. This delay term is thus minimized by scaling (reducing T_c), but bias current densities must increase in proportion to the square of the desired fractional improvement in f_τ.

The emitter charging time ($C_{je}[kT/qI_c]$ in eqn. 1) is a significant determinant of f_τ, and also plays a major role in ECL logic delay (section 3.1). If we were to assume that C_{je} were simply a depletion capacitance, it would be reasonable to expect that this charging time could be minimized simply by making the emitter-base depletion region very thick, by use of very low emitter doping, combined with a thick bandgap grading region in the base-emitter heterojunction. Clearly, this

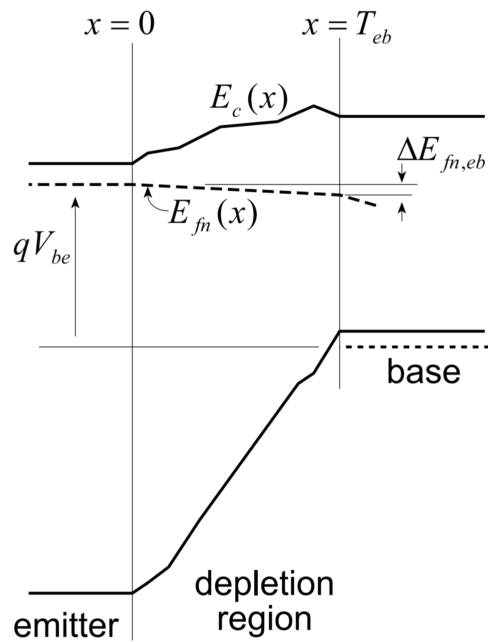


Figure 2: Band diagram of the HBT emitter-base junction. If the base-emitter junction thickness T_{eb} is excessive, HBT performance will be degraded by either stored charge or by excessive potential drops in the depletion layer.

approach must fail somehow in the limit of very large depletion thicknesses. We must examine design of the emitter-base junction in detail to determine the limits to the emitter-base depletion thickness, and to understand how the junction design must be modified as the transistor is scaled for increased device bandwidth.

In order to support a high emitter current density without a substantial potential drop in the emitter-base depletion layer, a high electron density $n(x)$ must be present within the emitter-base junction. In high speed HBTs the thickness T_{eb} of the emitter-base depletion layer must then be small if significant charge storage effects are to be avoided. Figure 2 shows a band diagram of the base-emitter depletion region. $n(x) = N_c \exp[-q(E_c(x) - E_{f,n}(x))/kT]$, where N_c is the conduction band effective density of states, $E_c(x)$ is the conduction-band energy and $E_{f,n}(x)$ the electron quasi-Fermi level. An arbitrary conduction-band profile $E_c(x)$ can be obtained through combined bandgap grading and doping. Under modulation of V_{be}, $\partial n(x)/\partial V_{be} = n(x)(q/kT)(x/T_{e,b})$. The ideality factor N is defined by the relationship $I_c \propto e^{qV_{be}/NkT}$; gradients in E_{fn} in the emitter-base depletion region result in N greater than unity, with

$$N = 1 + \frac{1}{q}\frac{\partial\left(\Delta E_{fn,eb}\right)}{\partial V_{be}} \ . \tag{7}$$

In the base-emitter depletion region, $dE_{fn}/dx = -J/\mu_{n,eb}n(x)$, while in the base $J_n = qn(T_{eb})D_n/T_b\Gamma$. Here, $\mu_{n,eb}$ is the electron mobility in the junction (due to the low doping in the grade, this mobility is significantly larger than that of the base) and $\Gamma = kT/\Delta E - (kT/\Delta E - D_n/v_{exit}T_b)e^{-\Delta E/kT}$ is a factor involving the base bandgap grading ($\Gamma \simeq 1$ for an ungraded base). Combining these relationships, the ideality factor is

$$N = 1 + \frac{T_{eb}}{T_b}\frac{\mu_n}{\Gamma\mu_{n,eb}}\int_0^1 \frac{n(T_{eb})}{n(\zeta T_{eb})}(1-\zeta)d\zeta, \tag{8}$$

Where $\zeta = x/T_{eb}$ is a normalized position variable, and μ_n is the electron mobility in the base. To obtain a low ideality factor, T_{eb}/T_b must not be large, and the electron density $n(x)$ in the junction must be kept high. Unless T_{eb}/T_b is kept small, the high $n(x)$ will result in significant charge storage. Using methods similar to those used to derive the collector transit time [17], [18] (eqn. 3),

$$C_{je}/A_e = \epsilon/T_{eb} + \frac{\partial}{\partial V_{be}}\left[\int_0^{T_{eb}} \left(x/T_{eb}\right) qn(x)dx\right]. \tag{9}$$

The term $(kT/qI_c)C_{je}$ in eqn. 1 can be then written as

$$\begin{aligned}(kT/qI_c)C_{je} &= \left(\frac{\epsilon A_e}{T_{eb}}\right)\left(\frac{kT}{qI_c}\right) \\ &+ \frac{\Gamma T_{eb}T_b}{D_n}\int_0^1 \frac{n(\zeta T_{eb})}{n(T_{eb})}\zeta^2 d\zeta.\end{aligned} \tag{10}$$

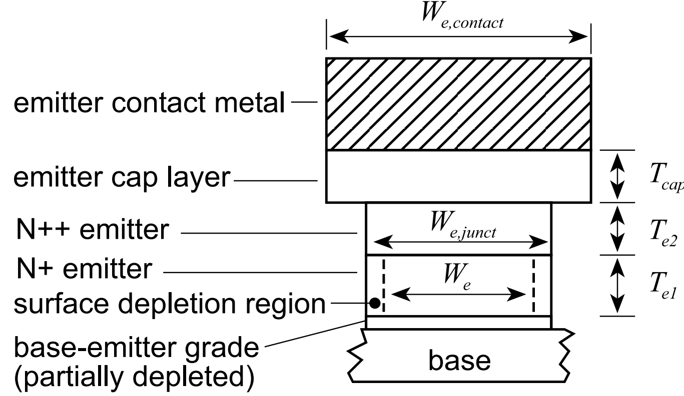


Figure 3: Cross-section of the emitter layers within a typical HBT, comprising an a heavily-doped semiconductor contact ("cap") layer, a low-resistance $N{++}$ emitter layer, and the $N+$ emitter. Lateral depletion of the $N+$ emitter can be significant in submicron devices.

The first term in eqn. 10 results from the depletion-layer capacitance, and is minimized using high bias current densities $J_e = I_e/A_e$; the second term reflects storage of mobile electron charge within the depletion layer, and is minimized by reducing $T_{eb}T_b$.

In eqn. 1, the delay term $R_{ex}C_{cb}$ is a major limit to HBT scaling for high f_τ. Further, R_{ex} contributes significantly to ECL logic delay. Because of the relative sizes of the emitter and collector Ohmic contacts, in a *well-designed* submicron HBT, R_c is 4:1 to 10:1 smaller than R_{ex} and R_cC_{cb} can be neglected in a first analysis. R_{ex} must first be calculated. The emitter layer structure of a typical HBT (fig. 3) contains a heavily-doped and narrow-bandgap contact ("cap") layer, and a heavily-doped $N{++}$ wide-bandgap emitter layer. A portion of the emitter layer may be more lightly ($N+$) doped for reduced junction capacitance, and may be of several hundred Å thickness to avoid dopant diffusion from the $N{++}$ layer into the emitter-base junction. If heterointerfaces are properly graded to avoid conduction-band barriers between layers, the parasitic emitter resistance is

$$\begin{aligned} R_{ex} &= \rho_{c,e}/L_eW_{e,contact} + \rho_{cap}T_{cap}/L_eW_{e,contact} \\ &+ \rho_{e2}T_{e2}/L_eW_{e,junct} + \rho_{e1}T_{e1}/L_eW_e \ , \end{aligned} \quad (11)$$

where $\rho_{c,e}$ is the emitter specific Ohmic contact resistivity, and ρ_{cap}, ρ_{e2}, and ρ_{e1} are the bulk resistivities of the cap, $N{++}$, and $N+$ emitter layers. For submicron emitters, the junction width $W_{e,junct}$ is significantly smaller than the contact width $W_{e,contact}$ due to lateral undercutting of the emitter during etching of the emitter-base junction, and the electrically-active emitter width W_e can be significantly smaller than $W_{e,junct}$ because of the presence of surface (edge) depletion regions of

width $(2\epsilon\phi/qN_{e1})^{1/2}$, where N_{e1} is the $N+$ layer doping and ϕ is the bandbending due to pinning of the Fermi energy at the surface. For simplicity in scaling analysis, we will approximate

$$R_{ex} \simeq \rho_e/A_e \tag{12}$$

where ρ_e is a fitted parameter, approximately $50\Omega{-}\mu\text{m}^2$ for submicron InAlAs/InGaAs HBTs fabricated to date at UCSB. In InAlAs/InGaAs HBTs we have fabricated, $\rho_{c,e} = 20\Omega - \mu\text{m}^2$ when InGaAs contacts at $10^{19}/\text{cm}^3$ doping are employed, and $\rho_{c,e} = 4\Omega - \mu\text{m}^2$ for contacts to InAs layers at $2 \times 10^{19}/\text{cm}^3$ doping. The $\rho_{e1}T_{e1} = 5.5\Omega - \mu\text{m}^2$ resistance of the $N+$ InAlAs layer ($8 \times 10^{17}/\text{cm}^3$ doping, 700 Å thickness) is significant in submicron devices for which W_e is 2:1 to 4:1 smaller than $W_{e,contact}$. To avoid such emitter size effects, deep submicron HBTs should use $\gg 10^{18}/\text{cm}^3$ emitter doping.

The $R_{ex}C_{cb}$ charging time can now be examined. Since $C_{cb} = \epsilon A_c/T_c$,

$$R_{ex}C_{cb} = \left(\frac{\epsilon\rho_e}{T_c}\right)\left(\frac{A_c}{A_e}\right) = 28 \ \text{ fs} \times \left(\frac{A_c}{A_e}\right), \tag{13}$$

if $\rho_e = 50\ \Omega - \mu\text{m}^2$ and $T_c = 0.2\ \mu\text{m}$. This a significant delay. In HBTs we have fabricated with 275 GHz peak f_τ, the substrate transfer process allows A_c/A_e to be kept small at 2.3:1, yet $R_{ex}C_{cb}$ still constitutes 11% of the total $1/2\pi f_\tau = 0.58$ ps forward delay. In mesa HBTs (fig. 1) A_c/A_e is often larger than 2.3:1 and hence $R_{ex}C_{cb}$ will contribute a larger delay. Because $R_{ex}C_{cb} \propto 1/T_c$, thinning the collector to reduce τ_c also increases $R_{ex}C_{cb}$.

To increase HBT current gain cutoff frequencies, the base and collector layers must be thinned and the bias current density increased. Thinning the collector increases $R_{ex}C_{cb}$, imposing a limit to scaling. Limits to bias current density imposed by device reliability, and loss in breakdown voltage with reduced collector thickness, are two further potential limits to scaling. Finally, unless the device structure of fig. 1 is laterally scaled, vertical HBT scaling for increased f_τ will result in *reduced* power-gain cutoff frequencies f_{max}.

2.2. *Lithographic scaling for high* f_{max}

Regardless of the value of f_τ, transistors cannot provide power gain at frequencies above f_{max}. Independent of f_τ, f_{max} defines the maximum usable frequency of a transistor in either narrowband reactively-tuned or broadband distributed circuits [22]. In more general analog and digital circuits (section 3.1), all transistor parasitics play a significant role. The f_τ and f_{max} of a transistor are then cited to give a first-order summary of the device transit delays and of the magnitude of its dominant parasitics.

In an HBT with base resistance R_{bb} and collector capacitance C_{cb}, the power-gain cutoff frequency is approximately $f_{max} \simeq (f_\tau/8\pi R_{bb}C_{cbi})^{1/2}$. The base-collector junction is a distributed network, and $R_{bb}C_{cbi}$ represents an effective, weighted time constant.

The base resistance (fig. 1) R_{bb} is composed of the sum of contact resistance R_c, base-emitter gap resistance R_{gap}, and spreading resistance under the emitter R_{spread}. With base sheet resistance ρ_s, and specific (vertical) contact access resistance ρ_c, we have

$$\begin{aligned} R_{bb} &= R_{b,cont} + R_{gap} + R_{spread} \\ R_{b,cont} &= \sqrt{\rho_s \rho_c}/2L_e \\ R_{gap} &= \rho_s W_{eb}/2L_e \\ R_{spread} &= \rho_s W_e/12L_e. \end{aligned} \tag{14}$$

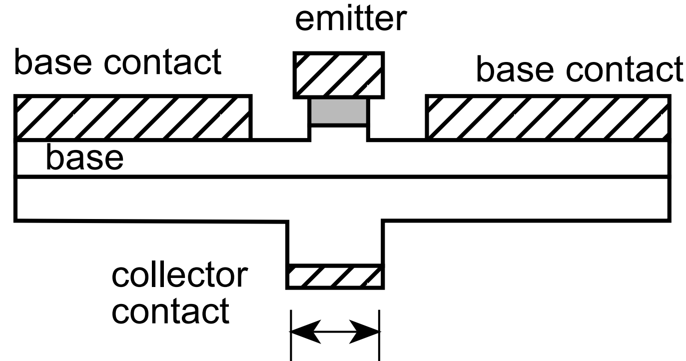
To compute f_{max}, we must find C_{cbi}. Because the base-collector junction parasitics are distributed, calculation of $R_{bb}C_{cbi}$ is complex, and will be deferred until section 2.3. As a first (and very rough) approximation, we will first compute $R_{bb}C_{cb}$, e.g. the product of the base resistance and the full capacitance $C_{cb} = \epsilon A_c/T_c$ of the collector-base junction,

$$\begin{aligned} R_{bb}C_{cb} &= \left[(\sqrt{\rho_s\rho_c} + \rho_s W_{eb}) \left(\frac{\epsilon}{2}\right) \left(\frac{L_c}{L_e}\right) \right] \left[\frac{W_c}{T_c} \right] \\ &+ \left[\left(\frac{\rho_s \epsilon}{12}\right) \left(\frac{L_c}{L_e}\right) \right] \left[\frac{W_c W_e}{T_c} \right]. \end{aligned} \tag{15}$$

Consider the influence of device scaling on the time constant $R_{bb}C_{cb}$. Decreasing the base thickness to reduce τ_b increases the base sheet resistivity ρ_c, increasing $R_{bb}C_{cb}$. Decreasing the collector thickness T_c to reduce τ_c directly increases $R_{bb}C_{cb}$, as is shown explicitly in eqn. 15.

Low $R_{bb}C_{cb}$, and consequently high f_{max}, is obtained by scaling the emitter and collector junction widths W_e and W_c to submicron dimensions. Reducing the emitter width W_e alone reduces towards zero the component of $R_{bb}C_{cb}$ associated with the base spreading resistance (the second term in eqn. 15). In the normal triple-mesa HBT (fig. 1), the base Ohmic contacts must be at least one contact transfer length ($L_{\text{contact}} = (\rho_c/\rho_s)^{1/2}$), setting a minimum collector junction width W_c. The component of $R_{bb}C_{cb}$ associated with the base contact resistance (the first term in eqn. 15) has a minimum value, independent of lithographic limits. Consequently, $f_{\max}$ does not increase rapidly with scaling. Given this minimum $R_{bb}C_{cb}$, attempts to obtain high f_τ by thinning the collector have resulted in decreased f_{max}, frustrating efforts to improve HBT bandwidths.

If the parasitic collector-base junction is eliminated, f_{max} will instead increase rapidly with scaling. The collector-base junction need only be present where current flows, e.g. under the emitter. We have fabricated such a device (figure 4) using substrate transfer processes. The emitter and collector junctions can be of equal



$$W_C = W_E = W$$

Figure 4: Cross-section of an idealized HBT with the collector-base junction lying only under the emitter. Such device structures can be formed using substrate transfer processes.

width, hence $W_c = W_e$. The base-collector time constant becomes

$$\begin{aligned} R_{bb}C_{cb} &= \left[(\sqrt{\rho_s \rho_c} + \rho_s W_{eb}) \left(\frac{\epsilon}{2} \right) \left(\frac{L_c}{L_e} \right) \right] \left[\frac{W_e}{T_c} \right] \\ &+ \left[\left(\frac{\rho_s \epsilon}{12} \right) \left(\frac{L_c}{L_e} \right) \right] \left[\frac{W_e^2}{T_c} \right]. \end{aligned} \tag{16}$$

With submicron scaling of the emitter and collector junction widths, the first term in eqn. 16 dominates, and f_{max} increases as the inverse square root of the process minimum feature size.

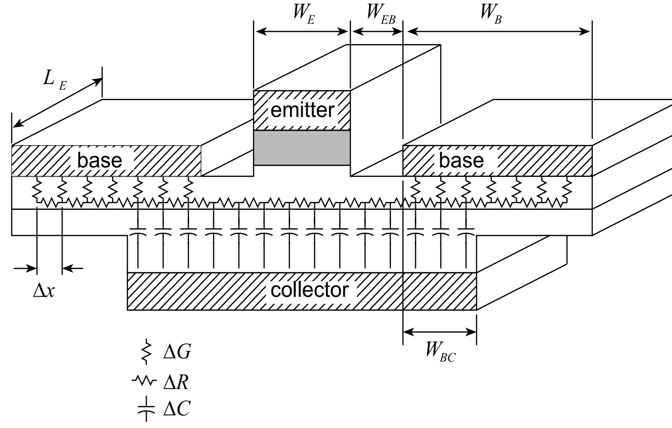


Figure 5: Distributed model of the HBT base-collector junction for accurate calculation of $R_{bb}C_{cbi}$. With mesh spacing Δx, $\Delta G = L_e \Delta x/\rho_c$, $\Delta R = \rho_s \Delta x/L_e$, and $\Delta C = \epsilon L_e \Delta x/T_c$

2.3. *Secondary Effects in* f_{max}

The formulas developed above are highly simplified and significantly underestimate the HBT f_{max}. Two significant corrections must be applied. First, the simple lumped RC model of the base-collector junction must be re-examined. Secondly, differential space-charge effects substantially reduce the collector-base capacitance under high-current conditions.

The HBT base-collector network is distributed, and is represented by the model of fig. 5. Using a small grid spacing, we have entered the resulting network into a microwave circuit simulator (HP-EESOF [23]) to calculate –without approximation– the HBT f_{max}. Alternatively, analytic expressions for f_{max} can be developed from hand analysis of the distributed network of fig. 5. Among these is the model of Vaidyanathan and Pulfrey [24], which provides good physical insight. The model of reference [24] is derived for a triple-mesa HBT; the authors of [25] have recently generalized the model to the case of transferred-substrate and lateral-etched-undercut collector [30] HBTs. We describe the Vaidyanathan / Pulfrey model below, and examine its predicted performance for HBTs with submicron emitter and collector junction widths.

