

Low Cost, Highly Flexible Complementary Bipolar Transistors Compatible with 0.18 or 0.13 μm CMOS Technology

E. J. Preisler, L. Lao, J. Zheng, P. Hurwitz and M. Racanelli

Jazz Semiconductor, Newport Beach, CA 92660, USA

Abstract: Both NPN and PNP bipolar transistors are modularly integrated at low cost into standard 0.18 μm and 0.13 μm CMOS process flows. The bipolar modules are of low complexity and thus cost, using only 4 dedicated masks for the NPN and only 2 for the PNP devices. The resulting devices cover an extremely wide range of application space. Devices range from around 2V BV_{CEO} and 115GHz F_T to 12V BV_{CEO} and 12 GHz F_T . Multiple device types can be co-integrated with the addition of simple implant masking steps.

I. INTRODUCTION

Complementary BiCMOS (hereafter CBiCMOS) processes can both provide flexibility in circuit design as well as allowing circuit topologies impossible in standard CMOS or NPN-only BiCMOS. An excellent summary of the uses of CBiCMOS is given by Monticelli [1]. Recently, higher performance SiGe BiCMOS devices have begun to be integrated with PNP devices to enable greater bandwidth in RF applications [2]-[4]. The work described below seeks to expand on the idea of SiGe CBiCMOS by exploring the limits of the flexibility of this process, within the constraint of considering only very low complexity, low mask-adder bipolar modules.

II. PROCESS

The basic process flow used to produce the devices measured for this study was as follows (See Fig. 1).

A deep N-Well mask is patterned and implanted before dielectric isolation to form junction isolation for the PNP devices and for triple well isolation of NFETs. This is followed by standard shallow trench isolation. Then follow standard CMOS process steps: wells, gate oxidation, gate poly deposition and patterning, then finally LDD and pocket implants.

At this point, the PNP module begins with collector and base implantation. Then the CMOS spacer is

deposited and patterned to define the PNP emitter windows and selectively implanted collectors. The details of the PNP process flow are outlined in reference [5].

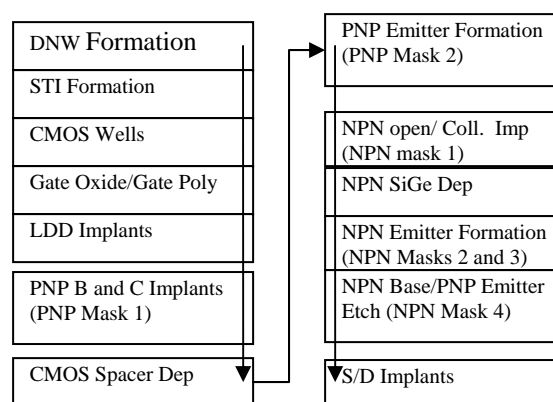


Figure 1: Schematic representation of the front end complementary BiCMOS process flow in either 0.13 or 0.18 μm technology nodes.

At this point the NPN module begins with patterning and etching of a spacer clear out mask which will define the active epitaxial SiGe regions for the NPN transistors. In some cases an implant is also performed through this mask to define the NPN collector or differentiate one type of NPN transistor from another. Next follows a standard quasi-self-aligned (QSA) SiGe HBT flow comprised of SiGe deposition, emitter window stack deposition and etching, emitter poly deposition and etching and finally base poly etching.

The emitter of the PNP device is partially formed by the NPN SiGe base deposition. It is patterned at the same time as the base poly is patterned in the NPN process, with the base poly etch serving as the CMOS spacer etch, NPN base poly etch and PNP emitter poly etch.

The NPN collector is formed by ion implantation only. The subcollector is formed by high energy implantation either through a dedicated mask or through the spacer clear out mask discussed above; similar to the process flow described in [6]. The collector sinker shares CMOS implants so that no dedicated mask is necessary. The intrinsic collector implants are implanted through either a dedicated mask or the spacer clear out mask. Two distinct devices, one in the 85 GHz ; 2-2.5V BV_{CEO} range and one in the 67 GHz ; 3-3.5V BV_{CEO} range, can be co-fabricated with a total of 4 additional masks (relative to the parent CMOS process). The higher voltage or higher speed devices require an additional mask in order to be co-integrated with the other devices. Of course any single device can be fabricated on its own with only 4 masks.

