This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

A Perspective on Symmetric Lateral Bipolar Transistors on SOI as a Complementary Bipolar Logic Technology

Tak H. Ning, *Life Fellow, IEEE,* and Jin Cai, *Member, IEEE*

***Abstract*— Recently published reports suggest that symmetric lateral bipolar transistors on semiconductor-on-insulator (SOI) is CMOS compatible in fabrication process, and can be much denser than CMOS due to their much larger (5 to 10x larger) drive-current capability. When used in traditional bipolar circuits, SOI bipolar offers much lower power dissipation and/or much higher maximum speed. With both NPN and PNP devices of comparable characteristics, SOI lateral bipolar suggests the possibility of complementary bipolar (CBipolar) circuits in configurations analogous to CMOS. In this paper, the performance vs. power dissipation of CBipolar circuits is examined using analytic equations. It is shown that for CBipolar to be superior to CMOS in both performance and power dissipation, narrow-gap-base heterojunction structures, such as Si emitter with Ge base or Si emitter with SiGe base, are required.**

Index Terms— CBipolar, Complementary bipolar, SOI bipolar, symmetric lateral bipolar

I. Introduction

THE idea of a symmetric lateral Si-on-insulator (Si-OI) bipolar transistor with a self-aligned base contact located on top of the intrinsic-base region, and base widths of about 2 μm, was first demonstrated almost thirty years ago [1]. With lithography capability now at 22 nm in manufacturing, it is possible to fabricate both NPN and PNP Si-OI lateral bipolar transistors (Fig. 1) with base widths much less than 100 nm using CMOS-like processes [2]. Measured data show that Si lateral bipolar devices have drive-current capability much higher than CMOS [3, 4], while model studies suggest that they are scalable in lateral dimensions like CMOS and could have f_{max} > 1 THz [5].

The emitter/collector symmetry makes SOI lateral bipolar transistors immune to base push out (into the collector region) and suitable for circuits that involve operation in deep saturation or in both forward-active (emitter–base diode forward biased) and reverse-active (collector–base diode forward biased) modes. The result is significantly reduced power supply voltage for conventional bipolar circuits, and the possibility of complementary bipolar (CBipolar) inverters (Fig. 2) operating with a power supply voltage *Vcc* equal to the emitter–base forward bias voltage, V_{BE}, needed to achieve the target on current for the circuit.

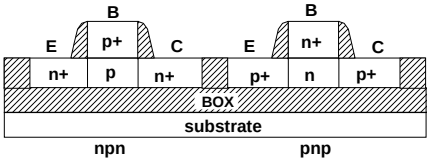

Fig. 1. Schematic illustration of the structure of complementary symmetric lateral bipolar transistors on SOI. (After [2])

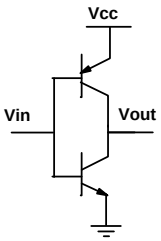

Fig. 2. Circuit schematic of a CBipolar inverter.

The operation of CBipolar inverters built using the transistor structure in Fig. 1 has been demonstrated experimentally [2, 3], and the concept of CMOS-like CBipolar circuits has been around for a long time [6]. It is an objective of this paper to develop insights into the operation of CBipolar circuits, using analytic current equations appropriate for SOI symmetric lateral bipolar transistors [4, 5]. Another objective is to examine the performance and power dissipation characteristics of CBipolar inverters, to see if CBipolar has the potential as an attractive digital circuit technology.

II. Symmetric Lateral Bipolar Transistors on SOI

As will be shown in Section III below, the ideal bipolar transistor characteristics for CBipolar applications are high on current (collector current) at low power supply voltage (*Vcc*), negligibly low off current at standby (V_{BE} = 0 and V_{CE} = *Vcc*), and very large current gain (β >> 100) when the transistors are turned fully on (at V_{BE} = *Vcc*). So far CBipolar (integrated NPN and PNP) has been reported only for Si-OI [2-4]. In this section we examine the properties of SOI lateral bipolar transistors as they apply to the operation of CBipolar circuits, using the reported data to illustrate both the status of CBipolar technology as well as the direction for future technology development. Several assumptions about the transistors are needed to make modeling using analytic equations tractable.

The authors are with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 (e-mail: ningth@us.ibm.com; jincai@us.ibm.com).

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

These assumptions are discussed and clearly stated as they are made.

A. *Typical Integrated Si-OI NPN and PNP*

Figures 3 and 4 show the Gummel plots for typical integrated Si-OI NPN and PNP devices with E/C regions formed by As or B implantation. Both devices show ideal currents (varying at 60 mV/decade) for voltages up to about 0.9 V. The current saturation at larger voltage is due to a combination of high-injection effect and parasitic resistances [4]. The PNP currents clearly saturate at a lower level than those of the NPN. This is due to the fact that doping by boron implantation results in a more graded E/C junction and higher E/C series resistance for the PNP device. From device physics considerations, a PNP device and an NPN device having the same doping profile should have about the same *I–V* characteristics. In the following, we focus our discussion on the NPN *I–V* characteristics (Fig. 3), and simply ***assume that comparable PNP devices will be available***.

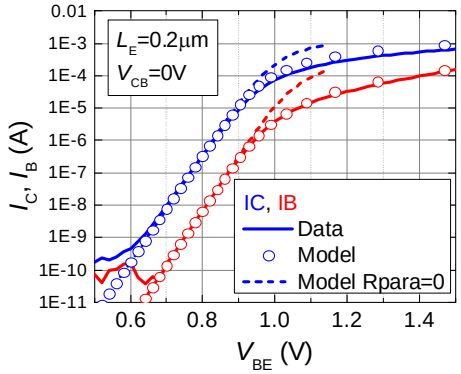

Fig. 3. Typical Si-OI NPN device. The device has T_{si} = 60 nm and N_E = N_C = 4E20/cm^3 formed by As implantation. The model currents were calculated using measured value of r_e = 267 Ω (see Fig. 7 for transistor equivalent circuit). Dash lines show calculated intrinsic device currents with no parasitic resistance. (After [4])

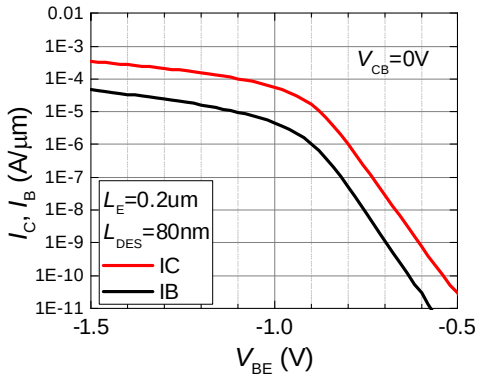

Fig. 4. Typical Si-OI PNP device. The device was integrated with the NPN in Fig. 3, and has N_E = N_C = 4E20/cm^3 formed by B implantation. (After [4])

In CBipolar inverter operation, the transistor in the off state is biased with V_{BE} = 0 and V_{CE} = *Vcc*. Fig. 5 is a plot of current as a function V_{BE} at fixed V_{CE} for the same NPN as in Fig. 3. It shows an off current of 0.1 μA/μm at V_{CE} = 1.0 V, increasing with V_{CE} to 0.7 μA/μm at V_{CE} = 1.5 V. Such levels of off current are comparable to those of state-of-the-art high-performance CMOS, but much too high for applications where low standby power is critical.

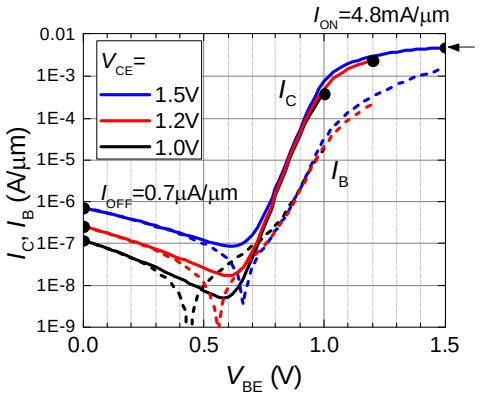

Fig. 5. Current characteristics for the same NPN as in Fig. 3 taken at fixed V_{CE} of 1.0, 1.2 and 1.5 V. (After [4])

The off currents at V_{BE} = 0 in Fig. 5 are caused by the leakage current in the reverse-biased B-C diode, as evidenced by the negative base current being equal to the collector current. Fortunately, reverse-bias diode leakage current is a function of the diode fabrication process. It is possible to obtain B–E and B–C diodes with negligible reverse-bias leakage currents. As an example, Fig. 6 shows the measured off current (i.e. collector current as a function of V_{CE} taken at V_{BE} = 0) of a Si-OI NPN device designed to have low off current. Figure 6 suggests an off current, including current due to instrument noise, of about 10 pA/μm at V_{CE} = 1.0 V. Such small off currents are low even by CMOS standards. Also, as will be shown later in Section III-B, such small device off currents can be ignored in consideration of the operation of CBipolar standby power dissipation. Therefore, in the rest of this paper, ***device off currents are assumed to be negligible and ignored completely***.