Referring to fig. 5, define three capacitances. $C_{cb,e} = \epsilon L_e W_e/T_c$ is the capacitance of the collector junction lying under the emitter. $C_{cb,gap} = 2\epsilon L_e W_{eb}/T_c$ is the capacitance of the collector junction lying under the gap between the emitter and the base contact. $C_{cb,ext} = 2\epsilon L_e W_{cb}/T_c$ is the capacitance of the collector lying under the base Ohmic contacts. Components of the base resistance are as defined in eqn. 14.

The collector-base capacitance under the emitter stripe $C_{cb,e}$ is charged through a resistance $(R_{b,cont}+R_{gap}+R_{spread})$. The collector-base capacitance under the gap

between the emitter and the base Ohmic contacts is charged through a resistance $(R_{b,cont} + R_{gap}/2)$.

The charging time constant associated with the collector-base junction capacitance $C_{cb,ext}$ lying under the base Ohmics requires more detailed scrutiny. $C_{cb,ext}$ can be charged by currents passing vertically through the base Ohmic contact above it; this path has a resistance $R_{b,cont,1} = \rho_c/2L_eW_{cb}$. Alternatively, $C_{cb,ext}$ can be charged by currents passing laterally from the base contact region lying outside the perimeter of the collector contact; this path has a resistance $R_{b,cont,0} = (\rho_s\rho_c)^{1/2}\coth((W_b - W_{bc})/L_{\text{contact}})$, where $L_{\text{contact}} = (\rho_c/\rho_s)^{1/2}$ is the base Ohmic contact transfer length.

In the limit of zero collector series resistance, Vaidyanathan and Pulfrey's model, [24], [25] reduces to

$$f_{max} = \sqrt{\frac{f'_\tau}{8\pi\tau_{cb}}}, \tag{17}$$

where

$$\frac{1}{2\pi f'_\tau} = \tau_b + \tau_c + \frac{kT}{qI_c}\,(C_{je} + C_{cb}), \tag{18}$$

and

$$\begin{aligned} \tau_{cb} &= C_{cb,e}\,(R_{b,cont} + R_{gap} + R_{spread}) \\ &+ C_{cb,gap}\,(R_{b,cont} + R_{gap}/2) \\ &+ (R_{b,cont,0} \| R_{b,cont,1})\,C_{cb,ext} \end{aligned} \tag{19}$$

Examining figure 5, the external collector capacitance $C_{cb,ext}$ is not charged through the resistances R_{gap} and R_{spread}. It is pessimistic to calculate f_{max} as $(f_\tau/8\pi R_{bb}C_{cb})^{1/2}$ in which the collector-base time constant includes the full collector-base capacitance. As indicated by Vaidyanathan and Pulfrey's model (eqn. 17), the external collector capacitance $C_{cb,ext}$ is in fact charged through a smaller associated resistance $(R_{b,cont,0} \| R_{b,cont,1})$. This model shows extremely good agreement with finite-element analysis (fig. 6).

Figure 7 compares the f_{max} of mesa and transferred-substrate HBTs, computed using the finite-element model. For the transferred-substrate device, f_{max} increases rapidly with deep submicron scaling. Experimentally, we observe a more rapid variation of f_{max} with collector width than is shown in fig. 6, and fig. 7 predicts a higher f_{max} than is experimentally observed for mesa HBTs. Series resistance in the base metallization and collector series resistance [24] (not modeled above, and not present in Schottky-collector transferred-substrate HBTs) are possible explanations for the discrepancy.

At high collector current densities, differential space-charge effects in the collector space-charge region result in C_{cb} smaller than $\epsilon A_c/T_c$, and increase the HBT f_{max}. The effect was predicted by Camnitz and Moll [27], and first experimentally observed by Betser and Ritter [26]. Similar effects have been observed in MESFETs

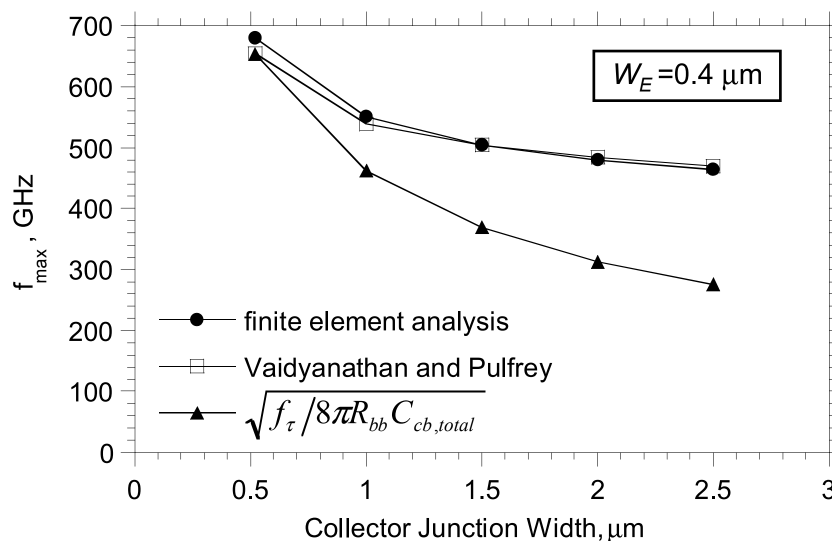


Figure 6: Comparison of f_{max} computed from a finite element model with Vaidyanathan and Pulfrey's model (Eqn.. 17) and a model using the total collector junction capacitance (Eqn.. 15). Except for W_c, the modeled HBT is that of figure 19, and has $W_e = 0.4\ \mu$m.

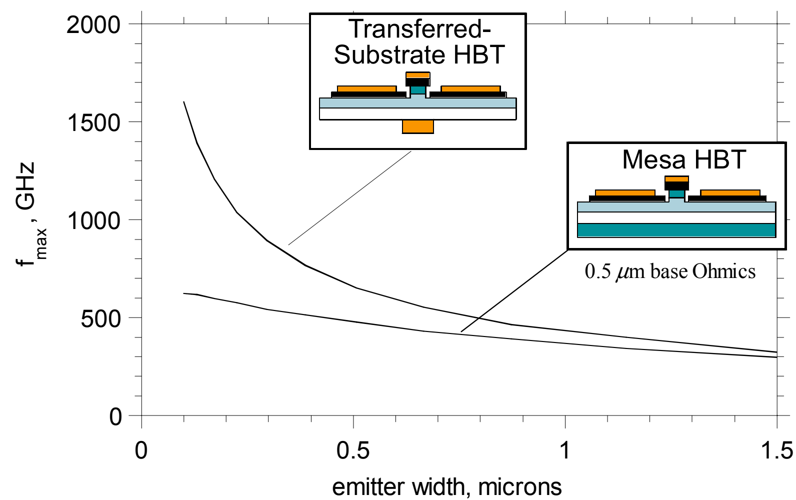


Figure 7: Lithographic scaling of transferred-substrate and mesa HBTs. f_{max} is calculated using fig. 5's finite-element model of the collector-base junction. Except for W_c and W_e, the HBT parameters are taken from the device of fig. 19. Current density and epitaxial layer thicknesses are held constant, resulting in constant f_τ

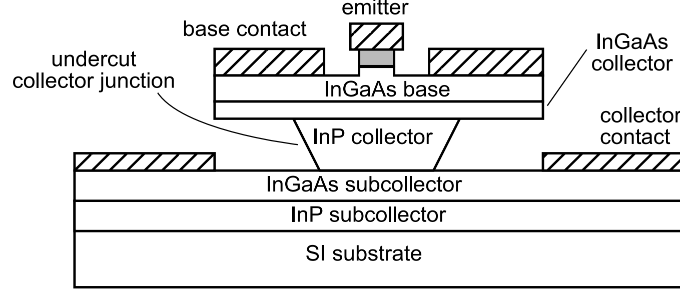


Figure 8: HBT with an undercut collector-base junction formed by selective etching of the InP collector in HCl

[28]. In III-V materials at high fields, electron velocity $v(\mathcal{E})$ decreases with increasing electric field. To a first approximation, $1/v(\mathcal{E}) \simeq \kappa_0 + \kappa_1 \mathcal{E}$. Modulating the collector voltage V_{cb} modulates the collector transit time τ_c (eqn. 3), and partially modulates the space-charge in the collector drift region. This modulated space-charge partially screens the base from modulations in the collector applied field, and $C_{cb,e}$ is reduced to

$$\begin{aligned} C_{cb,e} &= \epsilon A_e/T_c - I_c \frac{d\tau_c}{dV_{cb}} \\ &= \frac{\epsilon A_e}{T_c} - \frac{\kappa_1 J_c A_e}{2} \left[1 - \frac{\kappa_1 J_c T_c}{6\epsilon} \right], \end{aligned} \tag{20}$$

The quadratic dependence upon J_c results from internal collector field redistribution in the presence of the collector space-charge [27]. Current spreads laterally during transport through the collector, flowing through a region of width $\sim (W_e + T_c)$. The differential space charge effect strongly reduces the collector junction capacitance in regions below and adjacent to the emitter stripe. It thus has the strongest impact upon f_{max} in devices with minimal excess collector capacitance. Experimental data confirming C_{cb} cancellation will be shown in section 4.2. Capacitance cancellation is not instantaneous, but instead arises after a delay proportional to τ_c; HBT power gain must therefore increase at -40 dB/decade for frequencies above $\sim 1/2\pi\tau_c$. The effect can produce a $\sim$ 2:1 increase in f_{max}, hence a large increase in the attainable gain of tuned millimeter-wave amplifiers. In contrast, in digital circuits (section 3.1), many delay terms are significant, and a 2:1 reduction in C_{cbi} would produce only a $\sim$ 12% decrease in gate delay.

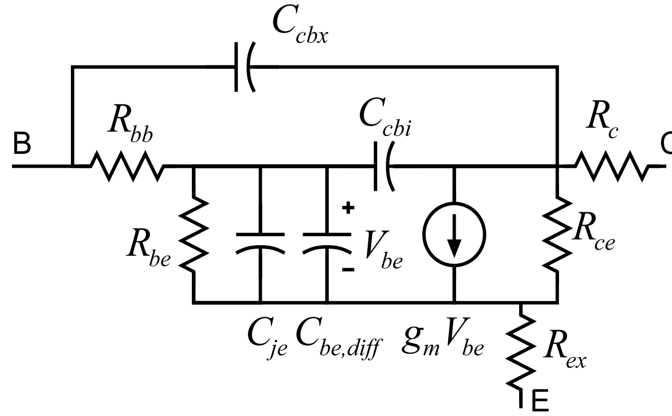


Figure 9: Hybrid-π small-signal HBT equivalent circuit. $C_{be,\text{diff}} = g_m(\tau_b + \tau_c)$. The element C_{cbi} does not represent capacitance of that fraction of the collector junction lying under the emitter, but is instead a parameter adjusted to obtain the correct f_{max} .

2.4. *HBT equivalent circuit model*

The HBT base-collector network is distributed, and accurate expressions for f_{max} are complex. Computer simulation of complex circuits requires a compact device model. Under small-signal operation, the Gummel-Poon model used in SPICE reduces to the simple hybrid-π model of figure 9. For this model, $f_{max} = (f_\tau/8\pi R_{bb}C_{cbi})^{1/2}$. It should be emphasized that C_{cbi} corresponds to no particular physical area in the collector-base junction. Specifically C_{cbi} is not equal $C_{cb,e}$, the capacitance of that fraction of the collector junction which lies under the emitter. Instead, in this model R_{bb} is given by eqn. 14, $(C_{cbx} + C_{cbi}) = \epsilon A_c/T_c$, and the intrinsic collector-base capacitance is set to $C_{cbi} = \tau_{cb}/R_{bb}$, where τ_{cb} is given by eqn. 19. Thus C_{cbi} is defined to be so that the simplified model predicts the correct device f_{max}. To correctly model common-base and emitter-follower input impedance at $f \simeq f_\tau$, the transconductance element must have an associated delay of $\sim (\tau_c + \zeta\tau_b)$, where the factor $\zeta \simeq$0.1–0.2 is dependent upon the degree of base bandgap grading.

2.5. *High f_{max} HBT designs*

To obtain simultaneous high values of f_τ and f_{max} the emitter and collector stripe widths must both be scaled. The substrate transfer process is an extremely aggressive method of reducing the parasitic extrinsic collector-base junctions, and requires a substantial departure from typical fabrication processes. There are alternatives requiring less radical processing. With GaAs/AlGaAs HBTs [29] deep proton implantation can reduce the extrinsic collector capacitance. The extrinsic collector junction can be undercut using selective wet chemical etches (fig. 8) [30], [31]. Collector capacitance under the base contact pad can be reduced using dielectric spacer

layers [32]. Alternatively, R_{bb} can be reduced by regrowing, prior to base contact deposition, thick extrinsic P+ contact regions on the exposed base surface [33], [34]. Finally, low $R_{bb}C_{cbi}$ can be obtained in mesa HBTs by reducing the size of the base Ohmic contacts. Using a CBr_4 doping source, we have grown by MBE InGaAs base layers with $> 10^{20}/\text{cm}^3$ carbon (P-type) doping. At such doping levels, ρ_c and hence the transfer length $L_{\text{contact}} = (\rho_c/\rho_s)^{1/2}$ are greatly reduced. The width of the base Ohmic contacts can be accordingly reduced.

3. HBT Digital Integrated Circuits

f_τ and f_{max} of scaled InP-based HBTs are significantly higher than Si/SiGe HBTs. Consequently, tuned and broadband amplifiers using InP-based HBTs show substantially higher bandwidths than those implemented in Si/SiGe [36], [37], [38]. Yet, in digital circuits the 2 competing technologies have held a rough parity for the past 3-4 years. Since analog/digital mixed-signal ICs (fiber optic transmission ICs, ADCs, DACs) are major HBT applications, we must examine in detail the relationship between logic gate delay and HBT design and scaling. The reader is also referred to gate delay analyes by Sano *et. al.* [49], and Enoki *et. al.* [50]. General methods of digital circuit delay analysis are discussed in Hodges and Jackson [51].

We compute below, as a function of HBT parameters, the maximum clock rate of an ECL master-slave (M/S) latch. M/S latches serve as timing control elements in digital ICs, as latched comparators in ADCs, and as decision circuits in fiber optic receivers. To benchmark their maximum clock frequency, M/S latches are configured as 2:1 *static* frequency dividers. It is important to distinguish between the maximum clock frequency of M/S latches configured as static dividers with that of dynamic 2:1 frequency dividers, which operate significantly higher clock frequencies, but have more restricted applications.

3.1. *Digital delay analysis*

A schematic diagram of an ECL MS latch is shown in fig. 10. The master latch has input stage Q1–4 and latch Q5–8, while the slave latch has input stage Q13-16 and latch Q17-20. The clock current is steered by Q9-12 and Q21-24. In our designs, signals between gates are routed on the collector nodes, using 100 Ω transmission lines terminated at sending and receiving ends in 100 Ω.

A logic voltage swing ΔV_L must be specified. Gate delay will vary with ΔV_L, but a minimum ΔV_L is necessary for adequate DC noise margin hence proper logic operation. In order for the differential pairs Q3–4 to properly steer the current of Q9, the difference in the *internal* V_{be} of the two transistors should be several times kT/q. As a first assumption, we set $\Delta V_{be,int} \sim 6kT/q$; this results in a e^6:1 ratio between the currents in the on and off states. In the presence of parasitic emitter resistance R_{ex}, the logic swing required for at least an e^6:1 current switching ratio is to

$$\Delta V_L \geq 6kT/q + I_0 R_{ex} = 6kT/q + J_0 \rho_e \ , \quad (21)$$

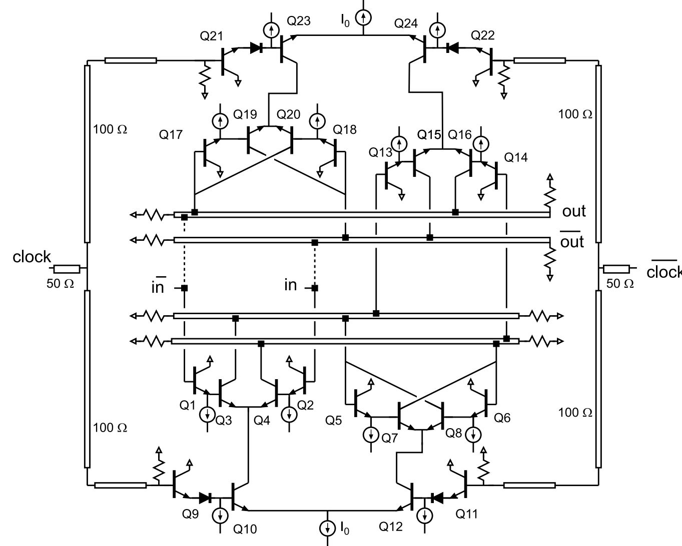


Figure 10: ECL master-slave flip-flop. The current sources are implemented with current mirrors. Except where marked, all resistors and transmission-line impedances are 100 Ω. Dotted lines indicate connections for a static 2:1 frequency divider.

where I_0 is the switched current, J_0 the emitter current density and ρ_e the emitter resistance normalized to a unit emitter junction area.

We compute, approximately, the gate delay using the charge control method, adding the charging times of each node associated with the signal path. The node charging time from the initial state to the (50%) switching point is $\Delta t \simeq \Delta Q/2I$, where ΔQ is the switched charge, and I the charging current [51], [52]. This is equivalent to analysis of a linearized version of the digital circuit, in which node impedances are modeled by $R = \Delta V/\Delta I$, $C = \Delta Q/\Delta V$, and $g_m = \Delta I_c/\Delta V_{be}$ [51]. Gain effects varying to second order in $(j\omega)$ in the circuit transfer function are neglected; this simplifying assumption introduces significant error by ignoring the effect of emitter-follower ringing.

