THz Bandwidth InP HBT Technologies and Heterogeneous Integration with Si CMOS

Teledyne Scientific Company
Thousand Oaks, CA, USA
Miguel.Urteaga@Teledyne.com

M.J.W. Rodwell
Department of ECE
University of California Santa Barbara
Santa Barbara, CA, USA

Abstract—Through aggressive lithographical and epitaxial scaling, the bandwidths of InP-based heterojunction bipolar transistors have been extended to THz frequencies. At 130nm emitter dimensions, transistors with maximum frequencies of oscillation (f_{max}) of >1THz have been demonstrated with accompanying circuit demonstrations at 670GHz. At 250nm emitter dimensions, high efficiency and high power density mm-wave power amplifiers covering E-band (71GHz) to G-band (235GHz) have been fabricated. The utility of these high performance transistors can be further enhanced through heterogeneous integration with Si CMOS. We have demonstrated wafer-scale 3D integration of InP and Si using a low temperature oxide-to-oxide bonding process with embedded Cu interconnects.

Keywords—Indium Phosphide; HBT; terahertz; heterogeneous integration

I. INTRODUCTION

Transistors fabricated in the InP material system have demonstrated the highest reported RF figures-of-merit and can be used for signal amplification in the THz frequency regime (0.3-3THz). Both InP high-electron mobility transistors (HEMTs) and heterojunction bipolar transistors (HBTs) have been reported with maximum frequencies of oscillation exceeding 1THz [1], [2], [3], [4], [5]. Recently, HEMT amplifiers have been demonstrated operating at >1THz [6], and HBT amplifiers have been demonstrated operating to >600GHz [7], [8], [9]. Increasingly sophisticated THz monolithic integrated circuits (TMICs) are being demonstrated as these technologies continue to mature.

Presently, system applications in the THz frequency range are primarily low-volume and limited to metrology, radiometry and spectroscopy. Therefore, there is substantial interest in exploiting high bandwidth InP HBT technologies for applications at lower frequencies. Examples of such applications include: wide bandwidth and high efficiency mm-wave power amplifiers, high-power density sub-mm-wave power amplifiers, high clock frequency mixed-signal data converters and high performance mm-wave beamformer integrated circuits for phased arrays. This application space can be significantly expanded through the integration of high performance InP with high-density logic devices like those available in SiGe BiCMOS processes.

There have been many recent demonstrations of heterogeneous integration of InP HBTs with Si CMOS using a variety of integration methods. These include monolithic growth on Si for in-plane integration [10], InP epitaxial transfer on top of the CMOS interconnect stack [11], [12], fine-pitch bonding of InP chiplets on CMOS wafers [13]. We have demonstrated wafer-scale 3D integration of our InP HBT technology with Si CMOS using Ziptronix Direct Bond Interconnect technology [14]. Our approach enables a “CMOS-on-top” architecture that has significant advantages for thermal management of the high power density InP HBTs.

In this paper, we review Teledyne’s THz bandwidth InP HBT IC technologies, including transistor and circuit demonstrations, and discuss our 3D heterogeneous integration of InP HBTs with a 130nm CMOS foundry process.

II. TERAHERTZ HBT TECHNOLOGIES

A. Device Technology

General scaling laws for InP HBTs have been developed and outlined in [15]. HBT bandwidth increases with vertical scaling of transistor epitaxy and lateral scaling of transistor dimensions provided Ohmic contact resistivities can be reduced concurrently. A critical challenge towards the development of THz bandwidth devices has been the development of high-yield self-aligned base-emitter contact schemes that meet targeted Ohmic contact resistivities. The contact resistance is a function of achievable doping levels in the n+-emitter cap and p+-base, Ohmic contact metallurgy and surface cleanliness prior to metal deposition. The contacts must also be thermally stable under high current operation.