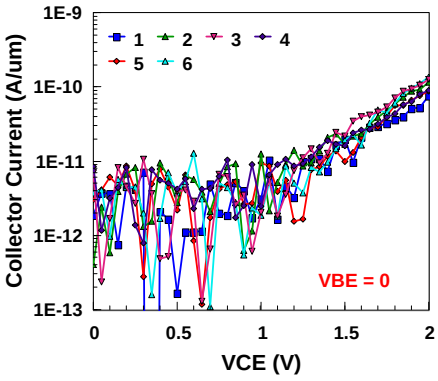

Fig. 6. Measured collector current per unit of L_E at V_{BE} = 0 as a function of V_{CE} for NPN transistors fabricated in the same experiment as those in Figs. 3 and 5, but using an E/C process designed to reduce B-E diode and B-C diode leakage current. The off current, including instrument noise, is 10 pA at V_{CE} < 1.0 V.

The measured base current in Fig. 3 behaves ideally, increasing with V_{BE} at 60 mV/decade, starting at less than 100 pA. The measured base current in Fig. 4 behaves ideally starting at less than 10 pA. The 60-mV/decade behavior indicates that the measured base current is the intrinsic base

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

current of the transistor, which will be discussed in the next subsection. However, it is not uncommon to find devices showing "excess base current", i.e. a base current component in addition to the intrinsic base current, in measured Gummel plots due to recombination in the E–B diode space-charge region. This recombination base current component can be recognized readily from its $\exp(qV_{BE}/2kT)$ dependence on V_{BE}. In the case of Si-OI devices, excess base current due to recombination can be readily reduced to a negligible level by process optimization, as indicated in the devices in Figs. 3 and 4.

In any case, as will be shown later (Section III-B), the standby current is determined by the base current when the transistor is turned on fully at $V_{BE} = Vcc$. At $V_{BE} = Vcc$, the excess base current due to recombination is negligible, due to its $\exp(qV_{BE}/2kT)$ dependence, compared with the intrinsic base current which varies as $\exp(qV_{BE}/kT)$. Therefore, for simplicity, we ***ignore completely the recombination base current component in this study***.

B. Analytic Equations for Collector and Base Currents

With both the C–B diode reverse-bias leakage current and the excess base current due to recombination assumed to be negligible, the Ebers-Moll model currents for a symmetric lateral NPN transistor have the forms [4, 5]:

$$I_C = I_{C0}(e^{qV'_{BE}/kT} - 1) - (I_{C0} + I_{B0})(e^{qV'_{BC}/kT} - 1), \quad (1)$$

and

$$I_B = I_{B0}(e^{qV'_{BE}/kT} - 1) + I_{B0}(e^{qV'_{BC}/kT} - 1). \quad (2)$$

Note that (1) and (2) are for an NPN transistor where all the device terminal voltages are positive quantities, and V_{BE} (= V_B – V_E) is positive. A corresponding set of equations can be written for a PNP transistor where all the device terminal voltages are negative quantities (with respect to the emitter terminal), and V_{BE} (= V_B – V_E) is negative. The PNP equations have ($-V_{BE}$) and ($-V_{BC}$) in the exponentials. To avoid possible confusion, we only show explicitly the equations that are applicable to NPN in this paper.

In general, the collector saturation current I_{C0} is a function of emitter–base junction voltage V'_{BE}, and is given by [4]

$$I_{C0}(V'_{BE}) = \frac{A_E q D_{nB} n_{ieB}^2}{N_B W_B} \times \left[1 + \frac{1}{4}\left(\sqrt{1 + \frac{4n_{ieB}^2 \exp(qV'_{BE}/kT)}{N_B^2}} - 1\right)\right]^{-1}, \quad (3)$$

where A_E is the emitter area, D_{nB} is the electron diffusion coefficient in the base region, n_{ieB} is the effective intrinsic carrier density in the base region, N_B is the base doping concentration and W_B is the quasi-neutral base width. Eq. (3) is valid for all injection levels, including high injection where the electron density in the p-type base, n_{pB}, is larger than N_B.

At low injections where $n_{pB} << N_B$, (3) reduces to the more familiar form of

$$I_{C0} = \frac{A_E q D_{nB} n_{ieB}^2}{N_B W_B}, \text{ (low injection)} \quad (4)$$

with no explicit dependence on V'_{BE}. Eq. (3) implies that the V'_{BE} value below which (4) is valid increases with N_B. For devices with N_B = 1E19/cm^3, the low-injection approximation is valid for V'_{BE} up to 1.0 V [4].

For the base current, with E/C regions doping typically larger than 1E20/cm^3, the low-injection approximation is valid for all practical V'_{BE} values, so we have

$$I_{B0} = A_E q D_{pE} n_{ieE}^2 / W_E N_E, \quad (5)$$

where n_{ieE} is the effective intrinsic carrier density in the emitter, D_{pE} is the diffusion coefficient for holes in the emitter, N_E is the emitter doping concentration, and W_E is an "effective emitter junction depth" parameter (see Fig. 7) determined from fitting to the measured base current [4, 5].

It should be noted that (3) varies as $\exp(-qV'_{BE}/2kT)$ at large V'_{BE} where $n_{pB} >> N_B$. That is, the collector current increases as $\exp(qV'_{BE}/kT)$ until it approaches high-injection condition, and then increases much more slowly, only as $\exp(qV'_{BE}/2kT)$ at very high injection levels. As a result, the current gain, which is the ratio of I_{C0}/I_{B0}, instead of being constant, decreases at high injection levels [4], just as the data in Figs. 3 and 5 suggest.

C. Parasitic Series Resistances

Equations (1) to (3) are in terms of the intrinsic B–E and B–C junction voltages V'_{BE} and V'_{BC}, respectively. In experiments, the voltages are applied to the device terminals. The transistor geometrical parameters and the device junction and terminal voltages are illustrated in Fig. 7.

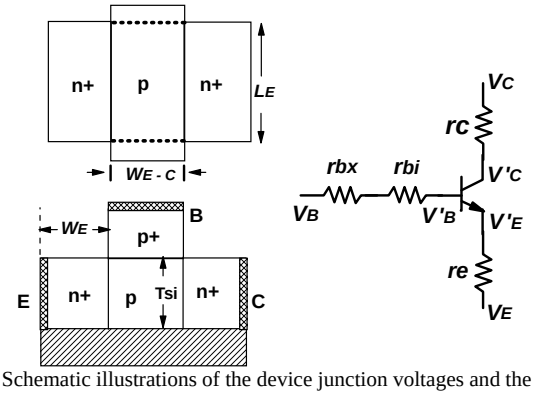

Fig. 7. Schematic illustrations of the device junction voltages and the terminal voltages of a bipolar transistor. r_e and r_c are the emitter and collector series resistances, respectively. r_{bx} is the extrinsic-base series resistance, and r_{bi} is the intrinsic-base resistance. (After [5])

The resistances r_e and (r_{bx} + r_{bi}) cause the Gummel plot currents to saturate at large V_{BE}. In Fig. 3, the saturation of the base current is due to r_e and (r_{bx} + r_{bi}), but mostly due to r_e because the base current is much smaller than the collector current. The saturation of the collector current is due to a

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

combination of the resistances, r_e and (r_{bx} + r_{bi}), and the high-injection effect discussed in the subsection above.

The intrinsic-base resistance r_{bi} is an integral part of a transistor design, and is given by [5]

$$r_{bi} = (T_{si} / 3L_E)\rho_B / W_B, \quad (6)$$

where ρ_B is the base resistivity, T_{si} is the semiconductor layer thickness, and L_E is the emitter stripe length (CMOS device width) indicated in Fig. 7. For a transistor of given areal dimensions, the collector current increases with T_{si}. The f_T of a transistor is relatively insensitive to T_{si}. On the other hand, the f_{max} of a transistor decreases rapidly with increase in T_{si} due to the increased r_{bi} with T_{si} [5].

The resistances r_e and r_{bx} depend on the device layout and fabrication process. As an example, Fig. 8 shows two possible layouts for a lateral bipolar transistor. It is clear that the layout on the right has smaller r_{bx}, and could result in device/circuit density advantage as well. All the data shown in this paper were taken with device layout in the CMOS style.