We assume a current density J_0 in the upper-level current-switch HBTs and a current density $J_0/2$ for all emitter followers and for the lower-level clock-steering current-switch HBTs. The upper-level differential current-switch transistors have emitter areas $A_{e,cs}$, the lower-level (clock switching) current-switch transistors have emitter areas $2 \cdot A_{e,cs}$, and the emitter followers have emitter areas $A_{e,ef}$. The currents that flow in these devices are therefore $I_0 = J_0 A_{e,cs}$, $I_0 = (J_0/2)(2A_{e,cs})$, and $I_{0,E} = (J_0/2)A_{e,ef}$ respectively. The base-emitter voltage in the on-stage is denoted as $V_{be,on}$. For simplicity, we assume a digital voltage swing ΔV_L at all upper-level collector nodes, although it is known that decreased MS latch delay can be obtained by using smaller switched currents (hence smaller ΔV_L) during

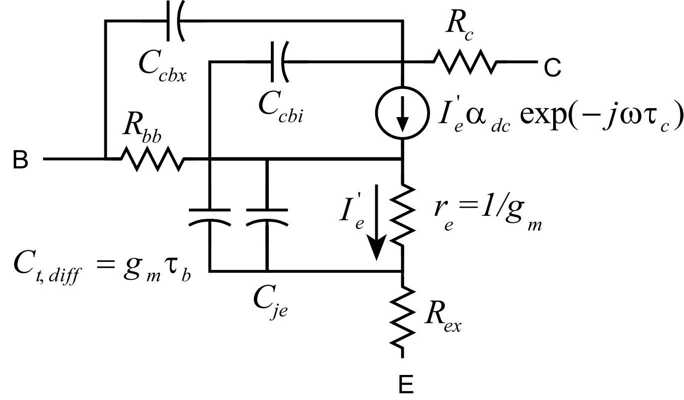


Figure 11: Simplified HBT common-base (T) equivalent circuit model used in the logic delay analysis. Note that τ_b is modeled as a diffusion capacitance while τ_c is modeled as a transport delay.

operation of the positive-feedback latch).

The large-signal base-emitter depletion capacitance is defined as $C_{je} = c_{je} * A_e$, where the average capacitance per unit emitter area is

$$c_{je} \equiv \frac{\Delta Q}{\Delta V} = \frac{1}{\Delta V}\int_{V_{be,on}-\Delta V}^{V_{be,on}} c_{je}(V)dV \ . \tag{22}$$

The base-collector capacitance $C_{cb} = c_{cb} * A_e$ is taken as proportional to the *emitter* area (thus assuming a fixed emitter-collector area ratio). The collector-base junction is fully depleted and operates at current densities below that causing base pushout. In the hybrid-π model (fig. 9), the large-signal base-emitter diffusion capacitance is $C_{be,\text{diff}} = I_0(\tau_b + \tau_c)/\Delta V$. For common-base switching paths the T-model (fig. 11) is employed; for that model, the base-emitter junction has a *small-signal* diffusion capacitance $C_{\text{t,diff}} = g_m\tau_b$, while under large signal drive the capacitance becomes $\tau_b I_c/\Delta V_L$.

Assume that the bases of Q1 and Q17 are at a logic high while the bases of Q2 and Q18 are at a logic low. At $t = 0$ the clock rises from low to high. The clock differential pair Q10/Q12 changes state, establishing after a propagation delay a collector current I_0 in Q10 . The current I_0 then charges the capacitances at the emitter node of Q3, driving the node negative until Q3 turns on (fig. 12a). After a propagation delay through Q3, I_0 is established as a collector current for Q3. The current I_0 then charges the capacitances at the collector node of Q3, driving the node negative with an charging time resulting from the node capacitances (fig. 12b). Finally (fig. 12c), the emitter followers (Q5,6,13,14) charge/discharge the base-emitter junctions of the master-stage latch current-steering pair (Q7,8) and the slave-stage input current-steering pair (Q15,16), with delay arising both from the emitter-followers and from the (Q7,8,15,16) base-emitter junction charging through R_{bb} Once this sequence is complete, the clock can change states and the sequence

repeats itself in the slave stage. Note that the delays associated with the clock differential pairs occur both in the master and in the slave and therefore, to a first approximation, do not affect the maximum clock frequency.

We first calculate the switching delay at the emitter of Q3 (fig 12,a). Q10 is turned on at $t = 0$. Q10 and Q3 are in series, and have equal on-state and off-state emitter currents. The logic voltage swings at the base-emitter junctions of Q10 and Q3 must therefore be identical, with $V_{be,on} - V_{be,off} = \Delta V_L = I_0 R_L$ for both transistors. Further, note that over the logic transition C_{je3} sees a voltage swing of $\Delta V_L - I_0 R_{ex3} = I_0(R_L - R_{ex3})$, while C_{cb10} sees a voltage swing of $2 \cdot \Delta V_L$. There is interconnect capacitance C_{m1} at the node; further Q3 and Q4 have substrate capacitances $C_{s,3}$ and $C_{s,4}$. The node charging time is

$$
\begin{aligned}
T_{\mathrm{Q3emitter}} &= \Delta V_L \left(\frac{C_{s3} + C_{s4} + C_{m1} + 2C_{cb10} + C_{je4}}{2I_0} \right) \\
&\quad + (\Delta V_L - I_0 R_{ex3}) \left(\frac{C_{je3}}{2I_0} \right) + \tau_b \\
T_{\mathrm{Q3emitter}} &= \Delta V_L \left(\frac{2c_s + C_{m1}/A_{e,cs} + 4c_{cb} + c_{je}}{2J_0} \right) \\
&\quad + (\Delta V_L - J_0 \rho_e) \left(\frac{c_{je}}{2J_0} \right) + \tau_b \ , \qquad (23)
\end{aligned}
$$

where the latter form is written using currents, capacitances and resistances normalized to a unit HBT emitter junction area ($\rho_e = R_{ex} A_e$, $r_{bb} = R_{bb} A_e$).

Second, we calculate the delay between the emitter and collector of Q3 (fig 12,b). Q3 operates in common-base mode, and its g_m element has delay τ_c. Capacitances C_{cb8} and C_{cb3} undergo a $2\Delta V_L$ voltage swing; other capacitances undergo a swing of ΔV_L. Adding collector transmission-line bus delay τ_{bus} (fig. 10), the node delay at the collector of Q3 is

$$
\begin{aligned}
T_{\mathrm{Q3coll}} &= \tau_c + \tau_{bus} \\
&+ \Delta V_L \left(\frac{2C_{cb8} + 2C_{cb3} + C_{cb5} + C_{cb13}}{2I_0} \right) \\
T_{\mathrm{Q3coll}} &= \tau_c + \tau_{bus} \\
&+ \Delta V_L \left(\frac{(4 + 2A_{e,ef}/A_{e,cs})c_{cb}}{2J_0} \right) \ . \qquad (24)
\end{aligned}
$$

Finally, we calculate the delay between the voltage transition at the collector of Q3 and the base (internal to R_{bb}) of Q15 (fig 12,c). In the figure, the emitter follower is represented by a T-model and the current-steering device by a partial hybrid-π model. The emitter followers Q6 and Q14, simultaneously undergoing a negative-going transition, are explicitly assumed to remain on during the switching event; this stipulates a minimum $(A_{e,ef}/A_{e,cs})$ area ratio, and always-on operation must be verified during design. The bias current in Q7,8 and Q15,16 must be

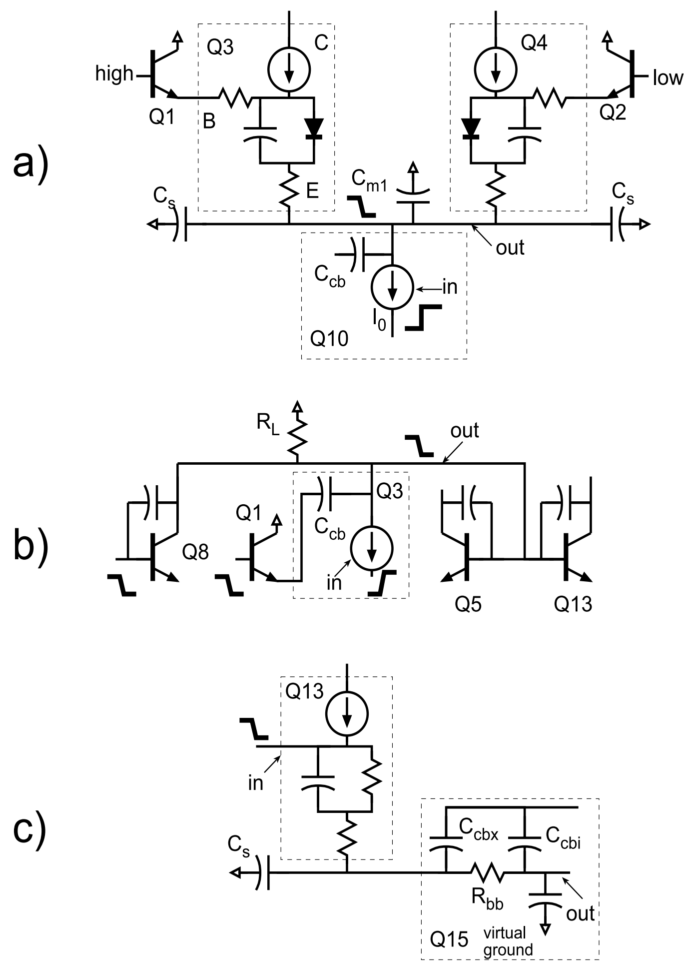


Figure 12: Equivalent circuit of the nodes in the signal path for calculating the M/S latch delay. Charging of the emitter node of Q3 (a). Charging of the collector node of Q3 (b). Charging of the base of the switching transistor Q15 (c).

Table 1: Delay coefficients a_{ij}, found by hand analysis, assuming gate delay of form $T_{gate} = 1/2f_{clock} = \Sigma a_{ij} r_i c_j$.

..	c_{je}	c_{cbx}	c_{cbi}	c_s	$\frac{C_{m1}}{A_{e,cs}}$	$\frac{\tau_f J_0}{\Delta V_L}$	$\frac{\tau_{bus} J_0}{\Delta V_L}$
$\Delta V_L/J_0$	1	6	6	1	0.5	1	1
kT/qJ_0	0.5	1	1	0.5	0	0.5	0
ρ_e	-0.25	0.5	0.5	0.5	0	0.5	0
r_{bb}	0.5	0	1	0	0	0.5	0

Table 2: Delay coefficients a_{ij}, found by SPICE, assuming gate delay of form $T_{gate} = 1/2f_{clock} = \Sigma a_{ij} r_i c_j$.

..	c_{je}	c_{cbx}	c_{cbi}	c_s	$\frac{C_{m1}}{A_{e,cs}}$	$\frac{\tau_f J_0}{\Delta V_L}$	$\frac{\tau_{bus} J_0}{\Delta V_L}$
$\Delta V_L/J_0$	0.8	4.3	4.3	1.9	0.7	1.6	1
ρ_e	-0.1	15	15	2.1	6	0.2	0
r_{bb}	0.7	2.2	5.2	0	0.1	0.7	0

examined carefully for this calculation. Under maximum-clock-rate operation, the base voltages of Q7 and Q15 change states only slightly before an emitter current is established in these transistors through the turn-on of Q12 and Q24 at the next clock high-low transition. A hand calculation here can only be approximate; we will take the emitter current of Q7 and Q15 to be I_0 during the base voltage transition. The delay is

$$\begin{aligned} T_{\mathrm{Q13/15}} &= (1/2)(kT/qI_{0,E} + R_{ex13}) \\ &\times (C_s + 2C_{cb15} + C_{je15} + \tau_f I_0/\Delta V_L) \\ &+ (1/2)R_{bb13}\,(C_{je15} + 2C_{cbi15} + \tau_f I_0/\Delta V_L) \\ T_{\mathrm{Q13/15}} &= (1/2)(2kT/qJ_0 + \rho_e)(A_{e,cs}/A_{e,ef}) \\ &\times (C_s/A_{e,cs} + 2c_{cb} + c_{je} + \tau_f J_0/\Delta V_L) \\ &+ (1/2)r_{bb}\,(c_{je} + 2c_{cbi} + \tau_f J_0/\Delta V_L)\ . \end{aligned} \tag{25}$$

The total gate delay is then

$$1/2f_{clock} = T_{total} = T_{\mathrm{Q3emitter}} + T_{\mathrm{Q3coll}} + T_{\mathrm{Q13/15}}\ , \tag{26}$$

and the maximum clock frequency f_{clock} is determined.

Both hand analysis and SPICE simulations indicate that f_{clock} exhibits a broad maximum as a function of the ratio of emitter follower to current switch emitter areas, with $A_{e,ef}/A_{e,cs} \simeq 2$ being optimum. We subsequently assume this ratio. Results of the hand calculations are summarized in table 1. Note that because the logic swing ΔV_L is large in comparison with kT/q, terms in kT/qJ_0 will be substantially smaller than terms in $\Delta V_L/J_0$ To somewhat simplify the tabulations,

Table 3: Delay components, found by SPICE, as a fraction of a total 4.9 ps latch delay, for the HBT of fig. 28. All emitter-followers and the lower-level current switch devices operate at 10^5 A/cm^3 current density, with the upper-level current switches operating at $2 \cdot 10^5$ A/cm^3. The logic swing is $\Delta V_L = 200$ mV

..	c_{je}	c_{cbx}	c_{cbi}	c_s	$\frac{C_{m1}}{A_{e,cs}}$	$\frac{\tau_f J_0}{\Delta V_L}$	$\frac{\tau_{bus} J_0}{\Delta V_L}$	total
$\Delta V_L/J_0$	6%	7%	5%	4%	1%	11%	10%	44%
ρ_e	0%	9%	7%	2%	3%	1%	0%	21%
r_{bb}	12%	1%	12%	0%	0%	10%	0%	35%
total	18%	16%	23%	6%	5%	22%	10%	100%

terms in kT/qJ_0 were combined in subsequent tables with those in $\Delta V_L/J_0$. The delays τ_f and τ_{bus} are written as effective capacitances $(\tau J_0/\Delta V_L)$, in order to represent the delay in the form $T_{gate} = 1/2f_{clock} = \Sigma a_{ij} r_i c_j$.

A large set of SPICE simulations were performed of M/S latch maximum toggle rate, using circuit models of μm-scale and submicron-emitter transferred-substrate HBTs. To the extent that the gate delay can be approximated by first-order delay terms, $T_{gate} = 1/2f_{clock} = \Sigma a_{ij} r_i c_j$, the delay coefficients a_{ij} can be found by varying the HBT model parameters in the simulations. Tables 2 and 3 show the results of this analysis for the HBT of fig. 28. Given the many simplifications involved in the hand analysis, the correlation between hand analysis and simulation is reasonable, except that in hand analysis smaller coefficients are found for the terms $r_{ex}c_{cbx}$, $r_{ex}c_{cbi}$, $r_{bb}c_{cbx}$, and $r_{bb}c_{cbi}$, terms which in the simulations are found to collectively contribute 29% of the latch delay.

Tables 1, 2, and 3 provide important points regarding HBT design for fast logic. In modern InP-based HBTs, τ_f is relatively small, and (for present UCSB HBTs) contributes only $\sim$ 20% of total gate delay, an amount comparable to the delay contributed by c_{je}, and much smaller than the $\sim$ 40% contributed by terms associated with $(c_{cbx} + c_{cbi})$. Low base and collector transit times, hence high f_τ, is –of itself– not indicative of high speed logic operation. This is because under logic operation, the change in base+collector stored charge is

$$\Delta Q_{b,c} = \tau_f I_0 = \left(\frac{\tau_f I_0}{\Delta V_L}\right) \Delta V_L = C_{ls} \Delta V_L \ , \qquad (27)$$

while under small-signal operation

$$\delta Q_{b,c} = g_m \tau_f \delta V_{be} = \left(\frac{\tau_f I_0}{kT/q}\right) \delta V_{be} = C_{ss} \delta V_{be} \ . \qquad (28)$$

Under logic operation, the base-emitter diffusion capacitance associated with τ_f is reduced in proportion to the ratio of logic swing ΔV_L to kT/q, a ratio of typically 10:1. In contrast, under logic operation the capacitances C_{je} and C_{cb} must be provided with charge $C_{je}\Delta V_L$, and $C_{cb}\Delta V_L$.

Examining the delay components in terms of real (r_{bb}, r_{ex}) and equivalent ($\Delta V_L/J_o$) resistances through which the depletion and diffusion capacitances are charged, we are faced with a significant discrepancy between hand and computer analyses. In either analysis, $\Delta V_L/J_o$ is dominant. A key conclusion is that high current densities are essential for fast HBT logic circuits. If the HBT is operated at a current density limited by the Kirk effect (eqn. 6), then the delay terms associated with charging the collector-base capacitance,

$$\begin{aligned}\frac{C_{cb}\Delta V_L}{I_0} &= \frac{\epsilon A_c}{T_c}\cdot\frac{\Delta V_L}{J_o A_e} \\ &= \frac{A_c}{A_e}\frac{\Delta V_L}{(V_{cb}+\phi)}\frac{T_c}{4v_{sat}}\ ,\end{aligned} \tag{29}$$

are minimized through use of *thin* collector layers. Delay associated with r_{bb} is also significant. Finally, note that while simulations associate 22% of the net delay with R_{ex}, this *underestimates* its effect; because adequate noise margin demands $\Delta V_L \geq 6kT/q + J_0\rho_e$ (eq. 21), reducing delay terms associated with $\Delta V_L/J_o$ through increased current density demands simultaneous improvements in ρ_e.

3.2. *Scaling for high speed logic*

As examined in sections 2.2 and 2.3, lithographic scaling of the emitter and collector junction widths progressively increases f_{max} if the parasitic collector-base junction is eliminated. If the lithographic dimensions are scaled while holding the base and collector epitaxial layer thicknesses constant, f_{max} increases rapidly while f_τ remains relatively constant. While such a device will produce gain at very high frequencies in reactively-tuned MIMICs, broadband analog circuits require simultaneous high values of f_τ and f_{max}.

In analyzing HBT logic speed (section 3.1), it is found that $\sim$ 10–15 equivalent RC delay terms are significant. In order to improve logic speed, all significant HBT capacitances and transit delays must be reduced. We now examine the scaling of HBT parameters required to increase bandwidth by a factor of $\gamma : 1$, using simplified expressions for HBT parameters in order to more clearly show the dominant trends. To ensure that bandwidth increases by $\gamma : 1$ for *all* circuits, digital and analog, using the scaled HBT, all transit times and all capacitances in figure 9 must be reduced by $\gamma : 1$, while maintaining constant all resistances, the transconductance, and the collector bias current I_c. Explicitly, $I_c \propto \gamma^0$ and $g_m \propto \gamma^0$.

The base-emitter diffusion capacitance is

$$C_{be,\text{diff}} = g_m(\tau_b + \tau_c) = (qI_c/kT)(\kappa_2 T_b^2 + \kappa_3 T_c)\ , \tag{30}$$

Here the terms κ_i represent parameters which do not change with scaling. To obtain $C_{be,\text{diff}} \propto \gamma^{-1}$ with fixed I_c, we must set $\tau_b \propto \gamma^{-1}$ and $\tau_c \propto \gamma^{-1}$. This requires $T_b \propto \gamma^{-1/2}$ and $T_c \propto \gamma^{-1}$.