At Teledyne, we have developed a self-aligned HBT process flow that utilizes an electroplated emitter contact with dielectric sidewall. Variants of this process flow have been used in HBT IC technologies with emitter junction dimensions scaling from 500nm to 130nm ([16], [17], [3]) demonstrating successive increases in HBT bandwidth. Fig. 1 shows a TEM cross-section of a 130nm InP HBT. The emitter contact is formed utilizing an Au-based electroplating process on a refractory emitter contacting layer. The straight sidewall profile from electroplating facilitates the formation of dielectric sidewall spacers that passivate the base-emitter junction and provide isolation from the self-aligned base contact. To achieve low base contact resistance, an extremely high carbon-doping concentration (>7x10^{19} cm^{-3}) is utilized in the InGaAs base.
layers closest to the contact. Grading of the base doping concentration is performed to maintain acceptable current gain and improve base transit time [18]. By controlling surface contaminants and limiting penetration of the Ohmic contact metallurgy, contact resistivities of ~5Ω·µm² can be realized on InGaAs base layers <30nm thick.

HBT fabrication is performed on 100mm InP substrates with epitaxial layers grown by molecular beam epitaxy. Our transistors are double heterojunction bipolar devices (DHBTs) that utilize an InGaAs base layer with superlattice (InGaAs/InAlAs) grading to remove conduction band discontinuities between both the InP emitter and collector regions. The base-emitter grade reduces HBT turn-on voltage and has been found to improve the forward current ideality factor (N_C) compared to an abrupt junction. A proper base-collector grade design is critical to avoid current blocking and enable high current density operation.

HBT device mesas are formed using wet-chemical etch processes that can selectively remove either InP or InGaAs/InAlAs containing layers. After the HBT formation, Benzocyclobutene (BCB) is used as a final HBT passivation and planarization layer. The BCB is etched back to expose the tail emitter contact eliminating the need for a precisely aligned via to this contact. Vias are formed to the base and collector contacts and an electroplated Au process is used to form the first interconnect level.

Owing to the wide bandgap InP collector, InP DHBTs have significantly higher breakdown voltages than comparable SiGe HBTs. At the 250nm scaling generation, our DHBTs have a 150nm N-collector thickness and demonstrate a common-emitter breakdown voltage BV_CEO ~4.5V (J_E =10mA/µm²). The transistors support high current and power densities. This is illustrated in the common-emitter I-V characteristics (Fig. 2) that show the transistor operating at current densities of >10⁹ A/cm², and power densities of >20 mW/µm². At the 130nm scaling generation, the N collector region is scaled to 100nm with a corresponding breakdown voltage BV_CEO ~3.5V. HBT current gain is limited by the high base doping used to minimize base contact and access resistances. Typical values of β are ~25 and ~18 for our 250nm and 130nm HBTs, respectively.

Measurements of highly-scaled HBTs are challenging because the reverse transmission measurements (S₁₂) are easily corrupted by coupling between on-wafer probes or by unwanted mode propagation in the semiconductor substrates. To mitigate this, we utilize extended reference plane thin-film microstrip calibration structures with the ground plane formed in first level metal (M1) and the signal line in the topmost metal. A multi-line TRL calibration and deembedding procedure is performed using on-wafer standards to place the measurement reference planes at the transistors’ M1 terminals.

Fig. 3 shows the measured transistor gains of a 0.13x2µm² InP HBT taken from [3]. Measurements were performed from 8-50GHz and 75-105GHz. The HBT figures-of-merit f and f_max are extrapolated from least squares fits to single-pole transfer functions of the measured h21 and unilateral power gain (U), respectively. Fits are performed on the data to 50GHz, as measurements of U show considerable variation at high frequencies (>75GHz). The HBT exhibits an extrapolated f/f_max of 521GHz/1.15THz. At the 250nm scaling generation, typical peak transistor figures-of-merit are f/f_max 400GHz/700GHz.

B. Interconnects

Our IC technology utilizes multi-level thin film wiring with a BCB interlayer dielectric (ε_r=2.7) and electroplated Au-based metallization. A 3- or 4-level wiring stack with 2µm ILD thicknesses is typically used for digital and mixed-signal ICs;

![Fig. 1. TEM cross-section of 130nm InP HBT.](image1)

![Fig. 2. Measured common-emitter I-V characteristics 0.25x4 µm² HBT.](image2)

![Fig. 3. Measured short circuit current gain (h21) and Unilateral power gain (U) of 0.13x2µm² InP HBT [3].](image3)
however; a modified process with a thicker (>5μm) top-most dielectric has also been developed to provide low-loss microstrip for mm-wave and THz circuit applications. For packaging and substrate mode control wafers are thinned to 50μm or 75μm with through substrate vias and backside metallization. We have also developed backside etch singulation processes that permits the realization of non-rectangular IC die. This process permits the formation of narrow InP extensions that can support E-plane waveguide probes (waveguide-to-chip transitions) that can extend into a rectangular waveguide channel with a minimal break in the waveguide sidewall. This process is used for packaging ICs in waveguide blocks at frequencies above 300GHz.