It should be noted that the emitter–collector separation W_{E-C}, being the sum of the base width W_B and the space-charge region widths on both sides of the base (see Fig. 7), is larger than W_B. For a given W_B, W_{E-C} is a function of the base doping concentration. For the transistor in Fig. 3, N_B is 2.5E18/cm^3, W_B is 10.3 nm and W_{E-C} is 57.3 nm. In general, with W_{E-C} significantly larger than W_B, there is adequate room to locate the metal contact directly on top of the intrinsic base.

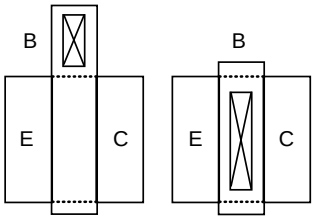

Fig. 8. Schematic illustrations of two possible layouts for SOI lateral bipolar transistors. The layout on the left is the same as CMOS, with metal contact via to the extrinsic base located not on top of the intrinsic base. The layout on the right has metal contact via to the extrinsic base located directly above the intrinsic-base region.

The NPN transistor in Fig. 3 has an emitter length of 0.2 μm and a measured emitter series resistance of 267 Ω, which translates to 55 Ω-μm. Such series resistance is quite a bit smaller than CMOS source-drain series resistances which are typically larger than 150 Ω-μm. In the case of CMOS, source-drain resistances are dominated by the "source-drain extensions" which are very shallow. In lateral bipolar devices, there are no "shallow emitter extension" regions.

At any rate, the data in Fig. 3 show that the currents are quite ideal, rising at 60 mV/decade, for V_{BE} up to about 1.0 V. The implication is that for V_{BE} less than 1.0 V, the parasitic resistances are not significant for the device in Fig. 3, at least in terms of device currents. Therefore, as a first order estimate of the operation and performance of CBipolar circuits, we ***ignore all resistances in this study***. For a well developed Si-OI CBipolar technology, the model results should be valid for *Vcc* values up to about 1.0 V.

In theory, the resistances r_e, r_c and r_{bx} could be reduced to a negligible level by process optimization, but the intrinsic-base resistance r_{bi} is an integral part of the device design, with a value given by (6) and is independent of any process optimization. Nonetheless, in our model using analytic equations, ***r_{bi} is assumed to be zero***. The effect of r_{bi} on circuit delay will be estimated and discussed in Section VI.

D. Collector Current and Power Supply Voltage

The quantity $(n_{ieB})^2$ in (4) is proportional to exp(–E_{gB}/kT), where E_{gB} is the bandgap of the base region. Therefore, the collector current has the form

$$I_C \propto \exp[(qV'_{BE} - E_{gB})/kT]. \quad (7)$$

Eq. (7) suggests that for a given device physical structure and dimensions, the voltage needed to obtain a target collector current scales with the energy bandgap of the base region. At this time, there is no report of SOI lateral transistors using semiconductors with bandgap smaller than Si. However, in the field of advanced CMOS, SiGe-IO and Ge-OI are often employed. Both SiGe and Ge have bandgaps smaller than Si. Figure 9 is a plot of sample calculated collector current density as a function of V_{BE} for Si-base and SiGe-base transistors. For the Si-base transistor, the model parameters are consistent with those used in Fig. 3. For the SiGe-base devices, the model parameters are assumed to be the same as the Si-base transistor, which is reasonable as a first-order estimate.

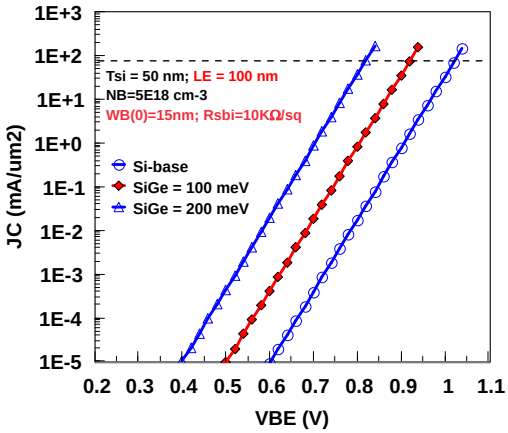

Fig. 9. Calculated collector current density vs. V_{BE} for a Si-OI and two SiGe-OI NPN transistors. All three devices have the same physical dimensions and base doping concentration. For the two SiGe-base devices, in one case the SiGe is assumed to have a bandgap 100 meV smaller than Si, in another case the SiGe is assumed to have a bandgap 200 meV smaller than Si.

Figure 9 clearly shows that if a Si-base device operates with a target collector current at *Vcc* of 1.0 V, a SiGe-base device having a bandgap 200 meV smaller than Si can be operated with the same collector current at *Vcc* of about 0.8 V. Since the bandgap of Ge is 0.46 eV smaller than that of Si, a Ge-base transistor can be operated with the same collector current at *Vcc* of only about 0.54 V. The important point is that the power supply voltage for a CBipolar circuit can be reduced substantially by using SOI of small bandgaps.

E. Current Gain and Emitter Engineering

Once the intrinsic base of a transistor has been designed to deliver a desired collector current at a target *Vcc*, the current gain is determined by the base current which is a function of the emitter parameters only, as indicated in (5). In other

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

words, the current gain of a transistor can be altered by engineering the emitter region.

With the recombination component of the base current assumed to be negligible (see discussion in Section II-A above), (4) and (5) give the current gain of an NPN transistor as

$$\beta = \frac{n_{ieB}^2}{n_{ieE}^2}\frac{D_{nB}}{D_{pE}}\frac{N_E W_E}{N_B W_B} \tag{8}$$

$$= \frac{D_{nB}}{D_{pE}}\frac{N_E W_E}{N_B W_B}\exp\left(\frac{E_{gE}-E_{gB}}{kT}\right)\exp\left(\frac{\Delta E_{gB}-\Delta E_{gE}}{kT}\right),$$

where E_{gE} is the emitter bandgap, ΔE_{gE} and ΔE_{gB} are the apparent bandgap narrowing parameters in the emitter and the base, respectively. For a homojunction transistor, i.e. a transistor with $E_{gE} = E_{gB}$, increasing N_E is an effective method for increasing current gain. However, the effectiveness of increasing N_E is tempered by the effect of heavy doping because the parameter ΔE_{gE} increases with N_E [7]. Reported data suggest that for homojunction lateral Si-OI devices, a current gain of around 50 should be realizable [3, 4].

As will be shown in Section III below, we really need very large current gains ($\beta >> 100$) for CBipolar to be of interest as a digital circuit technology. The most common approach to increase the current gain of a bipolar transistor is to develop a heterojunction device structure where the emitter bandgap is larger than that of the base. In this case, the $\exp[(E_{gE}-E_{gB})/kT]$ factor in (8) could increase current gain by a very large amount. A vertical Ge NPN bipolar transistor having an n-type GaAs as the emitter was demonstrated by Jadus and Feucht [8]. Polysilicon emitter is commonly employed in vertical SiGe-base NPN and PNP bipolar transistors.

So far, there is no report of heterojunction SOI lateral bipolar devices. For now, we simply make the conjecture that heterojunction SOI lateral bipolar devices with very large current gains can be realized, and proceed to ***assume heterojunction devices to be available*** in this study.

III. CBIPOLAR INVERTERS

Consider the CBipolar inverter in Fig. 2. When Vin is high (near *Vcc*), the NPN is turned on and the PNP is turned off, and Vout is pulled to near ground by the NPN transistor. When Vin is low (near ground), the NPN is turned off while the PNP is turned on, and Vout is pulled up to near *Vcc* by the PNP. In this section, we examine the operation of a CBipolar inverter and a CBipolar NAND gate. We also estimate the performance and power dissipation of CBipolar inverters. We employ analytic equations which allow us to get a deeper insight into the physics governing CBipolar circuit operation. As discussed in the previous section, to make the analysis tractable, and as a first order estimation, we ignore parasitic resistances of the devices and assume narrow-gap-base heterojunction device structures as needed. Also, when needed to simply the modeling, we ***may assume the NPN and PNP transistors to have identical characteristics.***

A. Quasi-Static Transfer Characteristics and Currents in an Inverter Chain

Consider a chain of identical CBipolar inverters with one inverter driving its neighbor down the chain, as illustrated in Fig. 10. To derive the quasi-static inverter transfer curves during switching, we need to establish the relationship between the input and output voltages of an inverter in terms of the device currents during switching.