An immediately apparent limit to collector scaling is loss of collector breakdown voltage. An AlInAs/GaInAs HBT with a 0.2 μm InGaAs collector thickness exhibits $V_{br,ceo} = 1.5$ V at 10^5 A/cm^2 bias. Semiconductors with higher products

($\mathcal{E}_{max} v_{sat}$) of breakdown field and electron velocity mitigate this limit; HBTs with InP collectors [35] exhibit τ_c comparable to devices with InGaAs collectors but have $\sim$ 5:1 increased breakdown. Regardless of the collector thickness, impact ionization cannot occur for V_{ce} less than the bandgap of the collector semiconductor. Further, unless the collector bandgap is small or the collector much thinner than 1000 Å, Zener tunneling currents will also be small for bias voltages below the collector bandgap energy. Even with 1000-Å collector layers, an InP/GaInAs/InP DHBT will exhibit $V_{br,ceo} > 1.2$ V , sufficient for current-mode logic. While important in power amplifiers and in mixed-signal (medium-voltage) ICs, loss of breakdown voltage may not pose a serious limit to the scaling of InP-collector DHBTs for low-voltage, high-speed logic.

The capacitance C_{je} is given by

$$C_{je} = C_{je1} + C_{je2} = \kappa_4 L_e W_e / T_{eb} + \kappa_5 T_{eb} T_b I_c \ . \tag{31}$$

Analysis of the partitioning of C_{cb} between $C_{cb,x}$ and C_{cbi} is complex (section 2.3), and in this section we therefore restrict the analysis to HBTs in which $C_{cb,x}$ is zero ($L_c \simeq L_e$ and $W_c \simeq W_e$) and $C_{cbi} = C_{cb}$. Such HBTs include transferred-substrate (figure 4) and undercut-mesa devices (figure 8), and mesa devices having very high base doping and hence requiring only a very small base Ohmic contact width. C_{cb} then scales as

$$C_{cb} = \epsilon W_c L_c / T_c \simeq \epsilon W_e L_e / T_c \ . \tag{32}$$

Because $T_c \propto \gamma^{-1}$, to obtain $C_{cb} \propto \gamma^{-1}$ we must set $W_c L_c \propto \gamma^{-2}$ and hence $W_e L_e \propto \gamma^{-2}$.

The base resistance R_{bb} is the sum of the terms (eqn. 14) $R_{b,cont}$, R_{gap} and R_{spread}. Correct scaling of C_{cb} requires that $W_e L_e \propto \gamma^{-2}$. It is desired that R_{bb} vary negligibly with scaling; we show here that this is obtained by setting $W_e \propto W_c \propto \gamma^{-2}$ and $L_e \simeq L_c \propto \gamma^0$. The base contact resistance term $R_{b,cont} = \kappa_6 \rho_c^{1/2} / L_e T_e$ is proportional to $\gamma^{1/4}$, while $R_{spread} = \kappa_7 W_e / L_e T_{eb} \propto \gamma^{-3/2}$. If we scale $W_{eb} \propto \gamma^{-1}$, then $R_{gap} = \kappa_8 W_{eb} / L_e T_b \propto \gamma^{-1/2}$. While the contact resistance term $R_{b,cont}$, the dominant term in R_{bb} for submicron devices, increases ($\propto \gamma^{1/4}$) slowly with scaling, the rapid decrease in R_{gap} and R_{spread} results in a total R_{bb} showing only a very slow increase with scaling.

To obtain $C_{je2} \propto \gamma^{-1}$ we must set $T_{eb} \propto \gamma^{-1/2}$. This results in $C_{je1} \propto \gamma^{-3/2}$, improving more rapidly than required for a $\gamma : 1$ scaling in transistor bandwidth.

The collector series resistance R_c is zero in transferred-substrate HBTs using Schottky collector contacts. In undercut-mesa devices, R_c has a similar geometric dependence as R_{bb}, and also varies only minimally with scaling.

Scaling thus requires that the emitter and collector stripe widths W_e and W_c be proportional to γ^{-2}, and that the emitter and collector stripe lengths L_e and L_c be independent of scaling. Because the collector current is constant ($I_c \propto \gamma^0$), the emitter current density increases quadratically with the desired improvement in transistor bandwidth ($J_e \propto \gamma^2$), as does the transistor's operating power density

$(P/A_e = J_e V_{ce} \propto \gamma^2)$. Limits to bias current density imposed by reliability concerns and dissipated power density are thus major impediments to scaling for high bandwidth.

The emitter resistance $R_{ex} = \rho_e / W_e L_e$ presents a major impediment to scaling. With $W_e L_e \propto \gamma^{-2}$, in order to maintain the desired constant R_{ex} the aggregate emitter resistivity $\rho_e \propto \gamma^{-2}$ must improve in proportion to the square of the intended improvement in HBT bandwidth. This will require substantial increases in emitter doping over those now typically used in HBTs, and use of low-resistivity (e.g. InAs) semiconductor contact layers.

The collector-emitter resistance is $R_{ce} = V_A / I_c$, where the Early voltage is $V_A = qN_a T_b T_c / \epsilon$ and N_a is the base doping. From these relationships $R_{ce} \propto \gamma^{-3/2}$, and does not scale as desired. Fortunately, for an HBT with $T_b = 300$ Å, $T_c = 0.2\mu$m, and $N_A = 5 \cdot 10^{19}/\text{cm}^3$ (a device with 275 GHz f_τ), $V_A \sim 500$ V. A $\gamma = 10 : 1$ scaling for a target 2750 GHz f_τ would still result in $V_A = 16$ Volts, which is acceptably large. In HBTs, degradation of R_{ce} through base-width modulation is not a significant impediment to scaling.

In scaling the device, we have set $W_e \propto \gamma^2$ and $L_e \propto \gamma^0$. If all other widths and lengths in the device layout are scaled in the same proportions, then the HBT area, and the area of a given circuit, are proportional to γ^{-2}. The average wire length within the circuit is proportional to the square root of the IC area, and hence is proportional to γ^{-1}. Wiring delays, whether transmission-line delays or $C_{wire}\Delta V/\Delta I$ charging times, thus also scale correctly. Because of the fixed bulk metal resistivity, interconnect parasitic series resistance does not scale correctly, increasing as γ^2.

In scaled HBTs, base current is dominated by surface recombination and by currents conducted on the surface between the base-emitter junction and the base Ohmic contact. Consequently , $I_b \propto n(T_{eb}) L_e$. Because $I_c \propto L_e W_e n(T_{eb})/T_b$, $\beta \propto W_e / T_b$. With the scaling laws above, $\beta \propto \gamma^{-3/2}$. Current gain decreases rapidly with scaling, and reduction of surface recombination and surface conduction is critical in deep submicron devices.

Finally, we reconsider scaling of the mesa HBT. For mesa HBTs, base and collector thickness, emitter and collector junction widths, emitter contact resistivity, and current density must all scale as discussed above for undercut-mesa and transferred-substrate HBTs. In particular, the base-collector junction width must still scale as γ^{-2}. For a normal triple-mesa device, this then requires that the widths W_b of the base Ohmic contacts (fig. 1) scale as γ^{-2}, while maintaining a fixed $R_{b,cont} = (\rho_s \rho_c)^{1/2}(1/2L_e)\coth(W_b/L_{\text{contact}})$. This can be accomplished by a combined reduction of both ρ_s and ρ_c, and hence a general analysis is exceedingly complex. As a limiting case, with a highly scaled HBT, W_c and must be very small, and hence W_b will be much less than L_{contact}. In this case $R_{b,cont} \simeq \rho_c / 2L_e W_b$, and hence constant $R_{b,cont}$ requires that the base Ohmic contact resistivity scale as $\rho_c \propto \gamma^{-2}$. Transferred-substrate and narrow-mesa HBTs do not require this improvement of base contact resistivity with scaling.

Table 4: Scaling laws for HBTs; required proportional change in key relevant HBT physical parameters in order to obtain a γ:1 increase in bandwidth in an arbitrary circuit. Additionally, for mesa HBTs, but not transferred-substrate or undercut-mesa devices, the base contact resistivity ρ_v must scale as γ^{-2}

parameter	symbol	scaling law
collector depletion layer thickness	T_c	γ^{-1}
base epitaxial layer thickness	T_b	$\gamma^{-1/2}$
emitter-base junction width	W_e	γ^{-2}
collector-base junction width	W_c	γ^{-2}
emitter-base depletion thickness	T_{eb}	$\gamma^{-1/2}$
emitter parasitic resistivity	$\rho_e = R_{ex}A_e$	γ^{-2}
emitter junction area	$A_e = W_e L_e$	γ^{-2}
emitter current	I_e	γ^0
emitter current density	J_e	γ^2
bias and signal voltages	V_{CE}, v_{ce}, v_{be}	γ^0
average interconnect length	L_{wire}	γ^{-1}
circuit area	–	γ^{-2}
device power density	–	γ^2
circuit power density	–	γ^2

To simultaneously increase HBT bandwidth in general circuits by $\gamma : 1$, emitter and collector junction widths must vary as γ^{-2} while maintaining constant junction lengths. Base thickness must vary as $\gamma^{-1/2}$ and collector thickness as γ^{-1}. Emitter current density and transistor and IC power density all increase in proportion to γ^2. The emitter contact structure must improve in proportion to γ^2. Power dissipation, reliability under high-current operation, required improvements in surface recombination velocity, and the required quality of the emitter Ohmic contact are the most significant impediments to scaling. These relationships are summarized in table 4.

3.3. *Design projections for* > 200 *GHz logic*

Following the design rules above, a scaling study of high speed M/S latches was pursued. Based upon measured parameters of tested HBTs, an HBT SPICE model was developed in which model elements (depletion capacitances, contact resistances, and carrier transit times) were calculated as a function of lithographic dimensions and layer thicknesses. ECL master-slave flip-flops were then simulated for maximum clock frequency. The results (table 5) start with the HBT design of fig. 28, and show progressive increases in clock rate as the emitter and collector stripe widths are reduced, base and collector layers thinned, the current density increased, and the emitter contact resistivity reduced. Thin collector layers are here required not primarily for low τ_c, but primarily so as to increase (eq. 6) the current density at base pushout, and hence decrease $C_{cb}\Delta V_L/I$ (eq. 29).

Table 5: SPICE simulation results of flip-flop clock speed as a function of transistor design. Interconnect capacitance and delay is not considered. Boldface indicates parameter changed from previous design..

Emitter		Collector				Base		Clock
width	parasitic resistance	width	thickness	current density	material	thickness	doping	
1.2	50	1.8	3000	1.0E+05	InGaAs	400	4E19 Be	115
0.7	50	1.5	3000	1.0E+05	InGaAs	400	4E19 Be	125
0.7	50	1.5	3000	1.0E+05	InGaAs	**300**	4E19 Be	128
0.7	50	**0.8**	3000	1.0E+05	InGaAs	300	4E19 Be	159
0.35	50	**0.45**	3000	1.0E+05	InGaAs	300	4E19 Be	176
0.35	50	0.45	3000	1.0E+05	InGaAs	300	**1E20 C**	182
0.35	**25**	0.45	2120	2.0E+05	InP	300	1E20 C	250
0.35	**12.5**	0.45	**1500**	**4.0E+05**	InP	300	1E20 C	285
μm	Ohm-μm^2	μm	Å	A/cm^2	--	Å	cm^{-3}	GHz

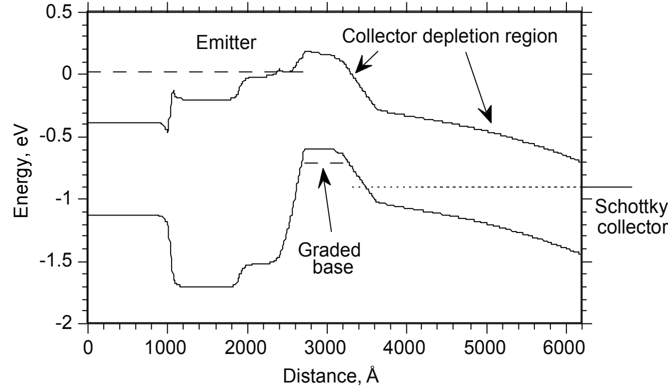


Figure 13: Band diagram, under bias, of a typical device.

4. Transferred-substrate HBTs

Wide HBT bandwidths are obtained by scaling. In scaling for high f_τ, significant limits include high power density and high current density, demands for very low emitter parasitic resistance, and the collapse of f_{max} due to the extrinsic collector-base junction. Using substrate transfer processes, this extrinsic junction can be reduced in size or eliminated. This permits either aggressive lithographic scaling *without* epitaxial scaling for greatly increased f_{max} at constant f_τ. Alternatively, if high values of both f_τ and f_{max} are sought, simultaneous lithographic and epitaxial scaling is required; with the extrinsic C_{cb} eliminated, operation at high current density and reduction of the emitter resistance are the key requirements for further scaling.

4.1. *Growth and fabrication*

The epitaxial layer structure is described by its band diagram (fig. 13). The InGaAs base is typically 300–400 Å thick, has $2kT$ bandgap grading, and is Be-doped at $5{\cdot}10^{19}/\text{cm}^3$. The InGaAs collector is 2000-3000 Å thickness. A collector N^+ pulse-doped layer placed 400 Å from the base delays the onset of base push-out at high collector current densities. Although such pulse-doped layers have been used as electron launchers [39] in GaAs-based HBTs, our experimental data shows no significant effect of the launcher upon τ_c for InGaAs-collector HBTs.

Devices typically use Schottky collector contacts [40], although HBTs with N+ subcollector layers (Ohmic-collector devices) have also been fabricated. While Ohmic-collector devices have non-zero collector series resistance, hence lower f_{max} [24], the 0.2 V barrier present in the Schottky-collector device increases the V_{ce} re-

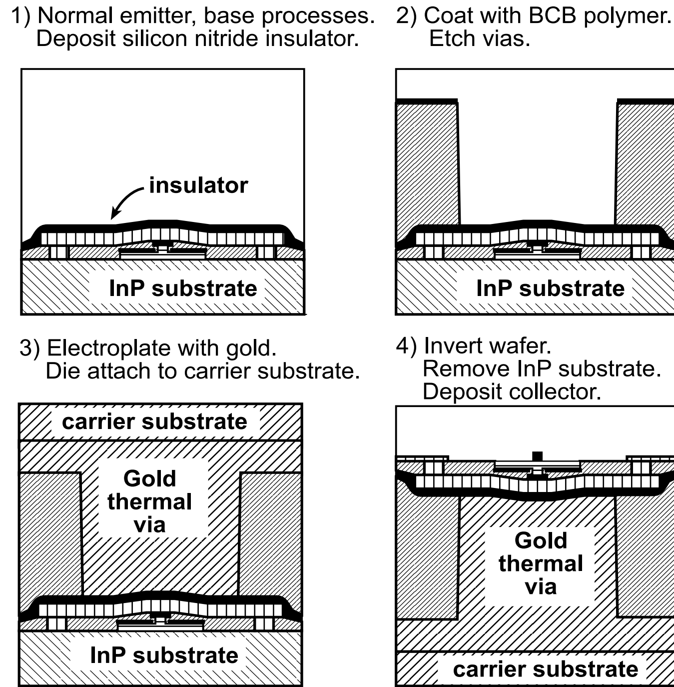


Figure 14: Transferred-substrate HBT process flow.

quired to suppress base push-out at high current densities. Ohmic-collector devices thus show higher f_{max} under the low-V_{ce} conditions associated with current-mode-logic (CML). Schottky-collector devices are used for emitter-coupled-logic (ECL), where the operating V_{ce} is higher.

Figure 14 shows the process flow. Standard fabrication processes [42] define the emitter-base junction, the base mesa, polyimide planarization, and the emitter contacts. The substrate transfer process commences with deposition of the PECVD Si_3N_4 insulator layer and the Benzocyclobutene (BCB) transmission-line dielectric (5 μm thickness). Thermal and electrical vias are etched in the BCB. The wafer is electroplated to metallize the vias and to form the ground plane. The wafer is then solder-bonded to a GaAs carrier substrate. The InP substrate is removed in HCl and Schottky collectors are deposited, completing the process. Fig. 15 shows a detailed device cross section.

For the emitter-base junction, deep submicron scaling requires tight control of lateral undercutting during the base contact recess etch. To form the emitter, reactive-ion etching in CH_4 / H_2 / Ar, monitored with a HeNe laser, first removes the N^+ GaInAs emitter contact layer. A HCl/HBr/Acetic selective wet etch then removes the AlInAs emitter, stopping on the AlInAs/GaInAs emitter-base grade. By etching at 10° C, the etch rate is slowed, and a controlled emitter undercut is formed. The undercut both narrows the emitter and serves (as normal) to define

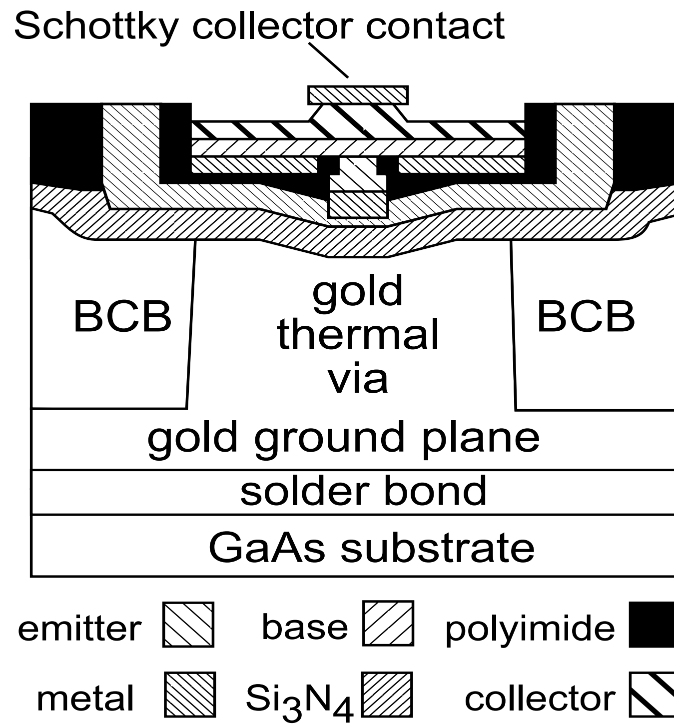


Figure 15: Schematic cross-section of a transferred-substrate HBT

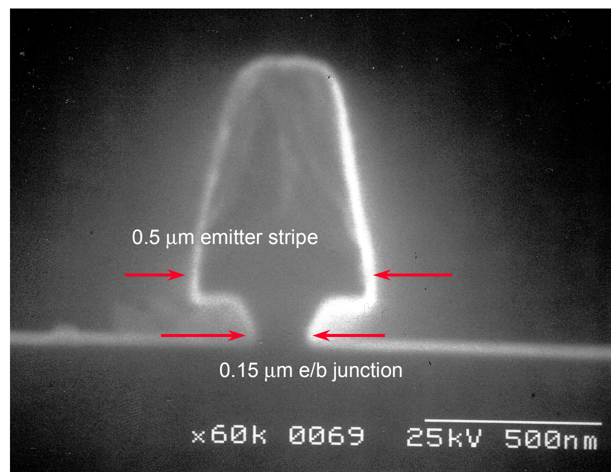


Figure 16: Cross-section of emitter-base junction. The 0.5 μm emitter metal was defined with a projection lithography system.