III. DEVICE PERFORMANCE

Figure 2 is a plot of F_T vs. BV_{CEO} for several NPN devices. The devices tested all had emitter areas of $0.28 \times 10 \mu\text{m}$ and had single emitter, double base, double collector configurations. The F_T was extracted from h_{21} at 10 GHz for all devices below 5V BV_{CEO} and at 5 GHz for the rest. V_{CB} was held at 1V in all cases so that the data for the highest breakdown devices somewhat underestimate the highest possible F_T achievable. Most of the devices shown fall between the constant $F_T \times BV_{CEO}$ contours of 200 and 225 GHz \times V. The higher voltage devices are competitive with full flow devices (buried subcollector etc.) in $F_T \times BV_{CEO}$ [7]. F_{MAX} ranges from 90 GHz to as high as 130 GHz. BV_{CBO} ranges from about 6 V up to nearly 20V for the devices with $BV_{CEO} \sim 10V$. The limitation at high speed is the tradeoff with F_{MAX} . At 90 GHz or so $F_T \sim F_{MAX}$, and going any higher in F_T causes F_{MAX} to be less than F_T . This is a consequence of the low-cost QSA NPN process which results in both a somewhat higher R_B and C_{BC} than fully self aligned (FSA) devices. The limitation at the high voltage end is process-related. For an implanted-subcollector process, the energies required to achieve very high breakdown start to become prohibitive due to photoresist thickness concerns, the need to switch to doubly ionized dopant sources, and possible damage to the silicon near the CB junction. Higher breakdown voltages can still be achieved within the implanted subcollector process by adding a dedicated mask near the beginning of the process flow so that adequate thermal budget is available to repair the implant damage. Current experiments along these lines are underway.

Figure 3 is a plot of F_T and F_{MAX} vs. $-J_C$ for a typical PNP device. Figure 4 is a plot of F_T vs. BV_{CEO} for

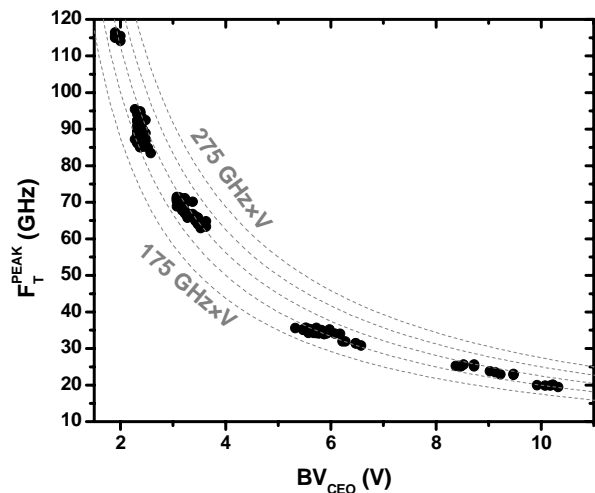


Fig. 2. F_T vs. BV_{CEO} for the NPN devices processed with 0.13 μm CMOS technology. Peak F_T is taken with $V_{CB} = 1V$ in all cases. All devices have an emitter area of $0.28 \times 10 \mu\text{m}$

several PNP devices. All the filled-circle data is from devices with emitter areas of $0.8 \times 10 \mu\text{m}$ and single emitter, double base, double collector configuration. The F_T was extracted from h_{21} at 3GHz in all cases. V_{CB} was held at -4V in all cases. The open circles at the high and low ends of the data indicated that slightly non-standard layouts were needed to create the devices that produced this data. However all devices were created using the same 2 mask-adder process flow. Most devices fall between the constant $F_T \times |BV_{CEO}|$ contours of 100 and 125 GHz \times V. The only limitation at the high speed end seems to be the deterioration in $F_T \times |BV_{CEO}|$. As in the NPN case, the limitation at the high voltage end is due to the energy of the implants required to achieve these breakdown voltages.

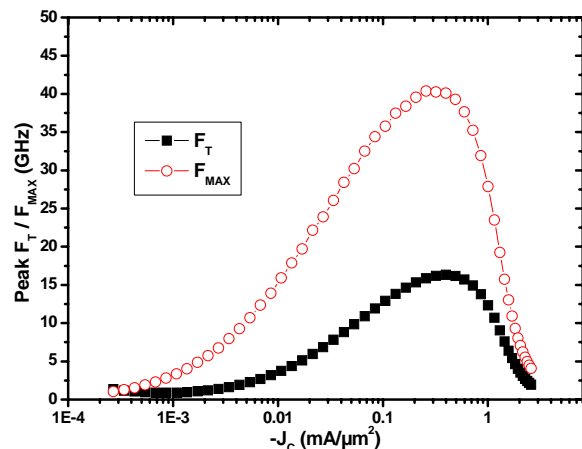


Figure 3. F_T and F_{MAX} vs. $-J_C$ for a typical PNP device processed with 0.18 μm CMOS technology. F_T and F_{MAX} are extracted from h_{21} at 3GHz and U at 3GHz respectively. V_{CB} was -4V. The device had an emitter area of $0.8 \times 10 \mu\text{m}$.