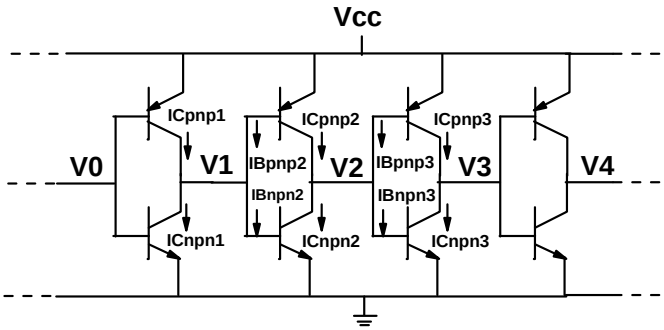

Fig. 10. CBipolar inverter chain. For clarity of illustration, only the case of FO = 1 is shown, and wire loads are not shown.

Let us consider inverter 2, with output V2, and inverter 3 with output V3. In general, we have, for FO larger than 1, FO base currents seen at node V2, so that

$$I_{Cpnp2} - I_{Cnpn2} = FO(I_{Bnpn3} - I_{Bpnp3}). \tag{9}$$

For a chain of identical inverters, inverter 1 and inverter 3 behave identically, so that

$$V_3 = V_1, \tag{10}$$

$$I_{Bnpn3} = I_{Bnpn1}, \tag{11}$$

and

$$I_{Bpnp3} = I_{Bpnp1}. \tag{12}$$

That is, (9) can be rewritten as

$$I_{Cpnp2} - I_{Cnpn2} = FO(I_{Bnpn1} - I_{Bpnp1}) \tag{13}$$

which relates the collector currents in inverter 2 to the base currents in inverter 1. Similarly we have

$$I_{Cpnp1} - I_{Cnpn1} = FO(I_{Bnpn2} - I_{Bpnp2}) \tag{14}$$

which relates the collector currents in inverter 1 to the base currents in inverter 2.

Referring to Fig. 10, and noting that V0 = V2 in a chain of identical inverters, we have the following voltage relations:

$$V_{BEpnp1} = V_2 - V_{cc}, \tag{15.1}$$

$$V_{BCpnp1} = V_2 - V_1, \tag{15.2}$$

$$V_{BEnpn1} = V_2 - V_{ee}, \tag{15.3}$$

$$V_{BCnpn1} = V_2 - V_1, \tag{15.4}$$

and

$$V_{BEpnp2} = V_1 - V_{cc}, \tag{15.5}$$

$$V_{BCpnp2} = V_1 - V_2, \tag{15.6}$$

$$V_{BEnpn2} = V_1 - V_{ee}, \tag{15.7}$$

$$V_{BCnpn2} = V_1 - V_2. \tag{15.8}$$

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

For given *Vcc* and *Vee*, these equations relate V_{BE} and V_{BC} of the transistors in inverters 1 and 2 to V1 and V2. Using these voltage relations in the current equations (1) and (2), and the corresponding current equations for PNP (not shown explicitly in this paper to avoid confusion), and applying the currents to the condition set by (13), we can derive the quasi-static transfer curve relating V1 (= Vin2) and V2 (= Vout2). Similarly, from consideration of the condition set by (14), we can derive the quasi-static transfer curve relating V1 (= Vout1) and V2 (= V0 = Vin1).

Figure 11 shows the modeled quasi-static transfer curve and the collector and base currents during switching, for an inverter with FO =1, assuming a current gain of 60 for both the NPN and the PNP. Such current gain is typical for a homojunction transistor [3, 4]. The transfer curve suggests that the noise margin is significantly smaller than *Vcc*/2. As can be inferred from (13) and (14), the transistors of the driving inverter must remain turned on sufficiently so that the collector currents are large enough to feed the base currents in the inverter being driven. For devices with a current gain of only 60, the transistors of the driving inverter have to maintain a large on current, and hence turn on and off somewhat gradually, as indicated by the transfer curve in Fig. 11.

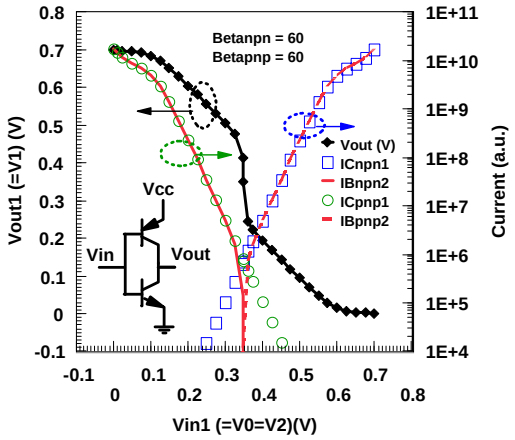

Fig. 11. Modeled transfer curve and transistor currents for a CBipolar inverter (inverter 1) with FO = 1 in an inverter chain. The transistors are assumed to have a current gain of 60.

Figure 12 is similar to Fig. 11, but for the case where both the NPN and the PNP have a current gain of 60,000. The 1,000x larger current gain represents a narrow-gap-base heterojunction device structure where the base bandgap is about 180 meV smaller than the emitter bandgap. For example, the device could have a Si emitter and a SiGe base where the SiGe bandgap is 180 meV smaller than that of Si. With a much larger current gain, the transistors of the driving inverter need to maintain a relatively small on current during switching, and hence turn on and off more abruptly. Comparison with Fig. 11 shows that larger current gain increases noise margin of the inverter.

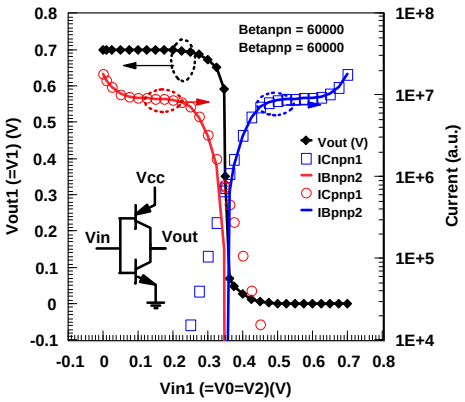

Fig. 12. Modeled transfer curve and transistor currents for a CBipolar inverter (inverter 1) with FO = 1 in an inverter chain. The transistors are assumed to have a current gain of 60,000.

B. Standby Current

Consider inverter 1 in Fig. 10 at standby with V1 = 0 and V0 = *Vcc*, i.e. with pnp1 off and npn1 on. When pnp1 is off, npn2 is also off and pnp2 is in full saturation. Therefore, (14) gives the standby collector current for npn1 as

$$I_{Cnpn1,sb} = FO \times I_{Bpnp2,sb}. \quad (16)$$

The base current for pnp2 in full saturation can be inferred from (2). Therefore, (16) can be re-written as

$$I_{Cnpn1,sb} = 2FO \times I_{B0\,pnp}(e^{qV_{cc}/kT} - 1). \quad (17)$$

Similarly, the standby collector current for pnp1 when npn1 is off is

$$I_{Cpnp1,sb} = 2FO \times I_{B0\,npn}(e^{qV_{cc}/kT} - 1). \quad (18)$$

The inverter standby current is not determined by the collector current of the on transistor, but by the emitter current of the on transistor. For the case with npn1 being on, we need to add the base current $I_{Bnpn1,sb}$ of npn1 to (17). Thus, the inverter standby current when npn1 is on is

$$I_{\text{standby}}(npn1on) = I_{Cnpn1,sb} + I_{Bnpn1,sb} \quad (19)$$
$$= 2FO \times I_{B0\,pnp}(e^{qV_{cc}/kT} - 1) + 2I_{B0\,npn}(e^{qV_{cc}/kT} - 1).$$

Similarly, for the case of pnp1 being on, we have

$$I_{\text{standby}}(pnp1on) = 2FO \times I_{B0\,npn}(e^{qV_{cc}/kT} - 1) + 2I_{B0\,pnp}(e^{qV_{cc}/kT} - 1). \quad (20)$$

Therefore, the average standby current for an inverter is

$$I_{\text{standby}} = (FO+1)I_{B0\,pnp}(e^{qV_{cc}/kT} - 1) + (FO+1)I_{B0\,npn}(e^{qV_{cc}/kT} - 1) \quad (21)$$
$$\approx (FO+1)(I_{B0\,pnp} + I_{B0\,npn})e^{qV_{cc}/kT}.$$

It should be noted that it is the base current at V_{BE} = *Vcc* that determine the standby current. Excess base current due to recombination, which is negligible at large V_{BE}, has little effect on the standby current. This justifies our ignoring the recombination component of the base current in our analyses, as discussed at the end of Section II-A. The standby power dissipation is simply the standby current multiplied by *Vcc*.

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

IV. CBipolar NAND Gates

In CMOS designs, the most commonly used circuit is the NAND gate. A CBipolar NAND gate is shown in Fig. 13. Just as in a CMOS NAND gate, the node Vx is floating. For a NAND gate to function properly, Vx has to reach *Vee*, or very close to it, when either Vin1 or Vin2 switches. In this section, we want to verify that a CBipolar NAND gate functions properly.