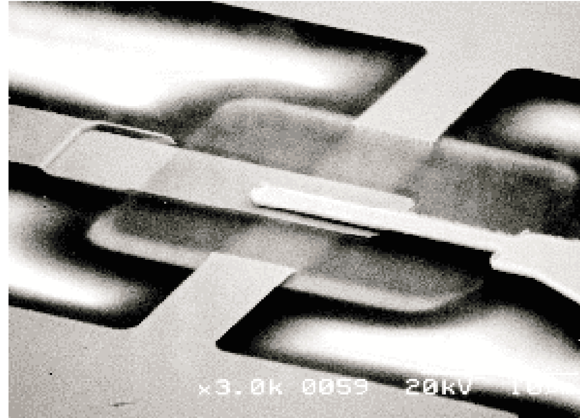


Figure 17: Transferred-substrate HBT defined by contact lithography

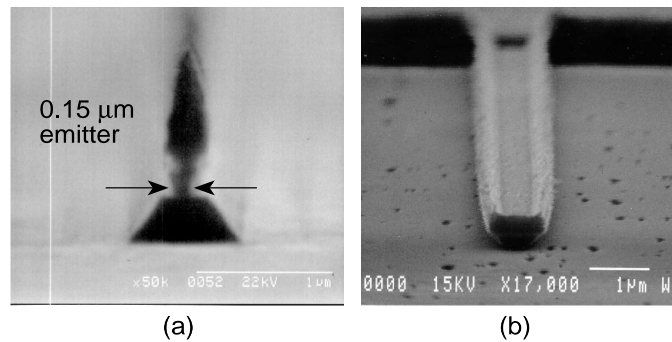


Figure 18: E-beam HBT: test structure with 0.15 μm emitter-base junction (a), and 0.4 μm Schottky collector stripe (b)

the liftoff edge in the self-aligned base contact deposition. A timed nonselective wet Citric/H_3PO_4/H_2O_2 etch then removes the base- emitter grade. Etch selectivity in both the RIE and HCl/HBr/Acetic etches aids in etch-depth control, and we are able to reproducibly etch ~100 Å into the base without use of surface contact resistance probing as a process monitor. Figure 16 shows the cross-section of a 0.15-μm emitter-base junction.

In defining submicron collector-base junctions, use of the Schottky-collector contact eliminates the need for an etch of similar precision through an N^+ collector Ohmic contact layer. The collector junction is defined by the stripe width of the deposited metal. Subsequent to collector deposition, a self-aligned wet etch of ~1000 Å depth removes the collector junction sidewalls (eliminating fringing fields) and reduces the collector junction width by ~2000 Å. The step, intended to reduce C_{cb}, generally provides a greater increase f_{max} than would be expected from the observed reduction in collector junction width.

Given the unusual features of the substrate transfer process, IC yield is a significant concern. The transistors and ICs reported here have all been developed by a team whose average size –over time– is approximately 12 Ph.D. students, working in a university cleanroom, and responsible for all aspects of technology, including crystal growth, processing, IC design, and testing. It is therefore difficult to separate yield difficulties inherent to the substrate transfer process with yield difficulties associated with limited manpower available to address process control, and the limited quality of university cleanroom equipment. Process failures do result from failure of the substrate transfer steps (failure of solder adhesion, failure –for unknown causes–of the substrate removal selective wet etch), but –equally– process failures arise in HBT fabrication steps unrelated to that of substrate transfer. Significant among these are excessive undercut in the emitter-base junction etch, failure of the emitter-base RIE or selective wet etches, emitter-base short-circuits forming during base contact liftoff, liftoff failures in interconnect metals, poor adhesion of resistor metal, and variation of resistor sheet resistivity. Given the resources available to a larger industrial group, various process difficulties –whether associated with or independent of substrate transfer– could be addressed. We believe the most serious fundamental difficulties are with the solder bonding and with the small wafer expansion after bonding (below), which most probably results from mechanical creep of the solder under exposure to stress and temperature cycles. Solder bonding also is presently limited to small wafer sizes (quarters of 50 mm wafers). More dimensionally stable alternatives, possibly spin-on-glasses, should be found for both the solder and the BCB dielectric.

Presently the largest working ICs fabricated in the process are 150-HBT ADCs and 250-HBT binary adders. The most significant process difficulty is dimensional change of the wafer during substrate transfer. Presently wafers show $3 \cdot 10^{-4}$ fractional expansion after transfer, resulting in ± 0.5 μm misregistration (during collector lithography) at the edges of the stepper exposure field if a 3 mm reticle is employed. We presently adjust the dimensions of the collector mask as a correction. At the expense of increased effort during collector lithography, a smaller exposure reticle size can be used for the collector lithography than for the steps preceding substrate transfer. The relative sizes of the emitter and collector junctions are determined by lithographic alignment tolerances, and the collector stripe width must exceed the emitter stripe width by twice the lithographic alignment tolerance. Our electron-beam lithography system can align to 0.1 μm registration, and our projection lithography system aligns to 0.1–0.3 μm registration, depending on the time since maintenance. Modern projection lithography systems are *much better*; 0.35-μm-resolution steppers have ~ 300 Å registration tolerance.

4.2. *Device results*

Transferred-substrate HBTs have been fabricated using contact lithography at 1–2 μm resolution, using a 0.5 μm stepper, and using electron-beam lithography. Fig. 17 shows a device defined by optical lithography [41]. Figure 18 shows HBT emitter-

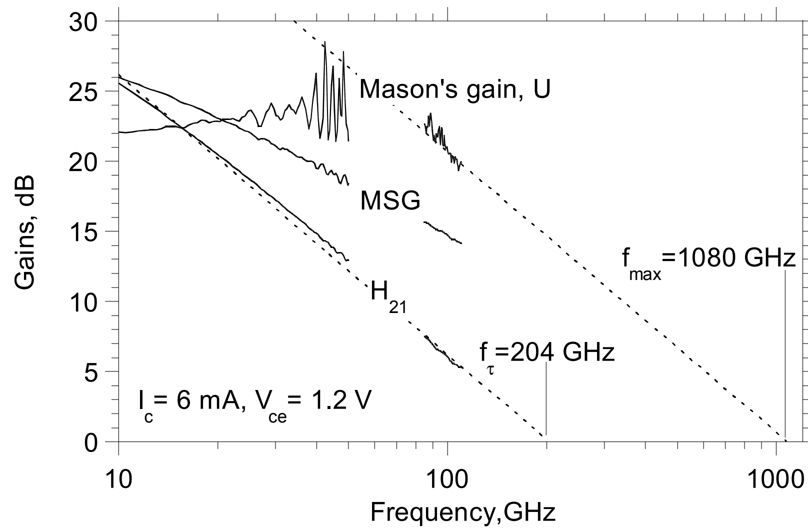


Figure 19: Gains of a 0.4 μm $\times$ 6 μm emitter and 0.7 μm $\times$ 10 μm collector HBT fabricated using electron-beam lithography. Theoretical -20 dB/decade (H_{21}, U) gain slopes are indicated. The device exhibits an *extrapolated* 1.08 THz f_{max}

base and collector-base junctions defined by electron-beam lithography.

Figure 19 shows microwave gains for a deep submicron device fabricated using electron-beam lithography, reported by Lee *et. al.* [43]. The base and collector layers are 400 Å and 3000 Å thick, while the emitter and collector junction dimensions are 0.4 μm $\times$ 6 μm and 0.7 μm $\times$ 10 μm. Biased at $V_{ce} = 1.2$ V and $I_c = 6$ mA ($J_e = 2.5 \times 10^5$ A/cm^2), the device exhibits 204 GHz f_τ. If extrapolated at -20 dB/decade, a 1080 GHz f_{max} is determined. We note, however, that such a 10:1 extrapolation must be treated with considerable caution.

We have extrapolated Mason's invariant (unilateral) gain at -20 dB/decade to determine the extrapolated f_{max}. Mason's gain [46] is invariant with respect to embedding the device in a lossless reciprocal network, and consequently is independent of pad inductive or capacitive parasitics and independent of the transistor configuration (common-emitter vs. common-base). For HBTs well-modeled by a hybrid-π equivalent circuit, Mason's gain conforms closely to a -20 dB/decade variation with frequency (fig. 20). In marked contrast, the maximum available / maximum stable gain is a function of the transistor configuration, and shows no fixed variation with frequency. f_{max} is unique; at $f = f_{max}$ the MAG/MSG and U are both 0 dB.

Device gains were measured over 45 MHz-50 GHz and 75-110 GHz using a microwave network analyzer and microwave wafer probes. To avoid uncorrectable

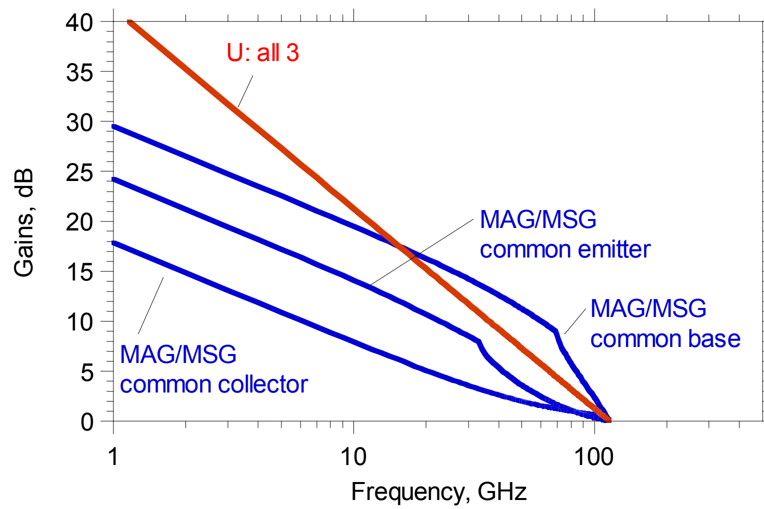


Figure 20: Variation of transistor gains with frequency, computed from a hybrid-π HBT model. Shown are the maximum available / maximum stable gains (MAG/MSG) in common-emitter, common-base, and common collector mode, and Mason's invariant, U, the unilateral gain

measurement errors (in S_{12}, hence U) arising from variable probe-probe electromagnetic coupling, the HBTs are separated from their probe pads by long on-wafer 50 Ω microstrip lines. On-wafer line-reflect-line calibration standards are used to de-embed the transistor S-parameters [44]. Before extracting HBT power gains to extrapolate f_τ and f_{max}, it is essential to verify the on-wafer calibration through measurement of known standards, to verify that the probe-probe parasitic coupling (as measured from the S_{12} of an on-wafer open-circuit standard) is at least 15-20 dB smaller than the measured transistor S_{12}, and to ensure that the transistor's measured S-parameters have a variation with frequency which conforms closely to that of a hybrid-π model. In the 75-110 GHz band, with high-f_{max} (hence very low S_{12}) HBTs, we have found that these requirements cannot be met using commercially-provided calibration substrates or with probe pads immediately adjacent to the transistor under test. The on-wafer LRM method is required, and the probe-probe separation must be at least 500 μm for all calibration test structures and for the device under test. In addition to the 10:1 extrapolation to 1.08 THz f_{max}, the very high power gain at 110 GHz also results in significant measurement variability, with repeated calibrations at the same bias point giving extrapolated f_{max} varying from 1.0 to 1.3 THz.

We have recently acquired a 140-220 GHz network analyzer with on-wafer probes, and are now developing methods to obtain precision HBT measurements in this band. Preliminary HBT measurements on a recently-processed submicron HBT wafer indicate $\sim$ 10 dB unilateral power gain and maximum stable gain at 200 GHz (the device is potentially unstable even at this high frequency) [36]. We have also recently demonstrated single-stage tuned HBT amplifiers at 185 GHz [36]; this indicates significant HBT gain at 200 GHz. Given current measurement data, the 1.1 THz extrapolated f_{max} is presently best viewed simply as an extremely high measured power gain at 100 GHz.

C_{cb} cancellation contributes substantially to the f_{max} obtained. At zero current, $C_{cb,e} = \epsilon A_e/T_c = 0.9$ fF. The measured variation of f_τ vs. V_{ce} (fig. 21) indicates $\partial\tau_c/\partial V_{ce} \sim 0.18$ ps/V, predicting $\sim$0.9 fF reduction in $C_{cb,e}$ from $I_c = 1$ mA to $I_c = 6$ mA. The *total* collector-base capacitance C_{cb} is determined from the measured variation with frequency of the imaginary part of the admittance parameter $\Im[Y_{12}] = j\omega C_{cb}$. The total C_{cb} determined from Y_{12} (fig. 22) shows a 0.64 fF decrease between 1 mA and 6 mA I_c. The measured variation in the total C_{cb} primarily reflects variation in the capacitance $C_{cb,e}$. The reduction $C_{cb,e}$ with bias current results in a rapid increase in f_{max} with bias (fig. 23).

Figure 24 shows the small-signal hybrid-π model. The measured S-parameters (fig. 25), h_{21}, and U, show good correlation with the hybrid-π model, and the model parameters are consistent with measured bulk and sheet resistivities and junction capacitances. The HBT output conductance is dominated by R_{cb}, which represents variation of collector-base leakage with bias. This is likely due to impact ionization. Base-width modulation in HBTs is negligible, hence R_{ce} is very large. $C_{be,poly}$ is a metal-polyimide-metal overlap capacitance between the emitter and base contacts

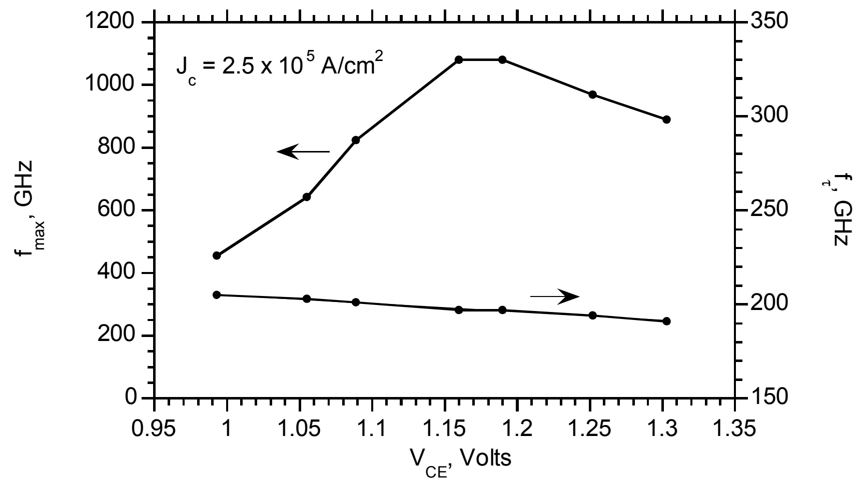


Figure 21: Variation of f_τ and $f_{\max}$ with collector-emitter voltage

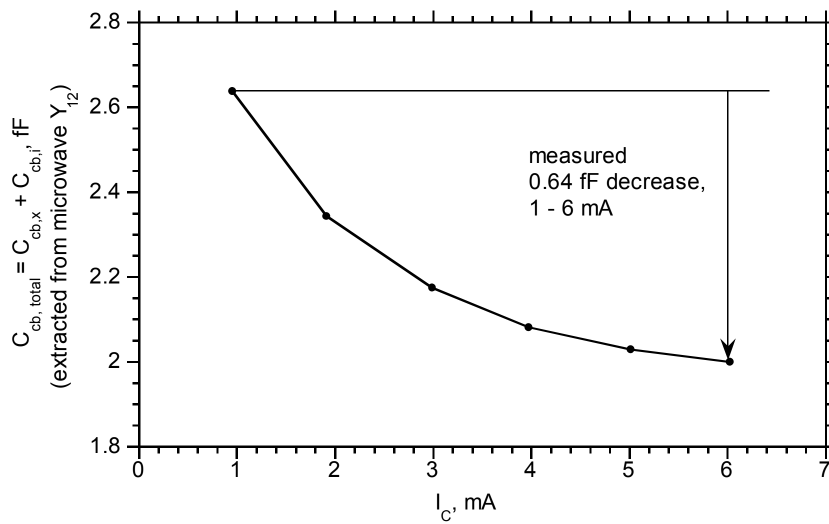


Figure 22: Collector-base capacitance extracted from Y_{12}, vs. emitter current

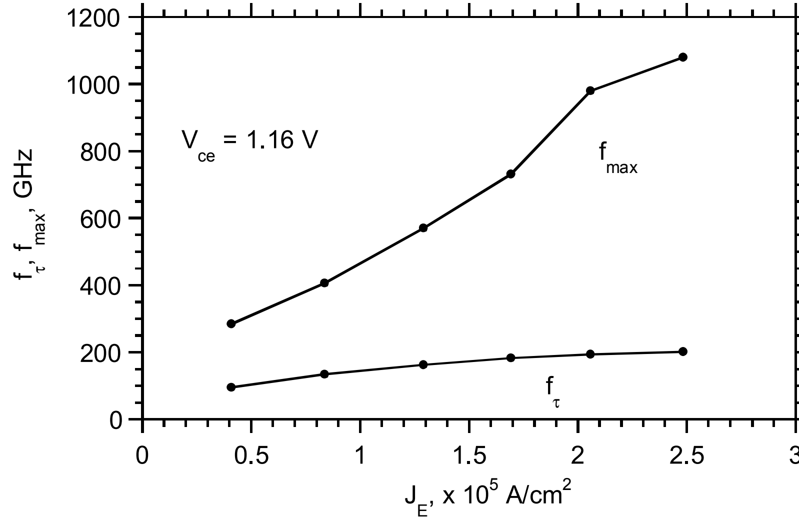


Figure 23: Variation of f_τ and $f_{\max}$ with emitter current density

(fig. 15) which contributes an additional $C_{be,poly}(R_{ex} + kT/qI_c) = 60$ fs to the transistor forward delay.

Neither contact lithography nor electron-beam lithography is suitable for fabrication of large ICs. We have fabricated HBT ICs using a 0.5 μm projection lithography system, and have obtained > 800 GHz f_{max} (fig. 26).

With the exception of reactively-tuned circuits, for which f_{max} is the sole determinant of circuit bandwidth, circuit design generally requires high values for both f_τ and f_{max}. Figure 27 [45] shows the forward delay of an HBT with 0.6 μm $\times$ 8 μm emitter and 2μm $\times$ 12 μm collector junctions, a 400 Å thick base with 52 meV bandgap grading, and a 2000 Å thick collector. The peak f_τ is 252 GHz, and RC charging terms constitute 35% of the forward delay. Figure 28 shows RF gains for a similar device with a thinner base, narrower emitter and collector junctions, and increased ($J_e = 2.5 \times 10^5$ A/cm^2) current density [47]. The device exhibits simultaneous 295 GHz f_τ and f_{max}. Significant terms in $\tau_{ec} = 1/2\pi f_\tau$ are $\tau_b + \tau_c = 395$ fs, $C_{je}/g_m = 82$ fs, $C_{cb}/g_m = 20$ fs, and $R_{ex}C_{cb} = 39$ fs. To obtain further increases in f_τ, the collector must be thinned, current density further increased and the emitter parasitic resistance improved.