C. Transistor Models

Transistor bandwidths have exceeded frequencies that discrete transistors can be accurately measured and we extract transistor model parameters from lower frequency measurements. The Keysight III-V HBT model is used for large signal modeling of our InP HBTs. The model is capable of capturing many of the unique properties of III-V bipolar transistors than are difficult to model with standard Si BJT models [19]. These properties include: collector transit time modulation by both collector current and base-collector voltage, collector-base capacitance cancellation and base-collector current blocking at high injection levels. Accurately modeling these phenomenon is important for predicting the HBT large signal characteristics particularly those related to gain compression under large signal drive. We have found that the HBT model remains accurate at frequencies approaching the transistor cutoff frequencies; an assertion that is supported by the good agreement we have observed between measurement and simulation for THz frequency IC designs.

III. CIRCUIT DEMONSTRATIONS

A. mm-Wave and sub-mm-Wave Power Amplifiers

With high power handling capability and wide RF bandwidths, our 250nm HBT process has been used to demonstrate mm-wave and sub-mm-wave power amplifiers. At E-band (71-86GHz) and W-band (92-95GHz), high RF output power densities (1-2W/mm normalized to emitter periphery) and high power added efficiency (PAE) can be realized with common-emitter gain stages biased for Class-A or Class-AB operation. In [20], we report the demonstration of a two-stage common-emitter E-band amplifier with 130mW $P_{\text{out}}$ and 40% PAE at 81GHz. The amplifier demonstrates >100mW $P_{\text{out}}$ with an associated PAE of >20% from 71-95 GHz. A similar two-stage design operating from 96-120GHz demonstrates $P_{\text{out}}$ of 85-110mW with an associated PAE of 25.5-18.9% across the band.

Higher-power E-band amplifier demonstrations in the technology include a 71-76GHz PA with 439mW $P_{\text{out}}$ and 26.9% associated PAE at 76 GHz [21], and an 81GHz PA demonstrating 470mW $P_{\text{out}}$ with 23.4% associated PAE [22]. In [22], a novel sub-quarter wavelength balun design is used to enable series power combining permitting a compact design that demonstrates an RF power density of 443mW/mm² normalized to die area (Fig. 4). The circuit and device-level power densities achieved in the technology are comparable to state-of-the-art Gallium Nitride (GaN) HEMT technologies at E-band and W-band.

At sub-mm-wave frequencies, 250nm InP HBT power amplifiers have demonstrated record levels of output power around 220GHz. Our basic design topology at these frequencies utilizes Cascode gain stages with Wilkinson power splitter/combiners and has demonstrated extremely high large signal bandwidths. A three-stage 4-W combined power amplifier has been demonstrated with 50-80mW $P_{\text{out}}$ and >20dB of large signal gain from 190-250GHz [23], and in [24] we reported the first MMIC with >200mW output power at >200GHz. Using a 16-way combined power amplifier an output power of 208mW at 210GHz was demonstrated (Fig. 5).

Our MMICs have been power combined to form high-power solid-state PA (SSPA) modules. In [25], 32 4-way combined power amplifier ICs were combined to demonstrated 710mW output power at 230GHz, and in [26] 16 8-way combined power amplifier ICs were combined to demonstrate 823mW at 216GHz. These results represent a 3-4x improvement in available output power from solid state sources operating in this frequency range.

The availability of high-power broadband sources is already advancing the state-of-the art in THz signal generation using diode multiplier chains. Virginia Diodes Inc. (VDI) has packaged a 220GHz 3-stage InP PA and demonstrated > 10 mW output power over the entire WR04 waveguide band (170-260GHz) [27]. This work has led to the development of the first ever 1.1-1.5 THz vector network analyzer extender system covering the full WR0.65 waveguide band [28].