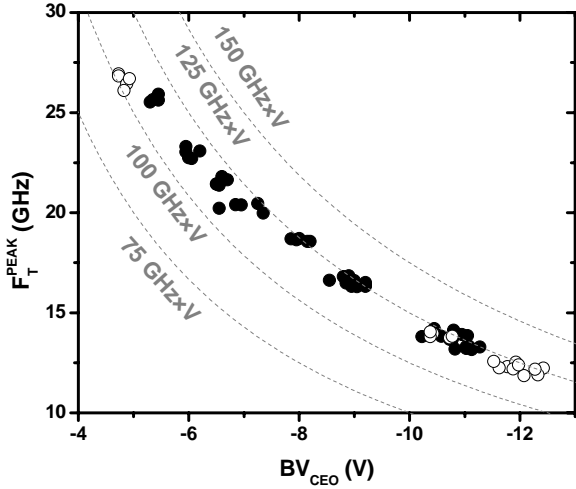


Figure 4. F_T vs. BV_{CEO} for the PNP devices processed with 0.18 μm CMOS technology. Peak F_T is taken with $V_{CB} = -4\text{V}$ in all cases. All filled-circle devices had emitter areas of $0.8 \times 10 \mu\text{m}$ with a single emitter finger, double base and double collector fingers.

IV: EFFECTS OF INTEGRATION

Experiments were run to verify the manufacturability of the bipolar devices within 0.13 and 0.18 μm CMOS technology and also to see the effect that the PNP module has on the NPN module.

Firstly splits were run with and without an NPN module in 0.18 and 0.13 μm CMOS process flows. The parametric shifts seen in 0.18 μm CMOS are well established in Jazz' mature 0.18 μm CMOS/BICMOS processes but the NPN thermal budget causes significant shifts in the CMOS for the 0.13 μm CMOS flow. Still, all CMOS data can be almost perfectly matched with simple implant adjustments (see figure 5 for some examples of matching data). Thus the low-cost NPN module is completely modular in 0.18 μm technology and "pseudo-modular" in 0.13 μm in the sense that no CMOS process steps need to be altered other than an adjustment of ion-implantation parameters.

Next the corollary effect of the CMOS process on the NPN devices was investigated. Figure 6 shows F_T vs. J_c plots of high speed NPNs built within the two different CMOS processes. With no adjustments the device processed with 0.13 μm CMOS is about 5% slower than the device processed with 0.18 μm CMOS. This is mostly due to the lower thermal budget of the 0.13 μm process which, all other things being equal, lessens emitter diffusion and thus widens the base. The third dataset in figure 6 shows another device processed with 0.13 μm CMOS that nearly matches the

performance of the 0.18 μm device with only a slight process change that doesn't affect any other devices.

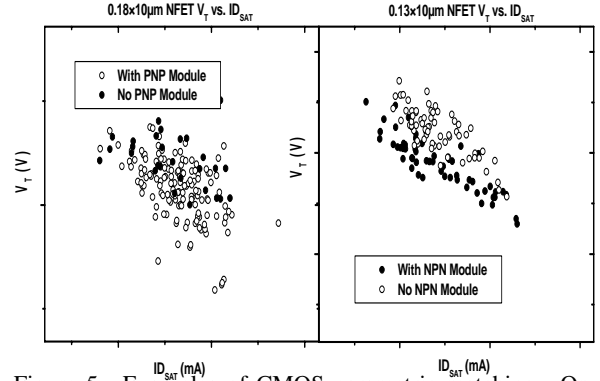


Figure 5. Examples of CMOS parametric matching. On the left is shown V_T vs. ID_{SAT} scatter data for 0.18 μm NFEETs with and without the PNP module processing. On the right is shown V_T vs. ID_{SAT} scatter data for 0.13 μm NFEETs with and without the NPN module processing.

The PNP module does not employ any significant thermal steps and thus does not significantly affect the CMOS devices (this was specifically verified only in the 0.18 μm technology; see figure 5). However the nature of the PNP integration may affect the NPN devices. To evaluate this independently, the PNP was integrated along with Jazz' most advanced in-production SiGe BiCMOS process; SBC18 [8]. The yield of large NPN arrays run with the PNP module was initially poor but some small process changes quickly recovered the NPN yield to be typically within a few percent of that of NPN-only lots.