Device scaling also reduces D.C. current gain. Base current in narrow-emitter InAlAs/InGaAs HBTs is predominantly due to conduction on the exposed InGaAs base surface between the emitter mesa and the base Ohmic contact. β decreases with emitter width, but increases as the base is thinned, as base bandgap grading

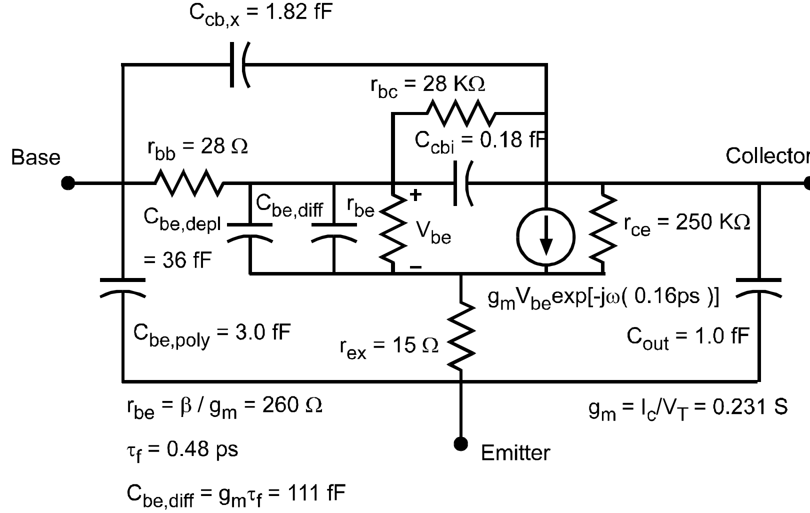


Figure 24: Device equivalent circuit model at $V_{ce} = 1.2$ V and $I_c = 6$ mA.

is increased, and (at the expense of f_{max}) as the emitter- base spacing is increased. $\beta > 50$ has been obtained with 0.2 μm emitters (fig. 29). Using 0.7 μm emitters and 300 Å base thickness with $2kT$ grading, $\beta \simeq 200$ is obtained.

4.3. *Interconnects and thermal management*

In developing an integrated circuit technology for microwave mixed-signal ICs, $\sim$ 100 GHz digital logic, and 100-300 GHz monolithic transmitters and receivers, significant issues in interconnects, packaging, and thermal management must also be addressed. Wiring parasitics, including line capacitance per unit length, line delay per unit length, ground via inductance, and parasitic ground return inductance, must all be minimized. Ground via inductance ($\sim$ 12 pH, or j7.5 Ω at 100 GHz) in standard 100-μm-substrate microstrip MIMICs makes low-impedance source/emitter grounding difficult in $>$ 100 GHz ICs. The interconnects must have low capacitance and low delay per unit length, and the wire lengths, hence transistor spacings, must be small. Given that fast HBTs operate at $\sim 10^5$ A/cm^2 current density, efficient heat sinking is then essential. To provide predictable performance, interconnects of more than a few ps length must have a controlled characteristic impedance. To prevent circuit-circuit interaction through ground-circuit common-lead inductance ("ground loops"), the IC technology must provide an integral low inductance –hence unbroken– ground plane for ground-return connections.

Ground-return inductance between the IC and package results in "ground bounce"

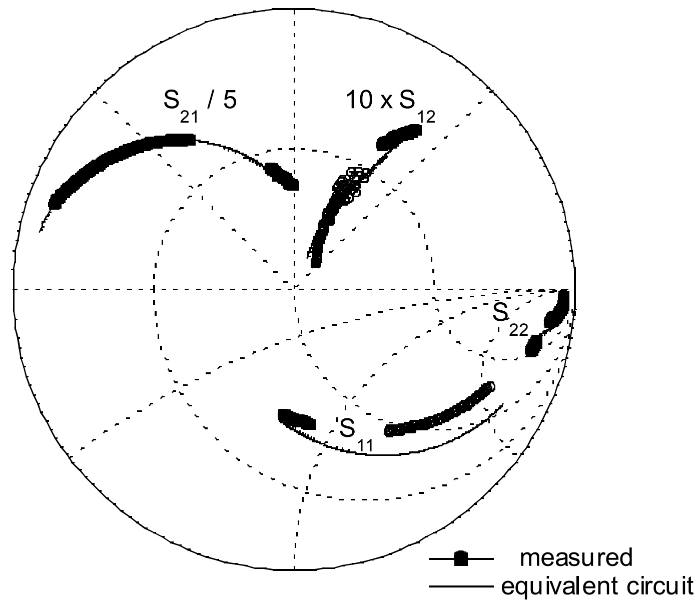


Figure 25: Measured 45 MHz–50 GHz and 75–110 GHz device S-parameters at $V_{ce} = 1.2$ V and $I_c = 6$ mA. The solid line represents S-parameters of the equivalent circuit model (fig. 24)

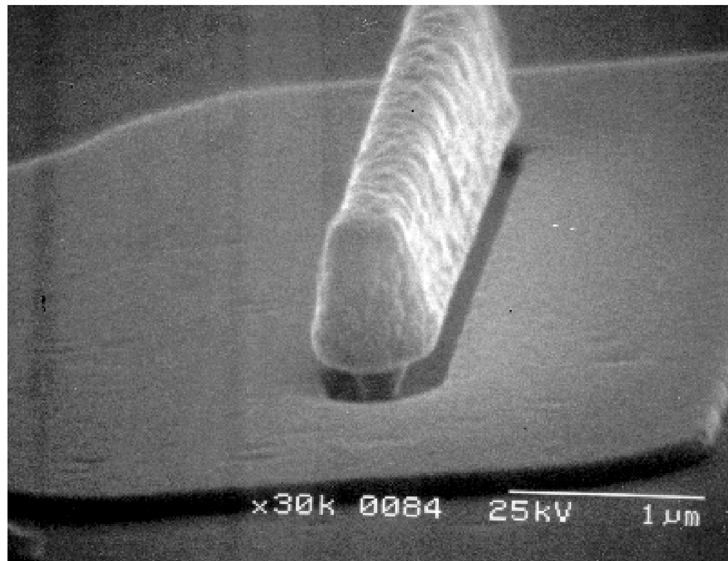


Figure 26: SEM from emitter side of a stepper-defined HBT with a 0.2 μm × 6 μm emitter.

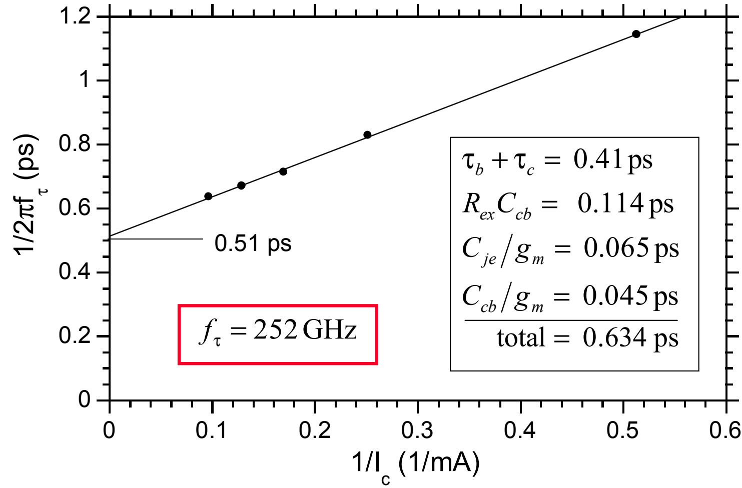


Figure 27: HBT forward transit delay *vs.* inverse emitter current for an HBT with a 2000 Å thick collector and a 400 Å thick base with $2kT$ bandgap grading. *RC* charging terms are significant in determining f_τ

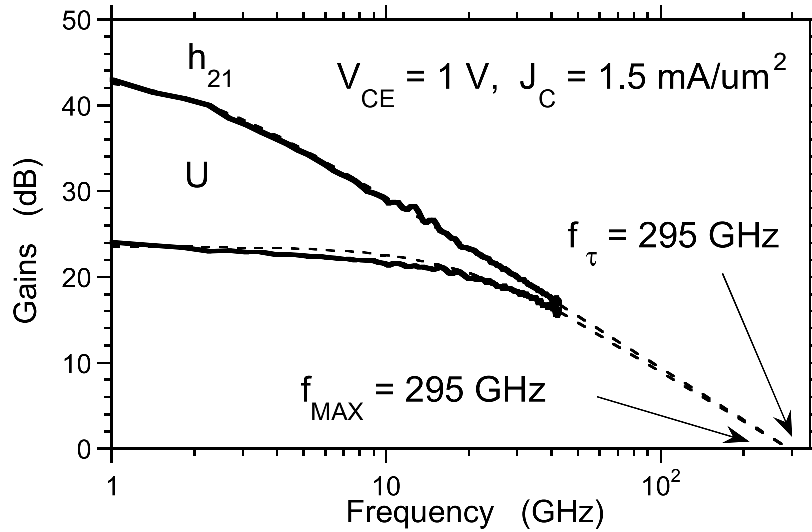


Figure 28: Measured RF gains for an HBT with a 300 Å base with 52 meV grading and a 2000 Åcollector.

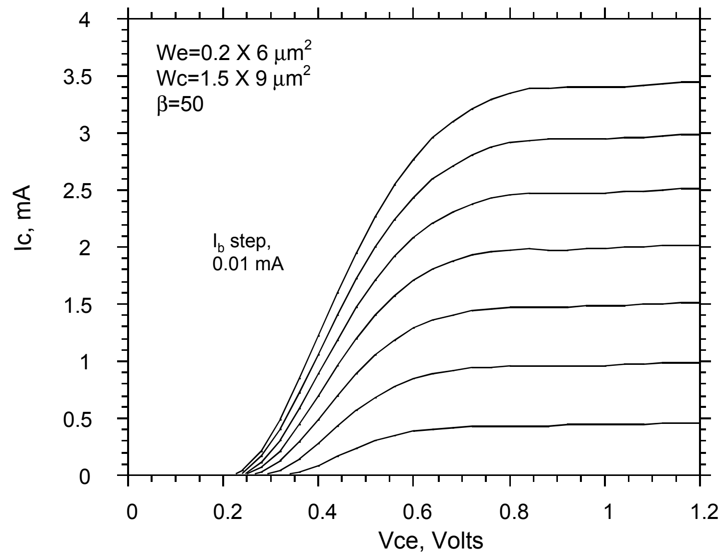


Figure 29: Common-emitter characteristics for a device defined by optical projection lithography. As a result of the 400 Å base with $2kT$ grading, β =50 is obtained even with a 0.2 μm emitter width.

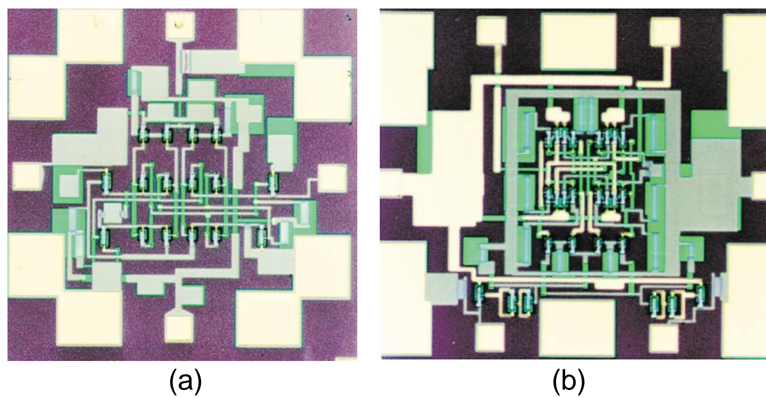


Figure 30: CML (a) and ECL (b) master-slave D-flip-flops.

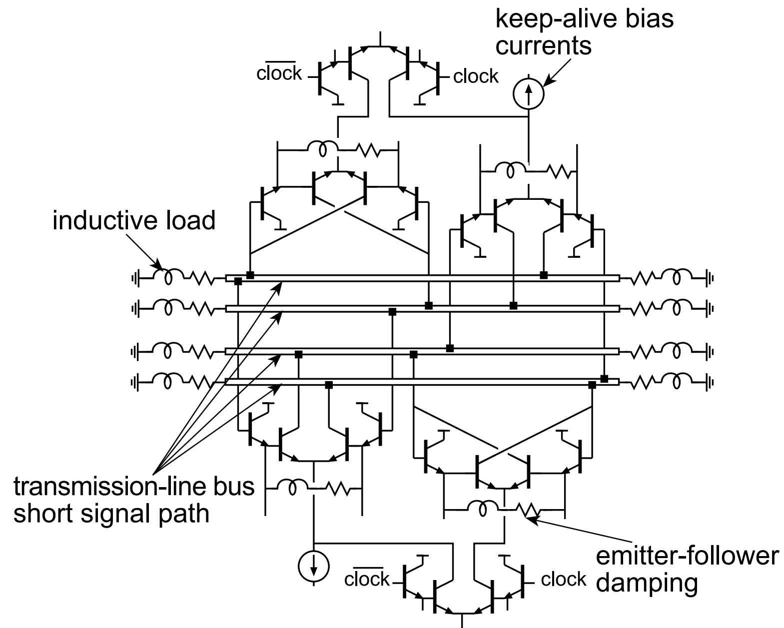


Figure 31: High-speed master-slave flip-flop; key features of the circuit design and physical layout.

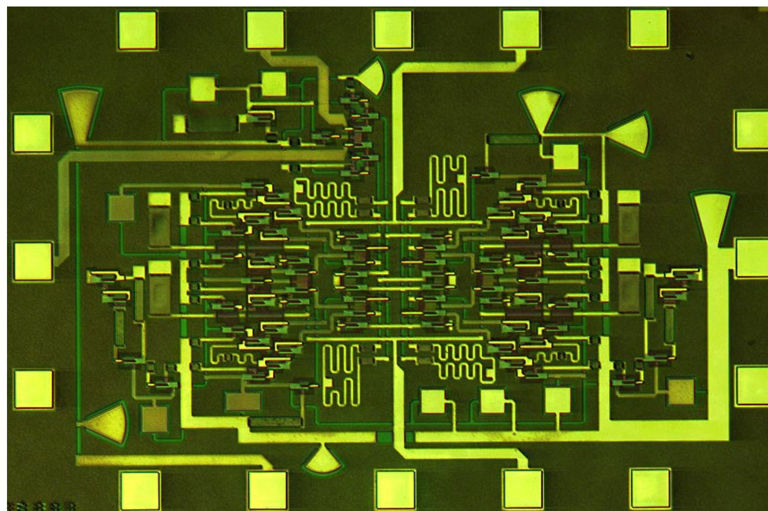


Figure 32: High speed master-slave flip-flop. The IC contains 70 HBTs.

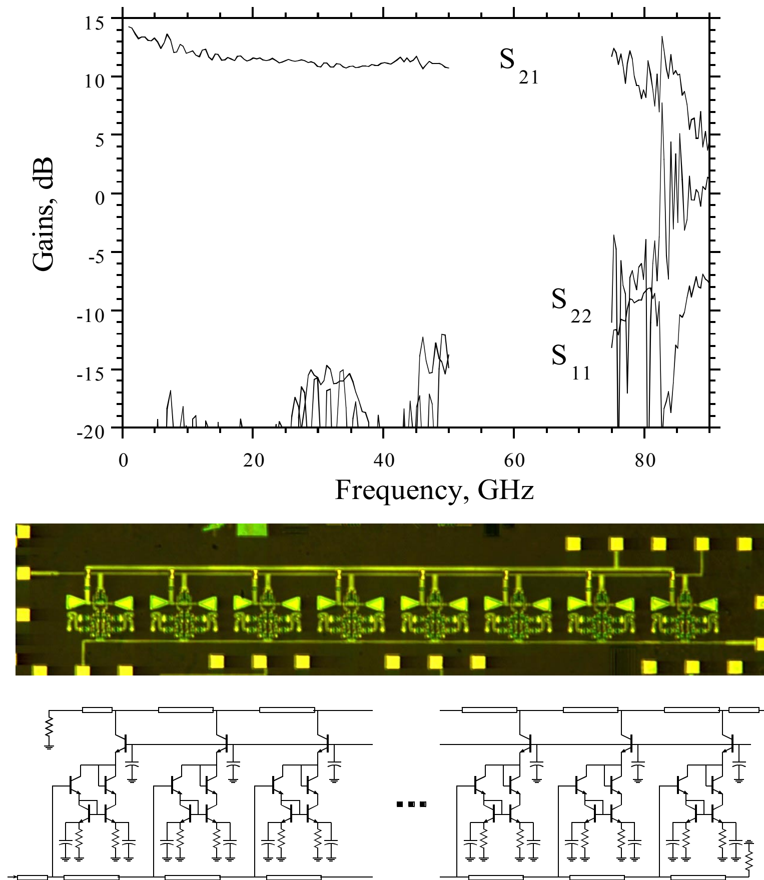


Figure 33: Distributed amplifier in the transferred-substrate process. The amplifier exhibits 11.5 dB gain and approximately 80 GHz bandwidth

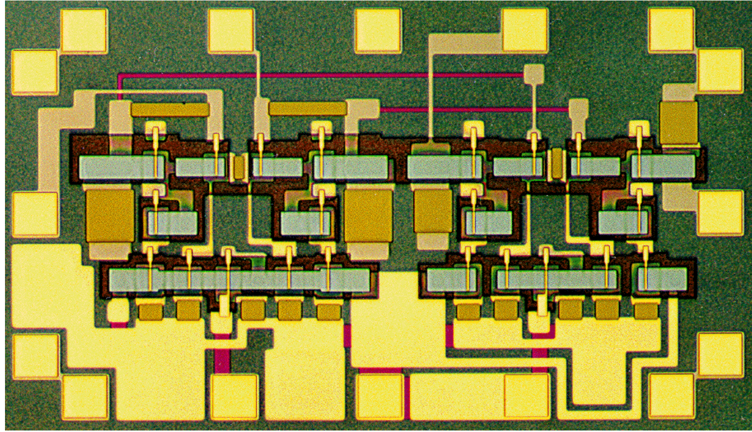
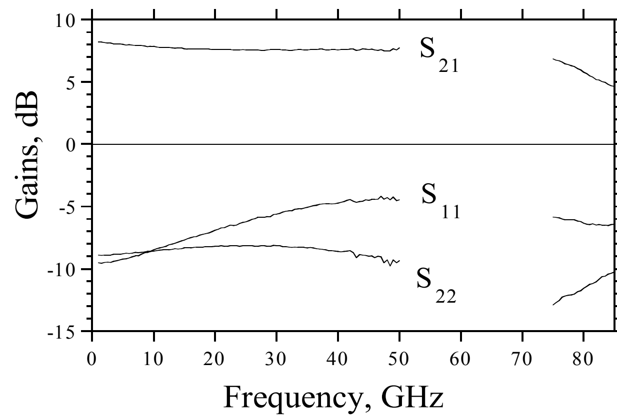


Figure 34: 11 dB gain, DC-50 GHz differential amplifier.

and hence interaction between the IC's input and output lines. For ICs with top-surface (coplanar-waveguide) ground connections and multiple input/output connections, ground bounce between IC and package will prevent 100 GHz operation. For an IC with N_{signal} signal lines of impedance Z_0, risetime ΔT, and voltage swing V_{signal}, and N_{ground} grounding bond wires of inductance $L_{\mathrm{bond}} \simeq 0.6\mathrm{pH}/\mu\mathrm{m}\cdot 300\mu\mathrm{m}$, the package-IC ground bounce is $V_{\mathrm{bounce}} = V_{\mathrm{signal}} N_{\mathrm{signal}} L_{\mathrm{bond}} / N_{\mathrm{ground}} Z_0 \Delta T$. For ground bounce equal to 10% of the signal amplitudes, a 100-GHz clock rate IC must have $N_{\mathrm{ground}}/N_{\mathrm{signal}}$ =5–10, and 80%–90% of the IC bond-pads must be devoted to IC grounding. Reported 10 GHz clock rate ICs devote $\sim$ 50% of IC pads for grounding. For mixed-signal and communications ICs, signal coupling through ground bounce must be much smaller than 10% of the digital I/O interface levels. Consequently, common-lead inductance between the IC and package ground systems must be made vanishingly small.