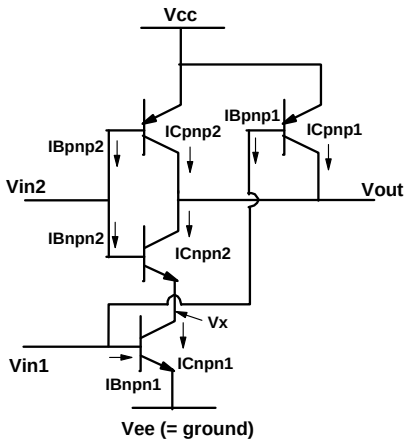

Fig. 13. Circuit schematic of a CBipolar NAND gate.

A. Case 1: Vin1 = Vcc and Vin2 Switching

The node voltage Vx is determined by the current continuity requirement

$$I_{Cnpn1} = I_{Bnpn2} + I_{Cnpn2}, \tag{22}$$

and the voltage relations:

$$V_{BEnpn1} = V_{cc} \tag{23.1}$$

$$V_{BCnpn1} = V_{cc} - V_x \tag{23.2}$$

$$V_{BEnpn2} = V_{in2} - V_x \tag{23.3}$$

$$V_{BCnpn2} = V_{in2} - V_{out}\,. \tag{23.4}$$

Substituting these voltages into the current equations (1) and (2), (22) gives Vx as a function of Vin2 and Vout. Figure 14 is a plot of Vx versus Vin2, with Vout as a parameter. It shows that Vx stays close to *Vee* (= ground) for most part of Vin2, rising to only about 20 mV above ground when Vin2 reaches *Vcc*, suggesting that the CBipolar NAND gate functions properly for the case of Vin2 switching.

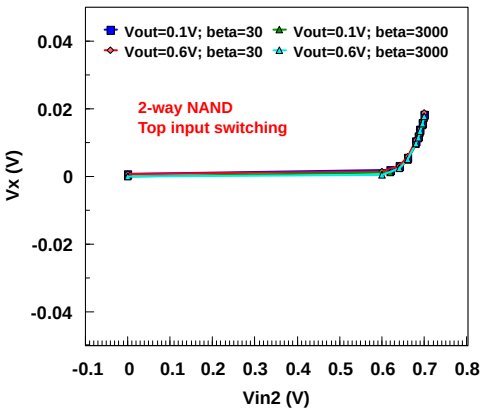

Fig. 14. Vx as a function of Vin2 for a CBipolar NAND gate for the case of Vin2 switching and *Vcc* = 0.7 V, with Vout as a parameter. In one case, the devices are assumed to have a current gain of 30. In another case, the devices are assumed to have a current gain of 3000.

B. Case 2: Vin2 = Vcc and Vin1 Switching

In this case, Vx is determined by the same current continuity requirement (22). The voltage relations are:

$$V_{BEnpn1} = V_{in1} \tag{24.1}$$

$$V_{BCnpn1} = V_{in1} - V_x \tag{24.2}$$

$$V_{BEnpn2} = V_{cc} - V_x \tag{24.3}$$

$$V_{BCnpn2} = V_{cc} - V_{out}\,. \tag{24.4}$$

Again, substituting these voltages into the current equations (1) and (2), (22) gives Vx as a function of Vin1 and Vout. Figure 15 is a plot of Vx versus Vin1, with Vout as a parameter. It shows that Vx reaches ground when Vin1 reaches *Vcc*, suggesting that a CBipolar NAND gate functions properly for the case of Vin1 switching.

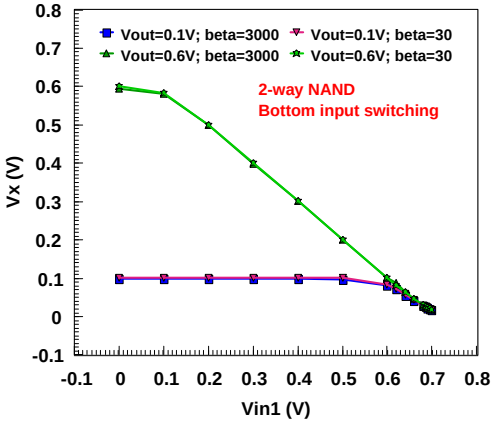

Fig. 15. Vx as a function of Vin1 for a CBipolar NAND gate for the case of Vin1 switching and *Vcc* = 0.7 V, with Vout as a parameter. In one case, the devices are assumed to have a current gain of 30. In another case, the devices are assumed to have a current gain of 3000.

V. Estimation of Inverter Chain Propagation Delay

It is instructive to examine the transfer curve and currents in Fig. 12. Vout1 is pulled down by npn1. Figure 12 shows that, as Vin1 increases, the collector current of npn1 remains substantially below its peak for most values of Vin1. When Vin1 is within about 50 mV of *Vcc*, the collector current of npn1 rises rapidly, approaching its maximum value as Vin1 reaches *Vcc*. The implication is that, in the switching of a CBipolar inverter, there is relatively little current to pull down the load at the output node until the input voltage gets very close to *Vcc*. The output is pulled down most rapidly after the input voltage reaches *Vcc* where the npn1 current is at its maximum.

This observation suggests that the propagation waveform for an inverter chain (Fig. 10) is as illustrated in Fig. 16. As V1 approaches *Vcc*, V2 starts to drop, being pulled down by npn2. Similarly, as V2 approaches ground, V3 starts to rise, being pulled up by pnp3.

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

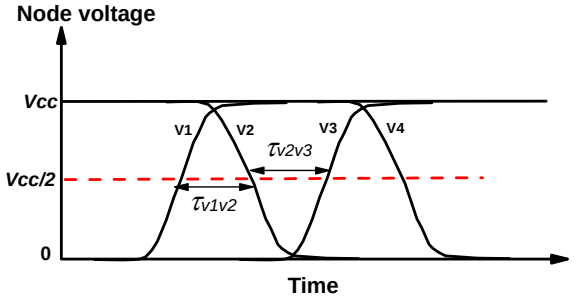

Fig. 16. Schematic illustration of the propagation waveform in an inverter chain.

A. Estimation of Inverter Propagation Delay Time τ_{delay}

Referring to Fig. 16, the transition time τ_{v1v2} can be written as the sum of two parts. One part is for V1 being pulled up from *Vcc*/2 to *Vcc*. The other part is for V2 being pull down from *Vcc* to *Vcc*/2. In general, these two parts overlap due to the fact that V2 starts dropping from *Vcc* towards *Vcc*/2 before V1 reaches *Vcc*, as illustrated in Fig. 16. One way to provide an ***upper-bound estimate for τ_{v1v2} is to assume these two parts do not overlap***, i.e. V2 starts its transition from *Vcc* towards *Vcc*/2 only after V1 has reached *Vcc*, as illustrated in Fig. 17. That is

$$\tau_{v1v2}(\text{upper bound}) = T[V_1 \text{ from } V_{cc}/2 \to V_{cc}] + T[V_2 \text{ from } V_{cc} \to V_{cc}/2]. \quad (25)$$

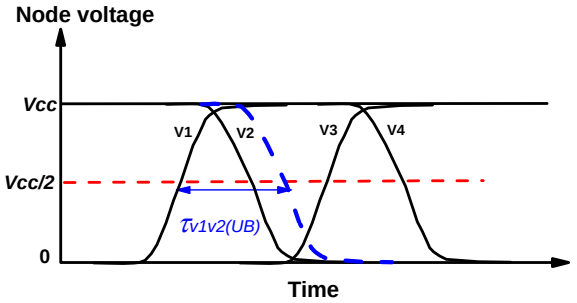

Fig. 17. Schematic illustrating an estimation of the upper bound of the transition time τ_{v1v2}.

Consider the second term in (25), with V2 being pulled down from *Vcc* to *Vcc*/2 by npn2. The load on node V2 is from npn3 and pnp3, multiplied by the number of fan-outs, plus any external load capacitance C_L. As can be inferred from Fig. 17, when V2 is being pulled down, V3 is at ground, with npn3 in saturation and pnp3 in forward-active mode. The total load on V2 is therefore

$$C_{v2down} = FO[C_{DEnpn3} + C_{DCnpn3} + C_{BEnpn3} + C_{BCnpn3} + C_{BEpnp3} + C_{BCpnp3}] + C_L , \quad (26)$$

where C_{DE} is the diffusion capacitance associated with the forward-biased emitter–base diode of npn3 and C_{DC} is the diffusion capacitance associated with the forward-biased collector–base diode of npn3. C_{BE} and C_{BC} are the sum of depletion layer capacitance and fringing capacitance of the base–emitter diode and base–collector diode, respectively. FO is the number of fan-outs attached to V2. The collector current pulling down V2 is I_{Cnpn2}. Therefore, the second term in (25) is given by

$$T[V_2 \text{ from } V_{cc} \to V_{cc}/2] = \int_{Vcc}^{Vcc/2} \frac{C_{v2down}}{-I_{Cnpn2}} dV_2 = \int_{Vcc/2}^{Vcc} \frac{C_{v2down}}{I_{Cnpn2}} dV_2 . \quad (27)$$

The (–) sign in the first integral is to account for the negative sign of the current for pull down.