Fig. 4. 470mW 81 GHz power amplifier utilizing sub-quarter wavelength baluns for series power combining. Chip area: 1.08x0.98mm². From [22]

Fig. 5. 3-stage 16-way combined power amplifier that demonstrated record 208mW output power at 210GHz. Chip area: 1.95x1.9mm². From [23]
B. THz Monolithic Integrated Circuits

At THz frequencies, chip-to-waveguide transitions are lossy and custom waveguide block assemblies are expensive to build and have large form factors. Single-chip transceiver solutions are desirable but require high levels integration without compromising performance. In our THz integrated circuit demonstrations, our goal has been to extend the complexity and functionality of modern microwave RFIC design to THz frequencies. We have developed building block ICs (amplifiers, oscillators, mixers, frequency dividers, PLLs) operating at THz frequencies and have been integrating them with increasing levels of complexity to demonstrate transceiver functionality.

Fundamental fixed frequency and voltage controlled oscillators have been realized operating to 570GHz in our 250nm HBT process [29] and to >670GHz in our 130nm process (Fig. 6). The oscillators use a differential series-tuned feedback topology with an integrated output buffer. For voltage controlled designs, the HBT base-collector diode is used as a tuning varactor. Differential topologies are attractive for THz circuit designs as the virtual ground along the plane of symmetry can minimize grounding inductance and makes designs insensitive to common-mode impedance such as base-emitter bias circuitry. We have used these types of designs extensively in our transceiver circuit designs.

The 130nm HBT process has also been used to demonstrate amplifier circuits operating at >600GHz. Fig. 7 shows the measured on-wafer S-parameters and a chip photograph of a 9-stage common-base amplifier [7]. The design uses thin-film grounded coplanar waveguide (CPW-G) wiring and consumes 168mW of DC power. The amplifier has 24dB gain at 670GHz and a 3dB bandwidth of >100GHz. The design demonstrated a saturated output power of -4.0dBm at 585GHz. Higher output power designs have been realized using a differential common-base topology that 4-way power combines dual differential branches [8]. On-wafer testing of these designs exhibited greater than 20dB gain to 620GHz and 2.8dBm saturated output power at 585GHz.

Integrated fundamental phase locked loop circuits (PLLs) have been demonstrated for on-wafer frequency synthesis. These include designs at 300GHz and 220GHz [30], [31]. The PLL designs use the previously described fundamental oscillators followed by a 2:1 dynamic frequency divider. Phase compensation is performed using a 5th-order sub-harmonic phase detector and the active loop filter is included on-chip. The 220GHz PLL [31], demonstrates an ~5GHz locking range and a measured phase-noise of -83dBc/Hz at a 100kHz offset with 463mW of DC power consumption.

We have incorporated the PLL circuits with sub-harmonic up-converting and down-converting mixers to realize 600GHz transmitter and receiver circuits. In [32] we reported an integrated 630GHz transmitter IC incorporating a 210GHz PLL and a 3rd-order sub-harmonic mixer. We have since incorporated output stage amplifiers on similar transmitter designs. Fig. 8 shows a chip photograph of a 590GHz transmitter circuit that incorporates a 195GHz PLL and the dual branch differential amplifier described in [8]. The IC contains 167-130nm HBTs and consumes 1.08W of DC power. The chip dimensions are 1.95x0.7mm² (Fig. 8). The close-in
output spectrum of a 590GHz down-converted transmitted signal is also shown in Fig. 8. The measurement is made using a down-converting sub-harmonic mixer. The phase-noise of the transmitted signal is -76dBc at a 100kHz offset. Power measurements were performed separately and the estimated RF output power is -2dBm.

IV. HETEROGENEOUS INTEGRATION WITH SI CMOS

Heterogeneous integration of high performance InP HBTs with fine-line Si CMOS significantly expands the potential applications of the technology by adding the functionality of calibration, memory and computation. As the technologies are brought physically closer together to decrease heterogeneous interconnect lengths, the challenge of maintaining the monolithic device performance increases. 3D integration offers an attractive approach for achieving transistor-level integration of InP and Si technologies. Utilizing Ziptronix Direct Bond Interconnect (DBI®) process [33] we have developed a wafer-scale low-temperature oxide-to-oxide bonding process to integrate InP and Si CMOS [14].