In addition to wide bandwidth transistors, the substrate transfer process provides thermal vias for HBT heatsinking, and microstrip transmission-line interconnects on a low dielectric constant substrate (ϵ_r=2.7) with vias, ground plane, and 3 levels of interconnects. At 5 μm length, the grounding vias are 20:1 shorter than in typical 100-μm-substrate microstrip MIMICs, reducing ground via inductance by over an order of magnitude. The process also incorporates NiCr resistors and Si_3N_4 MIM capacitors.

Presently, thermal resistance is dominated by temperature gradients internal to the transistor itself, arising from the low thermal conductivity of the InAlAs emitter and InGaAs base and collector layers. Thus, allowable power per unit HBT



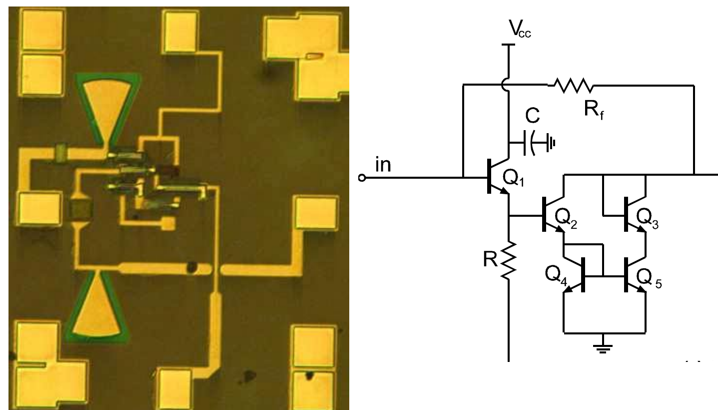


Figure 35: f_τ-doubler resistive feedback amplifier with 8.2 dB low-frequency gain and a DC-80 GHz 3-dB-bandwidth

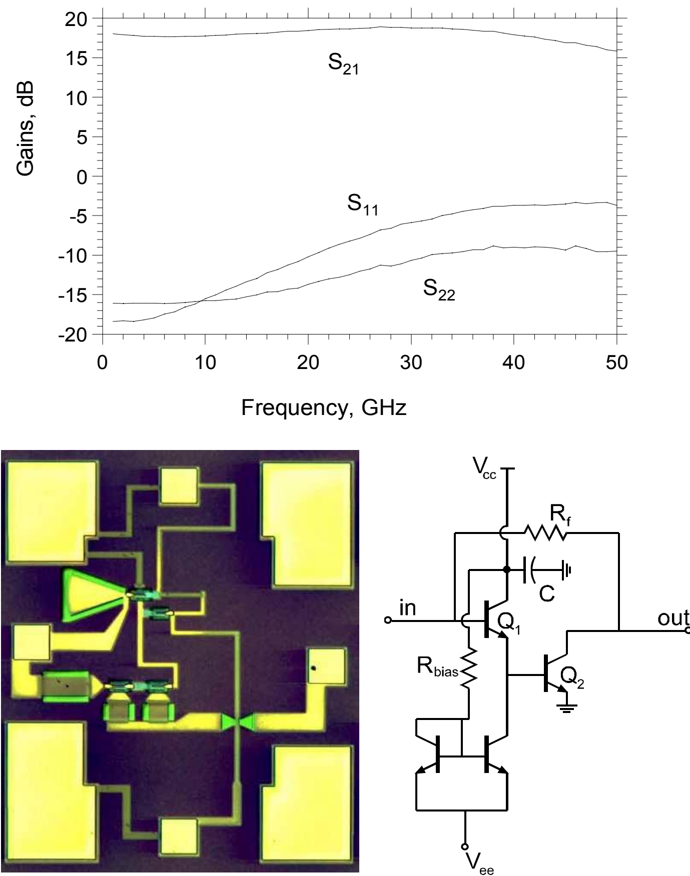


Figure 36: Measured S-parameters of a single-stage Darlington feedback amplifier. The amplifier exhibits 18 dB baseband gain, a 3-dB-bandwidth greater than 50 GHz, and greater than 400 GHz gain-bandwidth product.

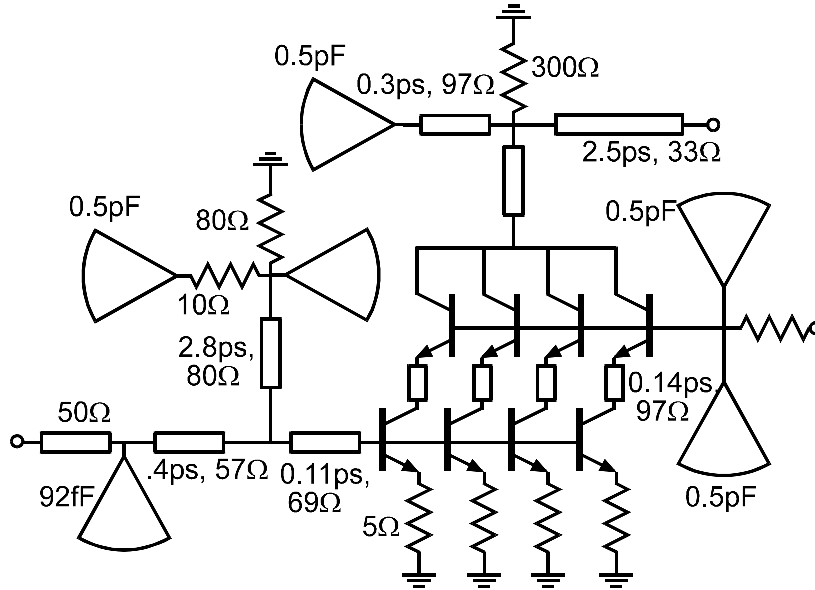


Figure 37: Circuit diagram of a W-band medium-power amplifier in the transferred-substrate HBT process.

emitter area remains comparable to mesa HBTs. For power transferred-substrate HBTs, use of high-thermal-conductivity InP emitter and collector epitaxial layers will greatly increase allowable power per unit HBT junction area. This is being pursued. To tolerate high power densities, the NiCr resistors must have thermal vias, which results in significant parasitic capacitance. Pull-up resistors in ECL do not require the thermal via.

5. Integrated circuit results

As a first demonstration of digital ICs in the transferred-substrate process, we fabricated ECL and CML master-slave flip-flops, configured as 2:1 static frequency dividers [53]. Circuits were fabricated using contact lithography, producing devices with 0.6 μm $\times$ 8μm emitters and 1.6 μm $\times$ 12 μm collectors. The devices operate at 1.25 mA/μm^2. The differential logic swing is 600 mV. The collector pull-up resistors are 50 Ω, hence the divider outputs directly drive 50 Ω output lines without buffering. For these initial designs, circuit design was entirely standard. The CML divider uses series-gated master and slave latches. Emitter-follower buffers are added to the CML clock and data ports to form the ECL divider. The ICs are shown in fig. 30. The ICs operated at maximum clock frequencies 47 GHz (CML) and 48 GHz (ECL) and dissipated 380 mW (ECL) or 75 mW (CML) from a -5 V supply.

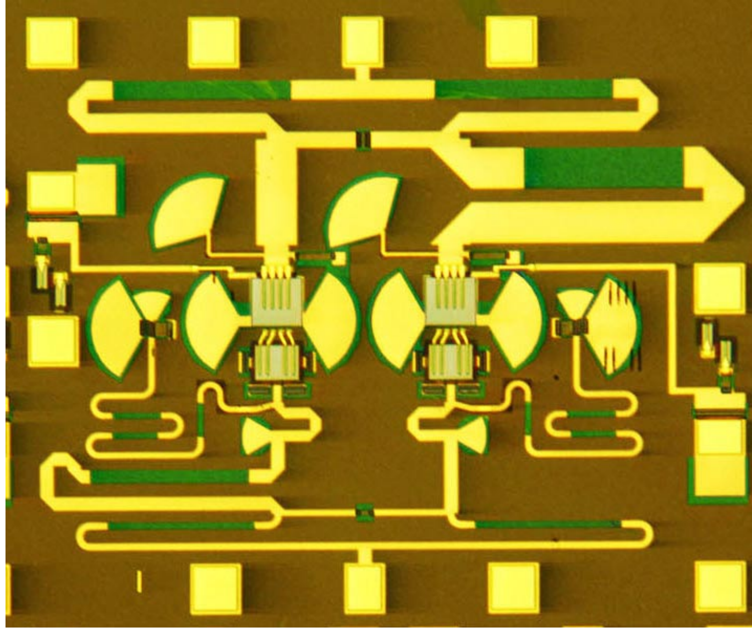


Figure 38: W-band balanced medium-power amplifier. The amplifier has 7 dB gain and produces 10.7 dBm saturated output power at 78 GHz.

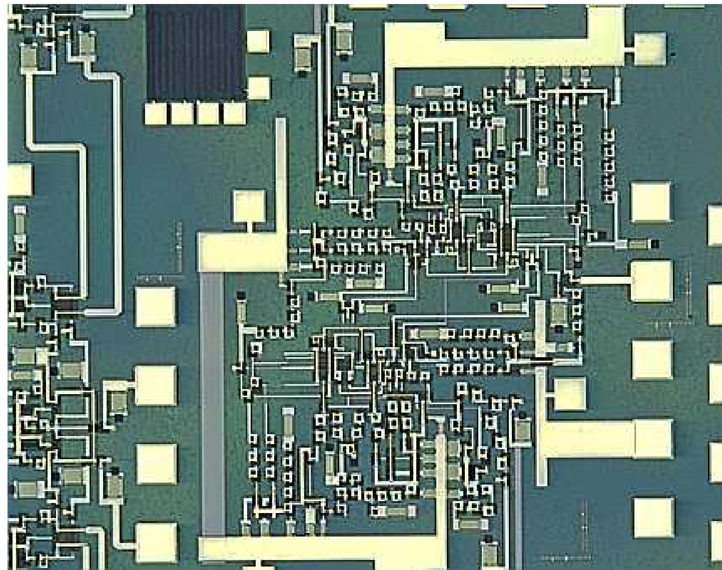


Figure 39: 2-bit carry generation logic circuits, developed as components of a microwave binary adder. The circuit contains 250 transistors.

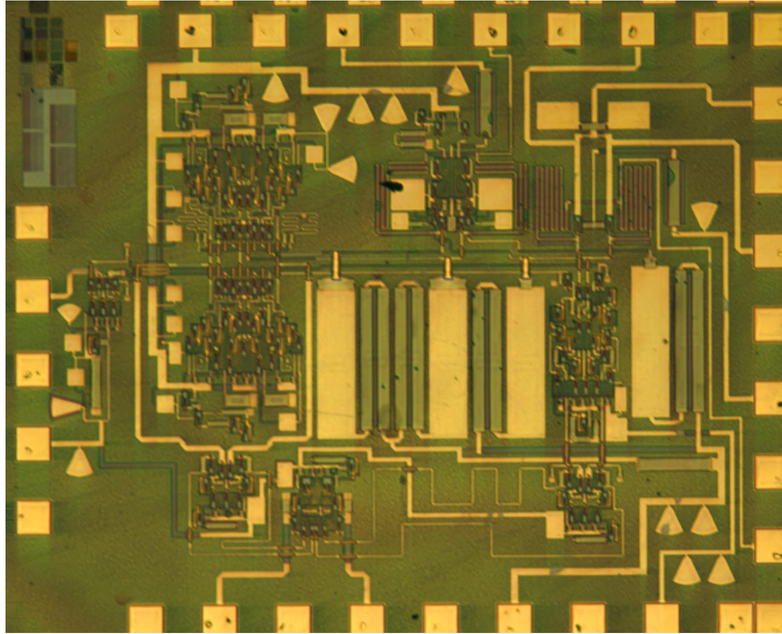


Figure 40: Δ–Σ ADC fabricated in the transferred-substrate process. The IC contains approximately 150 HBTs, and operates at 18 GHz clock rate.

Improved master-slave flip-flop designs were fabricated using optical projection lithography. These designs employed HBTs with 0.5 μm emitter and 1.5 μm collector junctions widths, with the devices operating at 2×10^5 A/cm^2 current density. Critical interconnects between stages are implemented as short doubly-terminated 100 Ω transmission lines at the center of the IC. The terminations use a *small* amount of series inductive peaking (fig. 31). Emitter-follower buffers increase logic speed but can induce strong ringing; L-R networks provide shunt loading of emitter-follower outputs and damp the emitter-follower pulse response. Keep-alive currents of 1/6 the logic currents keep the input stages weakly biased to minimize the input stage delays. The overall chip area is 1.0 x 0.4 mm, and consists of 76 transistors (fig. 32). The flip-flop dissipates 812 mW from a -5V supply, and the output buffer dissipates 38 mW from a -2V supply. Circuit simulations, which included all significant device and interconnect parasitics, predicted a 95 GHz maximum clock frequency when the latch is configured as a 2:1 static frequency divider. IC operation has been demonstrated to 66 GHz.

A number of high speed analog ICs have been fabricated in the transferred-substrate HBT process. Among these are 80 GHz distributed amplifiers [37] (fig. 33), 50 GHz broadband differential amplifiers for optical fiber receivers [56](fig. 34), and broadband Darlington and f_τ - doubler resistive feedback amplifiers (fig. 35). Figure 36 shows the measured gain vs. frequency of a Darlington resistive feedback

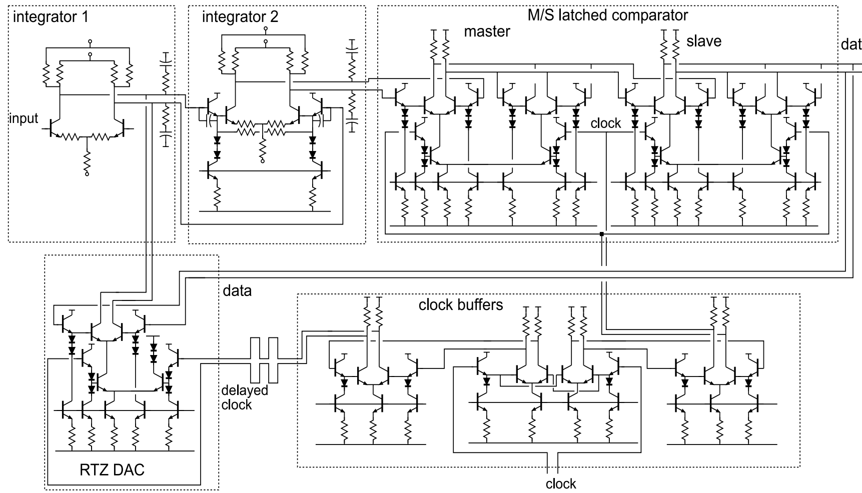


Figure 41: Simplified circuit diagram of the Δ–Σ ADC.

amplifier [57] [38]. Greater than 400 GHz gain-bandwidth product is obtain from a single Darlington stage. Tuned mm-wave amplifiers have also been demonstrated in the transferred-substrate process, including a 75 GHz amplifier [58] (figs. 38, 37) and, recently, a 185-GHz tuned amplifier [36].

Larger digital and mixed-signal ICs have also been fabricated in the transferred-substrate process. We have recently fabricated Δ–Σ ADCs [††] in the technology (fig. 40, fig. 41) [59]. These ICs have operated at an 18 GHz clock rate. Figure 42 shows the measured ADC signal/noise ratio and third-order distortion as a function of input power under two-tone test conditions. At a 990 MHz signal frequency, a peak signal/noise ratio of 120–125 dB (1 Hz) is obtained.

Larger digital circuits in development include sum and carry generation circuits for pipelined adder-accumulators (fig. 39). These circuits use 4-level series-gated current-steering logic and merged logic-latch circuits to obtain the equivalent of 2 AND, 2 OR, and 2 latching operations in a 50 ps clock period [60] .

6. Conclusions

Bipolar integrated circuit bandwidths have increased tremendously since the first demonstration of (bipolar) integrated circuits 40 years ago. Device, IC, and application bandwidths will continue to increase. With MOS transistors and III-V HEMTs (FETs), improved device bandwidths are obtained by lateral scaling (shorter gate lengths) combined with vertical scaling (thinner gate-channel insulating barriers), and progressive improvements in source/drain Ohmic contacts. With bipolar transistors, improved bandwidths are obtained by vertical scaling (thinner base and collector layers), combined with lateral scaling (narrower collector and emitter junc-

[††]For Δ–Σ converters, the terms "modulator" and "ADC" are used synonymously in the literature.