Next, let us consider the 1st term in (25). For simplicity, we ***assume the NPN and NPN to have the same characteristics***. For such symmetric inverters, the time needed to pull V1 from *Vcc*/2 to *Vcc* is the same as the time needed to pull V2 from *Vcc*/2 to 0, i.e.

$$T[V_1 \text{ from } V_{cc}/2 \to V_{cc}] = T[V_2 \text{ from } V_{cc}/2 \to 0] = \int_0^{Vcc/2} \frac{C_{v2down}}{I_{Cnpn2}} dV_2 . \quad (28)$$

Substituting (27) and (28) into (25), we have an upper bound for the propagation time τ_{v1v2} as

$$\tau_{v1v2}(\text{upper bound}) = \int_0^{Vcc} \frac{C_{v2down}}{I_{Cnpn2}} dV_2 . \quad (29)$$

Similar consideration can be applied to derive an upper bound for the propagation time τ_{v2v3}. Since we assume NPN and PNP devices to be identical in characteristics, we have $\tau_{v1v2} = \tau_{v2v3}$. An upper-bound estimate of the inverter propagation delay time is therefore

$$\tau_{delay}(\text{upper bound}) = \frac{1}{2}\tau_{v1v2}(\text{upper bound}) + \frac{1}{2}\tau_{v2v3}(\text{upper bound})] = \int_0^{Vcc} \frac{C_{v2down}}{I_{Cnpn2}} dV_2 . \quad (30)$$

As can be seen from the subsection below, except for C_L, the components of C_{v2down} are all functions of the base width W_B or the depletion layer widths W_{dBE} and W_{dBC}, which in turn depend on V_{BE} and V_{BC}. That is C_{v2down} is an implicit function of V2. Also, as can be inferred from Fig. 17, during the transition of V2 from *Vcc* to ground, V1 is at *Vcc*, so that npn2 is in forward-active mode with a collector current of $I_{Cnpn2} = I_{C0npn}\exp(qVcc/kT)$. Therefore, (30) can be re-written as

$$\tau_{deoay}(\text{upper bound}) = \frac{V_{cc}}{I_{C0npn} e^{qV_{cc}/kT}} \int_0^{V_{cc}} \frac{C_{v2down}(V_2)}{V_{cc}} dV_2 = \frac{V_{cc}\overline{C}_{v2down}(V_{cc})}{I_{C0npn} e^{qV_{cc}/kT}} , \quad (31)$$

where $\overline{C}_{v2down}(V_{cc})$, which is the integral in (31) and a function of *Vcc*, is the capacitance load on the node V2 averaged over V2 transitioning from *Vcc* to ground.

B. Homojunction Si CBipolar

Let us consider the device capacitance loading on V2, i.e. the term inside the bracket in (26). From the emitter–collector symmetry of the transistors, we have $C_{DCnpn3} = C_{DEnpn3}$. The

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

stored charge responsible for C_{DEnpn3} is due to the forward collector current in npn3, which has a V_{BE} value of V2, i.e.

$$Q_{DEnpn3}(V_{BEnpn3}=V_2)=\tau_{Fnpn}I_{Cnpn3}(V_{BEnpn3}) = \tau_{Fnpn}I_{C0npn}e^{qV_2/kT}, \quad (32)$$

where τ_{Fnpn} is the forward transit time of the NPN transistor [9]. The corresponding diffusion capacitance is a function of V2, and is given by

$$C_{DEnpn3}(V_2)=\frac{\partial Q_{DEnpn3}}{\partial V_{BEnpn3}}=\frac{q\tau_{Fnpn}}{kT}I_{Cnpn0}e^{qV_2/kT}. \quad (33)$$

The B–E capacitance C_{BE} and the B–C capacitance C_{BC} are given by [5]

$$C_{BE}(V'_{BE})=C_{dBE,tot}(V'_{BE})+L_E C_{BE,\,fringe} \quad (34)$$

and

$$C_{BC}(V'_{BC})=C_{dBC,tot}(V'_{BC})+L_E C_{BC,\,fringe}, \quad (35)$$

where $C_{BE,fringe}$ and $C_{BC,fringe}$ are the fringing capacitances per unit emitter stripe length L_E associated with the B–E diode and B–C diode, respectively. For CMOS, the fringing capacitance is typically 0.08 fF/μm. The same value is assumed here. The B–E diode depletion layer capacitance $C_{dBE,tot}$ and the B–C diode depletion layer capacitance $C_{dBC,tot}$ are given by

$$C_{dBE,tot}(V'_{BE})=A_E\varepsilon_{si}/W_{dBE}(V'_{BE}), \quad (36)$$

and

$$C_{dBC,tot}(V'_{BC})=A_E\varepsilon_{si}/W_{dBC}(V'_{BC}), \quad (37)$$

where ε_{si} is the permittivity of Si, A_E is the emitter area, and the depletion layer widths W_{dBE} and W_{dBC} are given by

$$W_{dBE}(V'_{BE})=\sqrt{2\varepsilon_{si}(\psi_{bi}-V'_{BE})/qN_B}, \quad (38)$$

and

$$W_{dBC}(V'_{BC})=\sqrt{2\varepsilon_{si}(\psi_{bi}-V'_{BC})/qN_B}. \quad (39)$$

In (38) and (39), ψ_{bi} is the diode built-in potential **given** by

$$q\psi_{bi}=E_g/2+kT\ln(N_B/n_i), \quad (40)$$

where E_g is the bandgap energy and n_i is the intrinsic carrier density.

The forward transit time τ_F is given by [5]

$$\tau_F=\tau_E+\tau_B+\tau_{BE}+\tau_{BC}, \quad (41)$$

where

$$\tau_E=I_B(V'_{BE},V'_{BC})W_E^2/3I_C(V'_{BE},V'_{BC})D_{pE} \quad (42)$$

is the emitter delay time,

$$\tau_B=W_B^2(V'_{BE},V'_{BC})/3D_{nB} \quad (43)$$

is the base delay time,

$$\tau_{BE}=W_{dBE}(V'_{BE})/2v_{sat} \quad (44)$$

is the B–E space-charge-region delay time, and

$$\tau_{BC}=W_{dBC}(V'_{BC})/2v_{sat} \quad (45)$$

is the B–C space-charge-region delay time, and v_{sat} is the electron saturation velocity.

The capacitance components in (26) averaged over the transition of V2 from *Vcc* to 0 can be readily calculated using the voltage dependence in (34) to (45). In terms of these averaged parameters, the propagation delay time has the form

$$\tau_{delay}(\text{upper bound})=2\bar{\tau}_{Fnpn}FO + \frac{V_{cc}}{I_{C0npn}e^{qV_{cc}/kT}}\left[(\bar{C}_{BEnpn3}+\bar{C}_{BCnpn3}+\bar{C}_{BEpnp3}+\bar{C}_{BCpnp3})FO+C_L\right], \quad (46)$$

with each of the averaged quantities a function of *Vcc*, as noted in (31). Equation (46) is plotted as function of *Vcc* in Fig.18 for a Si-OI CBipolar inverter using typical measured NPN device parameters [3, 4]. Also plotted is the standby power dissipation per inverter, using (21). It shows that τ_{delay} decreases approximately exponentially with *Vcc*, due to the pull up and pull down currents increasing exponentially with *Vcc*. The delay levels off to a minimum value of $2\bar{\tau}_{Fnpn}FO$.

Fig. 18 shows that, even for case of FO of 4 with C_L load (CMOS designers typically use FO = 4 as a metric to judge CMOS circuit speed), the inverter delay reaches below 10 ps. However, the standby power dissipation, larger than 10 μW at *Vcc* = 1 V, is much too large for Si-OI CBipolar to be of interest as a logic technology.