Fig. 9 shows a cross-section schematic of the integrated InP HBT/Si CMOS technology. In our 3D integration process, the front-end-of-line (FEOL) transistor process flows are unchanged. We have successfully integrated Teledyne’s 250nm InP HBT process with GlobalFoundries 130nm CMOS. The DBI® process utilizes a low temperature oxide-to-oxide bonding process. This is ideal for heterogeneous integration as alignment mismatches from differences in the coefficient of temperature expansion (CTE) between dissimilar materials can be minimized. The oxide-to-oxide bond requires chemical mechanical polishing (CMP) of the wafers to achieve a low roughness macroscopically planar surface. While CMP steps are common in silicon backend-of-line (BEOL) processes, they are typically not used in III-V interconnect processes like our previously described BCB-based interconnects. We have developed a final SiO₂ deposition and CMP planarization process on top of our BCB interconnect stack that can achieve the required planarization for wafer-scale bonding.

After SiO₂ planarization, embedded Cu-based contacts are defined on the wafer to form the hybrid bond surface. Matching contacts are patterned on the mating Si CMOS wafer. For this work laser coring was used to produce two 100mm Si wafers from the starting 200mm substrate. These wafers could then be face-to-face bonded with the 100nm InP wafer at room temperature using the DBI® process. After bonding, the Si substrate is thinned to ~10µm thickness (excluding the BEOL layers) using a grinding and CMP process. The backside of the Si wafer is then patterned and etched to reveal the CMOS interconnect metal. A final metallization layer is then patterned on the back of the thinned Si wafer. A key advantage of our integration approach is the final “CMOS-on-top” architecture where the chip I/O’s are routed through the Si substrate. This allows for efficient heat-sinking of the high power density InP HBTs. In many other integration approaches ([11], [12], [13]), the InP devices are placed on-top of the CMOS interconnect stack requiring the addition of thermal vias in the CMOS interconnects which can limit integration density and increase transistor thermal resistance. The heat removal path for InP devices in our approach is equivalent to a monolithic InP wafer and extracted thermal resistances for integrated and non-integrated devices are equivalent. Of course, this architecture moves the challenges of heat removal to the Si CMOS layer. However, in many heterogeneous integration applications, the CMOS is being utilized for low power density applications like memory or low-frequency digital logic for control and similar heat-sinking constraints are not present in a “CMOS-on-top” structure.

Fig. 10 shows a FIB/SEM cross-section of the 3D integrated InP/Si stack. The heterogeneous DBI® interconnect pitch is 5µm. Pre- and post- bond testing confirms that no degradation is observed in the DC or RF performance of the integrated devices. Fig. 11 shows the measured input characteristics of a 130nm NMOS device before and after integration and Fig. 12 shows the extracted RF figures-of-merit of a 250nm InP HBT before and after integration. Heterogeneous interconnect chains with up to 5000 interconnects between the InP and Si metallization layers have been characterized. From these measurements, the estimated resistance is ~0.04Ω per interconnect. S-parameter measurements of back-to-back interconnects show <0.2dB of insertion loss to 65GHz.

A fully-integrated Cadence-based design kit has been developed for the technology. This includes design rules for
Although recent efforts have been focused on circuit demonstrations and increasing levels of integration, paths are still available for further increasing InP HBT bandwidth. Ohmic contact resistivities targeted for 64nm and 32nm HBT scaling generations have been demonstrated [35] and established transistor process flows should continue to scale.

ACKNOWLEDGMENT

The author’s would like to thank R. Patti and S. Hong at Tezzaron Semiconductor, C. Pettaway at Novati Technologies and G. Fountain and P. Enquist at Ziptronix for support in the heterogeneous integration process development. IC processing and heterogeneous integration were performed at the Teledyne Scientific and Novati Technologies cleanrooms, respectively. Portions of this work were supported by the Defense Advanced Research Projects Agency (DARPA) under DARPA CMO Contract No. HR0011-09-C-0060 and AFRL Contract No. FA8650-13-C-7313. Program support from Dr. Dev Palmer (DARPA), Dr. Dan Green (DARPA), Dr. Alfred Hung (ARL) and Dr. Robert Fitch (AFRL) is gratefully acknowledged. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

REFERENCES


InP HBT IC technology for THz frequencies,

- Z. Griffith et al. “71-95GHz (23%-40% PAE) and 96-120GHz (19-22% PAE) high efficiency 100-130mW power amplifiers in InP HBT,” to be presented 2016 IEEE International Microwave Symposium, May 2016, San Francisco, CA.