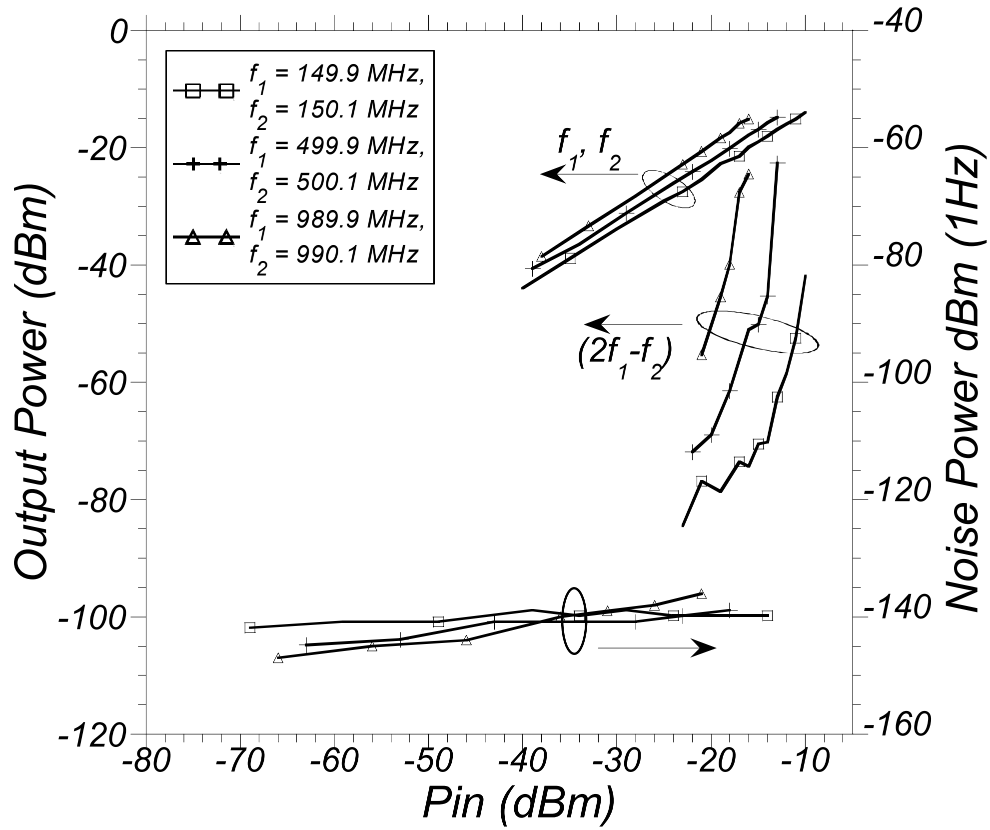


Figure 42: Δ-Σ ADC noise floor and third-order distortion power as a function of input power for different signal frequencies under two-tone test conditions.

tions), increased current density, and progressive improvements in emitter Ohmic contacts. While III-V HBTs benefit from strong heterojunctions, high mobilities, and high electron velocities, Si/SiGe bipolar transistors have been much more aggressively scaled, both in lithographic dimensions and emitter current density. Essential to the future success of III-V HBTs is submicron junction scaling and greatly increased current densities.

While bipolar ICs are much smaller than CMOS VLSI ICs, clock frequencies are much higher. In both technologies, thermal management and signal integrity are major limits to performance. As bipolar technologies evolve towards complex ICs operating at a 100 GHz clock, an increasing fraction of the total circuit connections will be terminated transmission lines of controlled characteristic impedance and minimal dielectric loading.

Acknowledgments

Work at UCSB was supported by the ONR under grants N0014-99-1-0041, N00014-01-1-0065, N00014-01-1-0066, N00014-01-1-0024, N00014-98-1-0750, and N00014-98-1-0830 (D. Purdy, D. VanVechten, M. Yoder, J. Zolper), by the AFOSR under grant F4962096-1-0019 (H. Schlossberg), and by the ARO under the Quasi-Optical MURI PC249806 (J. Harvey). JPL work was performed at the Center for Space Microelectronics Technology, JPL, Caltech, and sponsored by the NASA office of Space Science.

References

1. H. Kroemer, “Heterostructure Bipolar Transistors and Integrated Circuits”, Proc. IEEE, Vol. 70, No. 1, January 1982, pp. 13-25.
2. 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 (12), pp. 1709-1726, December 1993.
3. M. Sokolich, D. P. Docter, Y.K. Brown, A.R. Kamer, J.F. Jensen, W.E. Stanchina, S. Thomas III, C. H. Fields, D. A. Ahmari, M. Lui, R. Martinez, J. Duvall, “A low power 52.9 GHz static frequency divider in a manufacturable 180 GHz AlInAs/InGaAs HBT IC technology”, *Technical Digest*, IEEE GaAs IC Symposium , Nov. 1-4, 1998, Atlanta, Ga. , pp. 117-120.
4. P.K. Hughes, J.Y. Choe, J. Zolper, “Advanced Multifunctional RF system (AMRFS)”, *Technical Digest* (vol. XXV), Government Microcircuits Application Conference (GOMAC), Anaheim, CA., March 2000, pp. 194-197.
5. H. Suzuki, K. Watanabe, K. Ishikawa, H. Masuda, K. Ouchi, T. Tanoue and R. Takeyari, “InP/InGaAs HBT ICs for 40 Gbit/s optical transmission systems”, *Technical Digest*, IEEE GaAs IC Symposium, 1997, pp. 215-218.
6. H. M. Rein, E. Gottwald and T. F. Meister, “Si-bipolar - a potential candidate for high-speed electronics in 20 and 40 Gb/s TDM systems ?”, *Technical Digest*, Ultrafast Electronics and Optoelectronics Conference, 1997, pp. 118-120.
7. J.C. Candy and G.C. Temes, editors, *Oversampling Delta-Sigma Data Converters*, IEEE press, 1992, Piscataway, N.J.
8. D. C. Larson, “High speed direct digital synthesis techniques and applications”, *Technical Digest* IEEE GaAs IC Symposium, 1998, pp. 209-212.

9. J. Jensen, G. Raghavan, A. Cosand, R. Walden, “A 3.2 GHz second order sigma-delta modulator implemented in the InP HBT technology”, IEEE J. Solid-State Circuits, pp. 214-215, 1997.
10. S. Subbanna, J. Johnson, G. Freeman, R. Volant, R. Groves, D. Herman, B. Meyerson, “Prospects for Silicon Germanium based technology for very high speed circuits”, IEEE MTT-S International Microwave Symposium, Boston, MA, June 2000.
11. E.F. Crabbe, B.S. Meyerson, D.L. Harame, J.M.C. Stork, A. Megdanis, J. Cotte, J. Chu, M. Gilbert, C. Stanis, J. H. Comfort, G. L. Patton, S. Subbanna, “113-GHz f_τ graded-base SiGe HBTs”, 51st Device Research Conference, 1993; Abstract in IEEE Transactions on Electron Devices, Vol. 40, p. 2100, 1993.
12. K. Ohhata, T. Masuda, E. Ohue, K. Washio, “Design of a 32.7-GHz bandwdth AGC amplifier IC with wide dynamic range implemented with SiGe HBT”, IEEE J. Solid-State Circuits, Vol. 34, No. 9, Sept 1999, pp. 1291–1297.
13. E. F. Crabbe, J. H. Comfort, J. D. Cressler, J. Y.-C. Sun, and J. M.C. Stork, “High-Low Polysilicon-Emitter SiGe-Base Bipolar Transistors”, IEEE Electron Device Letters, Vol. 4, No. 10, October 1993, pp. 478–480
14. S. Yamahata, K. Kurishima, H. Ito and Y. Matsuoka, “Over-220-GHz-f_τ-and-f_{max} InP/InGaAs double-heterojunction bipolar transistors with a new hexagonal-shaped emitter”, *Technical Digest*, IEEE GaAs IC Symposium, 1995, pp. 163-166.
15. H. Kroemer, “Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region”, Solid State Electronics, vol. 28, pp. 1101-1103, 1985.
16. B.G. Streetman, *Solid State Electronic Devices*, third edition, Prentice-Hall, 1990.
17. S. Laux, W. Lee, “Collector signal delay in the presence of velocity overshoot”, IEEE Electron Device Letters, vol. 11, No. 4, pp. 174-176, 1990
18. T. Ishibashi, “Influence of electron velocity overshoot on collector transit times of HBTs”, IEEE Transactions on Electron Devices, vol. 37, no. 9, pp. 2103-2105, September 1990.
19. M. Littlejohn, K.W. Kim, H. Tian, “High-field transport in InGaAs and related heterostructures”, in *Properties of lattice-matched and strained Indium Gallium Arsenide*, P. Bhattacharya, editor, INSPEC, 1993, London.
20. E.P.O’Reilly, “Band structure of InP: Overview”, in *Properties of Indium Phosphide*, INSPEC, 1991, London.
21. C.T. Kirk, “A theory of transistor cutoff frequency (f_τ) fall-off at high current density”, IEEE Transactions on Electron Devices, ED-9, p. 164, (1962)
22. M.J.W. Rodwell, S.T. Allen, R.Y. Yu, M.G. Case, M. Reddy, E. Carman, J. Pusl, M. Kamegawa, Y. Konishi, and R. Pullela, “Active and Nonlinear Wave Propagation Devices in Ultrafast Electronics and Optoelectronics”, IEEE Proceedings, Vol. 82, No. 7, pp. 1037-1058, July 1994.
23. HP-EESOF Series IV microwave circuit simulation program. Hewlett-Packard Company, 3000 Hanover Street, Palo Alto, CA 94304, USA
24. M. Vaidyanathan and D. L. Pulfrey, “Extrapolated f_{max} of heterojunction bipolar transistors”, IEEE Transactions on Electron Devices, Vol. 46, No.2, February 1999.
25. M. Vaidyanathan and D. L. Pulfrey, *private communication*
26. Yoram Betser and Dan Ritter, “Reduction of the base collector capacitance in InP/GaInAs heterojunction bipolar transistors due to electron velocity modulation”, IEEE Transactions on Electron Devices, vol. 46 , no. 4, April 1999.
27. L. H. Camnitz and N. Moll, “An Analysis of the Cutoff-Frequency Behavior of Microwave Heterojunction Bipolar Transistors ”, In *Compound Semiconductor Transistors* , edited by S. Tiwari, pp. 21-45, IEEE Press, Piscataway, 1992.
28. R. W. H. Engelmann and C. A. Liechti, “Bias Dependence of GaAs and InP MESFET

Parameters", IEEE Transactions on Electron Devices, vol. ED-24, no. 11, pp. 1288-1296, Nov. 1977.

29. M.-C Ho, R.A. Johnson, W. J. Ho, M.F. Chang, P. M. Asbeck, ". High-performance low-base-collector capacitance AlGaAs/GaAs heterojunction bipolar transistors fabricated by deep ion implantation", IEEE Electron Device Letters, vol.16, (no.11), Nov. 1995. pp.512-14.
30. W. Liu, D. Hill, H. F. Chau, J. Sweder, T. Nagle and J. Delany, "Laterally etched undercut (LEU) technique to reduce base-collector capacitance in heterojunction bipolar transistors", *Technical Digest* , IEEE GaAs IC Symposium, pp. 167-170, 1995.
31. A. Gutierrez-Aitken et. al., 1999 International Electron Device Meeting, December, Washington, DC.
32. T. Oka, K. Hirata, K. Ouchi, H. Uchiyama, K. Mochizuki, T. Nakamura, ". Small-scaled InGaP/GaAs HBTs with WSi/Ti base electrode and buried SiO_2 " IEEE Transactions on Electron Devices, vol.45, (no.11), Nov. 1998, pp.2276-82.
33. H. Shimawaki, Y. Amamiya, N. Furuhata, K. Honjo, "High f_{max} AlGaAs/InGaAs and AlGaAs/GaAs HBT's with p^+/p Regrown Base Contacts", IEEE Transactions on Electron Devices, Vol. 42, No. 10, October 1995, pp. 1735-1744.
34. T. Oka, K. Hirata, K. Ouchi, H. Uchiyama, T. Taniguchi, K. Mochizuki, T. Nakamura, "Advanced Performance of Small-Scaled InGaP/GaAs HBTs with f_τ over 150 GHz and f_{max} over 250 GHz", In *Proceedings*, 1998 IEEE International Electron Device Meeting, December 6-9, San Francisco, pp. 653-656
35. Y. Matsuoka, S. Yamahata, K. kurishima and H. Ito, "Ultrahigh-speed InP/InGaAs Double-Heterostructure Bipolar Transistors and Analysis of Their Operation", Japanese Journal of Applied Physics, vol. 35, pp.5646-5654, 1996.
36. M. Urteaga, D. Scott, T. Mathew, S. Krishnan, Y. Wei, M.J.W. Rodwell, "185 GHz Monolithic Amplifier in InGaAs/InAlAs Transferred-Substrate HBT Technology", *Submitted to the* 2001 MTT-S International Microwave Symposium, submitted Dec. 2000.
37. S. Krishnan, S. Jaganathan, T. Mathew, Y. Wei, M.J.W. Rodwell "Broadband HBT amplifiers", 2000 IEEE Cornell Conference on High Speed Electronics.
38. S. Krishnan, D. Mensa, J. Guthrie, S. Jaganathan, T. Mathew, R. Girish, Y. Wei and M.J.W. Rodwell, "Broadband lumped HBT amplifiers." IEE Electronics Letters, pp.466-7, Vol 36, No.5.
39. S. Yamahata, Y. Matsuoka, T. Ishibashi, "Ultrahigh-speed AlGaAs/GaAs ballistic collection transistors using carbon as a p-type dopant", Electronics Letters, Vol. 29, No. 22, 28 October 1993, pp. 1996-1997.
40. R. P. Smith, S.T. Allen, M. Reddy, S.C. Martin, J. Liu, R. E. Muller, and M.J.W. Rodwell, "0.1 μm Schottky-Collector AlAs/GaAs Resonant Tunneling Diodes", IEEE Electron Device Letters, Vol. 15, No. 8, August 1994.
41. Q. Lee, B. Agarwal, R. Pullela, D. Mensa, J. Guthrie, L. Samoska, M. Rodwell, "A $>$ 400 GHz f_{max} transferred-substrate heterojunction bipolar transistor IC technology", IEEE Electron Device Letters, vol. 19, pp. 77-79, 1998.
42. W. E. Stanchina, *et. al.*, "An InP-based HBT fab for High-Speed Digital, Analog, Mixed-Signal, and Optoelectronic ICs" *Technical Digest,*GaAs IC Symposium, 1995, pp. 31-34
43. Q. Lee, S.C. Martin, D. Mensa, R.P. Smith, J. Guthrie, S. Jaganathan, Y. Betser, T. Mathew, S. Krishnan, L. Samoska, and M.J.W. Rodwell, "Submicron transferred-substrate heterojunction bipolar transistors with greater than 1 THz f_{max}.", postdeadline paper, 1999 IEEE Device Research Conference, June, Santa Barbara, CA.
44. Q. Lee, S.C. Martin, D. Mensa, R.P. Smith, J. Guthrie, and M.J.W. Rodwell, "Submicron transferred-substrate heterojunction bipolar transistors", IEEE Electron Device Letters, Vol. 20, No. 8, August 1999, pp. 396-398.

45. D. Mensa, Q. Lee, J. Guthrie, S. Jaganathan, and M.J.W. Rodwell, "Transferred-substrate HBTs with 250 GHz current-gain cutoff frequency ", *Proceedings*, 1998 International Electron Device Meeting, San Francisco, December.
46. S. J. Mason, "Power gain in feedback amplifier", IRE Trans. Circuit Theory, vol. CT-1, pp. 20-25, 1954.
47. Y. Betser , D. Mensa, S. Jaganathan, T. Mathew and M. Rodwell, "InAlAs/InGaAs HBTs with Simultaneously High values of Ft and Fmax for mixed analog/digital applications", *To be published*, IEEE Electron Device Letters, submitted July 2000.
48. J. Guthrie, D. Mensa, T. Mathew, Q. Lee, S. Krishnan, S. Jaganathan, S. Cerhan, Y. Betser, M.J.W. Rodwell, "A 50 mm Copper/Polymer Substrate HBT Technology for > 100 GHz MIMICs", 1999 IEEE Conference on InP and related materials, May, Davos, Switzerland.
49. E. Sano, Y. Matsuoka, T. Ishibashi, "Device Figures-of-Merit for High Speed Digital ICs and Baseband Amplfiers" IEICE Transactions on Electronics, E78-C (1995) pp. 1182-1188.
50. T. Enoki, E. Sano, T. Ishibashi, "Prospects of InP-based IC technologies for 100 Gb/s-class lightwave communications systems", International Journal of High Speed Electronics and Systems, *this issue*.
51. D.A. Hodges and H.G. Jackson, *Analysis and Design of Digital Integrated Circuits, 2nd Edition*, McGraw-Hill, 1983, ISBN 0-07-029153-5
52. P.K. Tien, "Propagation delay in high speed silicon bipolar and GaAs HBT digital circuits", International Journal of High Speed Electronics and Systems, 1(1) pp. 101-124, 1990.
53. R. Pullela, D. Mensa, B. Agarwal, J. Guthrie, M. Rodwell, "47 GHz static frequency divider in Ultrafast transferred-substrate heterojunction bipolar transistor technology", 1998 Conference on InP and Related Materials, May, Tsukuba, Japan.
54. Q. Lee, D. Mensa, J. Guthrie, S. Jaganathan, T. Mathew, S. Krishnan, S. Cerhan and M.J.W. Rodwell, "66 GHz static frequency divider in transferred-substrate HBT technology", 1999 IEEE RF/Microwave monolithic circuits symposium, June, Anaheim, CA.
55. B. Agarwal, R. Pullela, Q. Lee, D. Mensa, J. Guthrie, M. J.W. Rodwell, "80 GHz Distributed Amplifiers with transferred-substrate heterojunction bipolar transistors", 1998 IEEE MTT Microwave Symposium, June, Baltimore Md.
56. B. Agarwal, Q. Lee, R. Pullela, D. Mensa, J. Guthrie, M. J.W. Rodwell, "A transferred-substrate HBT wideband differential Amplifier to 50 GHz", IEEE Microwave and Guided Wave Letters, June 1998.
57. D. Mensa, R. Pullela, Q. Lee, B. Agarwal, J. Guthrie, S. Jaganathan, M.J.W. Rodwell, "Baseband amplifiers in the transferred-substrate HBT technology". 1998 IEEE GaAs IC symposium. Nov. 1-4, Atlanta, Ga.
58. J.R. Guthrie, M. Urteaga, D. Scott, D. Mensa, T. Mathew, Q. Lee, S. Krishnan, S. Jaganathan, Y. Betser, M. Rodwell, "HBT MMIC 75 GHz power amplifiers", 2000 IEEE Conference on Indium Phosphide and Related Materials, May, Williamsburg, Va.
59. S. Jaganathan, D. Mensa, T. Mathew, Y. Betser, S. Krishnan, Y. Wei, D. Scott, M. Urteaga, M. Rodwell, "An 18 GHz clock rate continuous-time Δ-Σ modulator implemented in InP transferred-substrate HBT technology", 2000 IEEE GaAs IC Symposium, November, Seattle, Wa.
60. T. Mathew, S. Jaganathan, D. Scott, S. Krishnan, Y. Wei, M. Urteaga, M. J. W. Rodwell, S. Long, "2 bit adder carry and sum logic circuits at 19 GHz clock frequency in Transferred Substrate HBT technology", *submitted to the* 2001 IEEE Conference on Indium Phoshide and related materials, May, Nara, Japan.