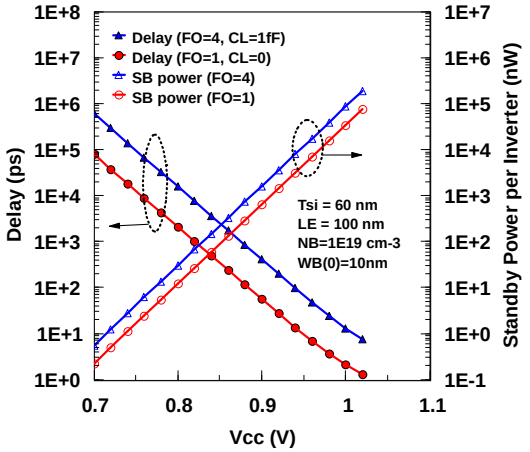

Fig. 18. Calculated propagation delay and standby power dissipation as a function of *Vcc*, for a case of FO = 1 and C_L = 0, and a case of FO = 4 with C_L = 1 fF. Si-OI devices are assumed, with device parameters as indicated.

C. *Sensitivity to Power Supply Variation*

The calculated delays in Fig. 18 stops at *Vcc* = 1.02 V. The delay for FO = 1 and C_L = 0 shows signs of leveling off starting at *Vcc* of about 0.98 V. If we had extended the calculation to larger values of *Vcc* (which would require inclusion of device resistances and CAD tools and hence is beyond the scope of this study), we would see the delay leveling off towards a value of $2\bar{\tau}_{Fnpn}$. The leveling off should occur rather rapidly for a couple of reasons. First, a combination of high-injection effect and emitter-resistance effect causes the device on current to saturate rapidly with increasing *Vcc*, as indicated by the measured device current in Fig. 3. Second, the forward transit time τ_F is not truly constant. It increase slowly as *Vcc* increases.

For designs where minimum sensitivity to power supply variation is required, *Vcc* should be large enough so that the circuit operates close to its minimum delay. For the example in Fig. 18, that would mean using *Vcc* of 1.04 V or slightly larger. A larger *Vcc* value leads to smaller delay, larger active power dissipation, and significantly larger standby power

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

dissipation since standby power dissipation increases exponentially with *Vcc*. To ensure standby power dissipation remains attractively low, narrow-gap-base heterojunction device structures are needs, as discussed in the next subsection.

D. Narrow-Gap-Base Heterojunction CBipolar

Figure 18 shows that for the standby power to be less than 1 nW, *Vcc* needs to be less than 0.75 V. Unfortunately, for Si-base devices, the delays at *Vcc* < 0.75 V are greater than 10 ns even for an unloaded inverter. Such slow circuits are probably of limited interest.

There are two approaches to engineering SOI CBipolar to achieve both attractive circuit speed and low standby power dissipation. The wide-gap-emitter approach is most commonly considered for suppressing base current [10]. In this approach, silicon remains the material for the base region, but a wide-gap semiconductor is used for the emitter. With a bandgap larger than Si as the emitter, the effect is to shift the standby power plots in Fig. 18 to the right, thus reducing the base current and standby power at a given *Vcc*. There is little or no change to the delay plot in Fig. 18. As a result, the wide-gap-emitter approach does not reduce *Vcc*, and hence has little effect on active power dissipation which is proportional to V_{cc}^2.

Another approach is to employ narrow-gap base but still keeping Si as emitter [11]. As discussed in Section II-D, narrow-gap base enables *Vcc* to be reduced without reducing the desired collector current when the transistor is turned on. By keeping Si as the emitter, the standby power dissipation as a function of *Vcc* is not changed. The net effect is shifting the delay vs. *Vcc* curves to the left in Fig. 18 without shifting the plots for standby power dissipation.

Both the wide-gap-emitter and the narrow-gap-base approaches probably yield about the same minimum circuit delays, the narrow-gap-base approach is preferred because it leads to both significantly lower standby power dissipation and lower active power dissipation.

The device parameters, such as apparent bandgap narrowing and minority carrier mobility as a function of doping concentration, *etc*., needed to calculate the collector current are not as well studied for Ge and SiGe as for Si. Instead of calculating the delays for Ge-base and SiGe-base CBipolar inverters, as we have done for Si-base case, we can obtain a projected estimate for the delays by using (7) to shift the calculated delay plots for Si-base inverters. This is done in Fig. 19. The plots suggest that with narrow-gap-base heterojunction structures, the power supply voltage for CBipolar could be reduced towards 0.5 V, thus achieving low active power dissipation as well as very low standby power dissipation. For Si-emitter heterojunction CBipolar, operation at *Vcc* of 0.7 V could mean standby power dissipation of about 0.1 nW. Operation at *Vcc* of 0.5 V could mean standby power dissipation of only 0.1 pW. Such projections of performance and power dissipation suggest narrow-gap-base heterojunction CBipolar could be an attractive candidate for "post-CMOS" technology [11].

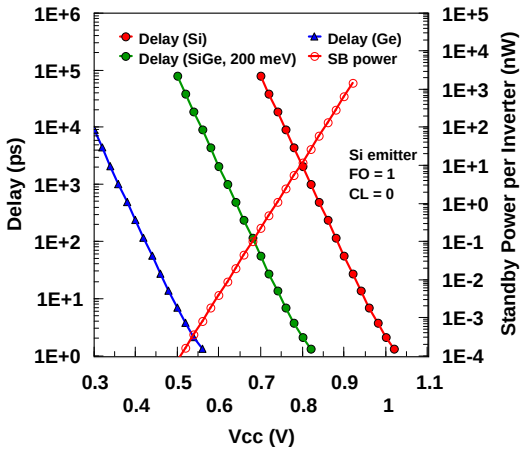

Fig. 19. Comparison of calculated propagation delay and standby power dissipation for Si homojunction CBipolar with projections for narrow-base heterojunction CBipolar for the case of FO = 1. All devices are assumed to have the same Si emitter region and base doping and base width as in Fig. 18. The Ge-base delay is projected by shifting the Si plot to the left by 0.46 V. The SiGe-base delay is projected by shifting the Si plot to the left by 200 mV, equivalent to assuming the SiGe bandgap to be 200 meV smaller than that of Si.

VI. Rethinking Device and Circuit Density (*vis-à-vis* CMOS)

It has been reported that a symmetric lateral bipolar transistor with T_{si} of 60 nm can achieve drive currents of almost 5 mA/μm, about 5x that for a typical CMOS device [3, 4]. The higher drive-current capability implies smaller device area for a bipolar transistor than for a CMOS device. This is illustrated in Fig. 20.

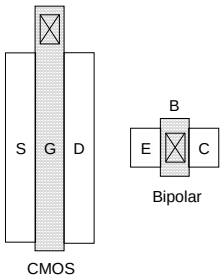

Fig. 20. Layout schematics comparing the size of a CMOS transistor and a SOI lateral bipolar transistor, both delivering the same peak drive current. The bipolar transistor is assumed to carry 5x as much current per device width as the CMOS transistor. Base contact above the intrinsic-base region is allowed for bipolar.

One way to appreciate the inherent drive-current advantage of lateral bipolar transistors compared to CMOS, or field-effect transistors (FET) made of 2-D materials such as graphene, MoS_2, *etc*., is to consider the areal charge carrier density. For CMOS, and FET in general, the maximum charge density in the inversion channel is typically about 1E13/cm^2. To our knowledge, the highest reported inversion layer charge density is 6E13/cm^2 for graphene in a Hall bar structure gated with polymer electrolyte [12]. For the bipolar transistor in Figs. 18 and 19, with T_{si} = 60 nm and N_B = 1E19/cm^3, the areal charge density is 6E13/cm^2. If we consider a different design point, say doubling either T_{si} or N_B, or both, the charge density for the bipolar transistor could exceed 1.2E14/cm^2.

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

Indeed, the semiconductor thickness T_{si} is a design/process parameter that can be used to substantially improve CBipolar device/circuit density. For example, by doubling T_{si}, the emitter length L_E (CMOS device width) can be reduced by half.

A. Limitation on T_{si}: Current Crowding Consideration

The value of T_{si} should not be too large to cause appreciable current crowding, where the collector current density in the intrinsic base is substantially larger near the top (close to extrinsic base) than near the bottom (close to the BOX). Current crowding is negligible if the condition

$$[R_{Sbi}(0)J_B(V_{cc})T_{si}^2/2] < kT/q \quad (47)$$

is met [5], where $R_{Sbi}(0)$ is the intrinsic-base sheet resistivity at V_{BE} = 0, and $J_B(V_{cc})$ is the operating base current density when V_{BE} = *Vcc*. Equation (47) gives the maximum T_{si} for negligible current crowding as

$$T_{si,\max} = \sqrt{\frac{2kT}{qR_{Sbi}(0)J_B(V_{cc})}}. \quad (48)$$

For the device in Figs. 18 and 19, we have $R_{Sbi}(0)$ = 9 kΩ/□. For *Vcc* = 1.0 V, J_B = 1.47E5 A/cm^2 and $T_{si,\max}$ = 62.5 nm. This verifies that the transistor, with T_{si} = 60 nm, has negligible current crowding when operated at *Vcc* of 1.0 V. For lower *Vcc*, e.g. *Vcc* of 0.7 V for operation of a SiGe-base heterojunction device, J_B is exponentially lower and there is no current crowding concern at all for practical T_{si} values. That is, for narrow-gap-base heterojunction CBipolar, the limit on large T_{si} is the ability to manufacture the devices, not due to current crowding considerations.