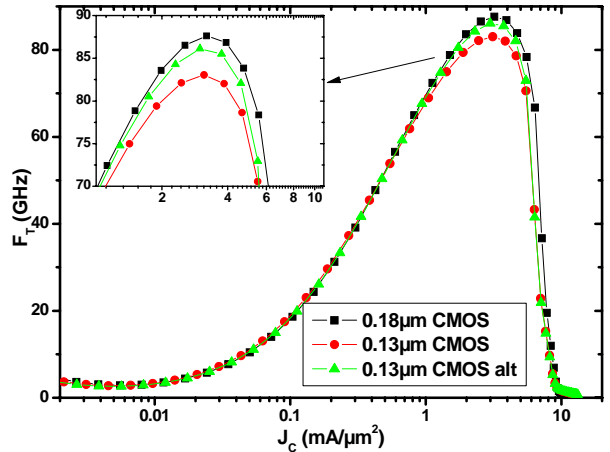


Figure 6. F_T - J_c plots for identical NPN devices processed with 0.13 or 0.18 μm CMOS. The device indicated by the triangle data was processed with 0.13 μm CMOS and only one small process change relative to the 0.18 μm device.

Figure 7 shows RF data for a SBC18 standard breakdown NPN device with and without the PNP process. There is a small shift that goes along with a slight parametric shift in the NPN data in general, but shifts are small compared to natural process variation.

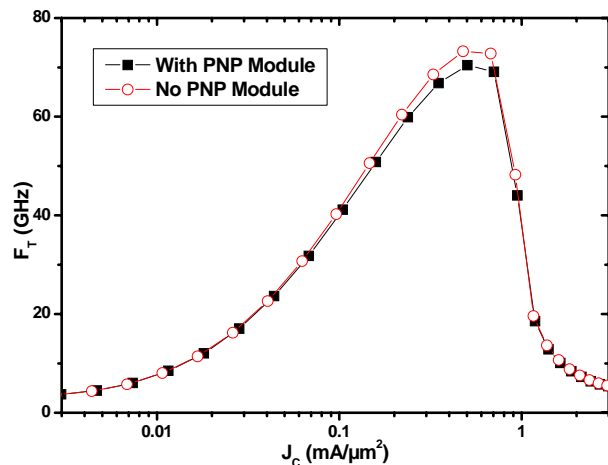


Figure 7. F_T vs. J_C plot for SBC18 standard voltage NPN devices (0.18 μm CMOS technology) processed with and without the PNP module.

Finally the effect of CMOS process node on the PNP is shown in Figure 8. The device processed in 0.13 μm CMOS technology suffers very slightly in F_T but there is a corresponding increase in BV_{CEO} such that the $F_T \times |BV_{CEO}|$ product is slightly superior in 0.13 μm technology. While the F_T shift is not large compared to natural process variation, the BV_{CEO} shift is. Exact matching of the DC parameters can be achieved if necessary with a slight change in PNP process parameters.

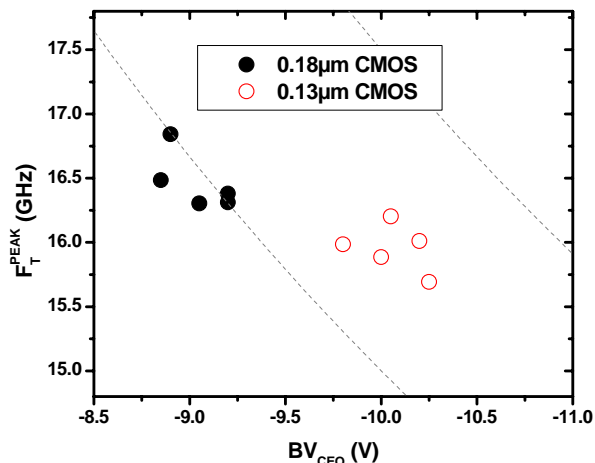


Figure 8. F_T vs BV_{CEO} plot for PNP devices processed with 0.13 and 0.18 μm CMOS.

VII. CONCLUSION

NPN and PNP bipolar devices covering a very wide range of application space have been fabricated using both 0.13 μm and 0.18 μm CMOS process backbones. The low-cost NPN devices are completely modular in 0.18 μm technology and pseudo-modular in 0.13 μm technology. They require a minimum of 4 additional masks to create various devices with F_T as high as 115 GHz or BV_{CEO} as high as 10V. The PNP devices are completely modular in 0.18 (and likely also 0.13 μm) technology as long as triple well isolation is part of the standard CMOS process offering. They require 2 additional masks to create devices with F_T as high as 27 GHz or BV_{CEO} as high as 12V.

REFERENCES

- [1] D. Monticelli *Proc. 2004 BCTM*, p. 21, (2004).
- [2] B. Heinemann et al. *Proc. 2003 IEDM*, p. 117, (2003)
- [3] B.T. Voegeli et al., *Proc. 2005 BCTM*, p. 136, (2005).
- [4] B. Heinemann et al. *Proc. 2006 ISTD*, p. S153, (2007)
- [5] U.S. Patent #6933202
- [6] E. J. Preisler et al. *Proc. 2007 BCTM*, p. 202, (2007).
- [7] M. Racanelli et al., *Proc. 2001 IEDM*, p. 336, (2001).
- [8] Ibid.