B. Limitation on T_{si}: RC Delay Associated with r_{bi}

Besides current crowding, another concern for employing large T_{si} is the *RC* time delay associated with the intrinsic-base resistance r_{bi}, which is proportional to (T_{si}/L_E), as indicated in (6). When T_{si} is increased by 2x in order to reduce L_E by 2x and increase device/circuit density by 2x, r_{bi} of the resulting device is increased by 4x. We need to ensure that the *RC* time delay associated with r_{bi} does not limit circuit delays.

The resistor r_{bi} and the diode capacitors $C_{dBE,tot}$ and $C_{dBC,tot}$ form a *RC* network. For the effect of r_{bi} to be negligible, the delay associated with this *RC* net work should be small compared to the CBipolar circuit delay. It can be inferred from (6), (36) and (37) that the delay time given by $r_{bi}(C_{dBE,tot} + C_{dBC,tot})$ is independent of L_E and varies as $(T_{si})^2$.

For a given transistor, the emitter–collector separation W_{E-C} is fixed, as can be inferred from Fig. 7. W_{E-C} is equal to the sum of the quasi-neutral base width W_B, the B–E diode depletion layer width W_{dBE}, and the B–C diode depletion layer width W_{dBC} [5], i.e.

$$W_{E-C} = W_B(V'_{BE}, V'_{BC}) + W_{dBE}(V'_{BE}) + W_{dBC}(V'_{BC}). \quad (49)$$

Equation (49) implies that, during device switching, as W_B increases, W_{dBE} and W_{dBC} decrease. That is, as r_{bi} decreases, the depletion layer capacitances C_{dBE} and C_{dBC} increase. In other words, during device switching, the product $r_{bi}(C_{dBE,tot} + C_{dBC,tot})$ is not expected to change much. For the device in Figs. 18 and 19, with T_{si} = 60 nm, N_B = 1E19/cm^3 and $W_B(0)$ = 10 nm, the *RC* product is 0.19 ps at V_{BE} = V_{BC} = 0, 0.3 ps at V_{BE} = 0 and V_{BC} = −1 V, and 0.24 ps at V_{BE} = V_{BC} = 1.0 V. These *RC* delay times are small compared with the minimum circuit delays in Figs. 18 and 19. These results justify the assumption of neglecting r_{bi} in modeling CBipolar inverter delays (see discussion in Section II-C).

VII. Comparison with CMOS

At the power-performance level, a comparison between CMOS and CBipolar can be made using the projected CMOS numbers in the 2013 *ITRS* report [13]. This is shown in Table I. Here we choose the MG CMOS, i.e. FinFET, numbers for comparison because MG CMOS represents the best CMOS device structure. The FinFET numbers are from the projection for year 2017, instead of later years, because it is the furthest out year where the projection column contains no red-color entry (which means no known solution). *ITRS* reports show CMOS device performance in terms of "*CV*/*I*" and not in terms of inverter speed. The quantity *CV*/*I* represents the intrinsic delay of a device. It correlates with the intrinsic delay of a CMOS inverter. Quantitatively, the intrinsic delay of a CMOS inverter should be larger than the device *CV*/*I* because the load on the output of a CMOS inverter comes from an nFET and a pFET. The CMOS inverter standby power is simply the product of device I_{off} and *Vdd*. The entry "V^V", i.e. $(Vdd)^2$ and $(Vcc)^2$, is meant to compare the active power dissipation.

It is clear from Table I that both CMOS and CBipolar can provide circuits of comparably high speed. The main difference is in standby power dissipation, which has been a critical issue for CMOS for some time. Si-OI CBipolar is simply not competitive because of its very high standby power dissipation. Both the power and the performance of SiGe-OI CBipolar, with 200 meV smaller bandgap, are competitive with high-performance CMOS. The implication is that increasing the Ge percentage, thus further decreasing the base bandgap and *Vcc*, SiGe-OI CBipolar could have both power and performance better than CMOS. Ge-OI CBipolar is clearly far superior to CMOS, with standby power dissipation even lower than that of low-power CMOS.

Table I. Comparison of projected CMOS (FinFET) with calculated CBipolar inverter. Device widths are 100 nm for CBipolar and 500 nm for CMOS, consistent with the discussion in Fig. 20.

	CMOS (FinFET) (2013 ITRS projection for 2017)		CBipolar (Si emitter/collector) (From Fig. 19 for FO = 1, CL = 0)		
	Low Power	High Perf	Si-base	SiGe-base (200 meV)	Ge-base
Vdd; *Vcc* (V)	0.8	0.8	1.02	0.82	0.56
Standby Power	4 pW	40 nW	78 μW	28 nW	0.83 pW
V^2	0.64	0.64	1.04	0.67	0.31
CV/I (ps)	1.6	0.50			
Inverter Delay (ps)			1.4	1.4	1.4

VIII. Summary

The symmetric lateral bipolar transistor on SOI technology enables CBipolar circuits to be considered for practical use for the first time. The inverter operation, propagation delay, and standby power dissipation were modeled using analytic equations. When scaled to base widths of about 10 nm, the

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2014.2361696, IEEE Journal of the Electron Devices Society

propagation delays can reach sub-10 ps even for loaded inverters. However, for CBipolar standby power dissipation to be low enough to be of interest, narrow-gap-base heterojunction structures are needed. Heterojunction structures, with Si emitter and Ge base or SiGe base were considered, and the results suggest that such CBipolar technologies could have maximum speed comparable to the best CMOS and standby power dissipation lower than CMOS. In addition, due to the much larger drive-current capability of symmetric lateral bipolar transistors, CBipolar could have significant device/circuit density advantage over CMOS. It is our hope that the results from this study will stimulate the R&D investments needed to develop CBipolar into a mainstream high-performance and low-power logic technology.

REFERENCES

[1] J. C. Sturm, J. P. McVittie, J. F. Gibbons, and L. Pfeiffer, "A lateral silicon-on-insulator bipolar transistors with a self-aligned base contact," *IEEE Electron Device Lett*., vol. EDL-8, pp. 104-106, 1987.

[2] J. Cai, T. H. Ning, C. D'Emic, K. K. Chan, W. E. Haensch, J.-B. Yau, and D.-G. Park, "Complementary thin-base symmetric lateral bipolar transistor on SOI," *IEDM Tech. Dig.*, pp. 386-389, 2011.

[3] J. Cai T. H. Ning, C. D'Emic, J.-B. Yau, K. K. Chan, J. Yoon, K. A. Jenkins, R. Muralidhar and D.-G. Park, "SOI lateral bipolar transistor with drive current >3mA/μm," *IEEE S3S Conference*, 2013.

[4] J. Cai, T.H. Ning, C.P. D'Emic, J.-B. Yau, K.K. Chan, J. Yoon, R. Muralidhar, and D.-G. Park, "On the device design and drive-current capability of SOI lateral bipolar transistors," *IEEE J. Elect. Dev. Soc.*, vol. 2, No.5, pp. 105-115, 2014.

[5] T. H. Ning and J. Cai, "On the performance and scaling of symmetric lateral bipolar transistors on SOI," *IEEE J. Elect. Dev. Soc.*, vol. 1, No.1, pp. 21-27, 2013.

[6] H.H. Berger and S.K. Wiedmann, "Complementary transistor circuit for carrying out Boolean functions," U.S. Patent 3,956,641, May 1976.

[7] D. B. M. Klaassen, J. W. Slotboom, and H. C. de Graaff, "Unified apparent bandgap narrowing in n- and p-type silicon," *Solid-State Electron*. vol. 35, pp. 125-129, 1992.

[8] D. K. Jadus and D. L. Feucht, "The realization of a GaAs-Ge wide band gap emitter transistor," *IEEE Trans. Electron Devices*. vol. ED-16, pp. 102-107, 1969.

[9] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*, 2nd edition, Cambridge University Press, 2009.

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

[11] T. H. Ning, "A perspective of future nanoelectronic devices," *Proc. VLSI-TSA*, pp. 153-154, 2013.

[12] A. Pachoud, M. Jaiswal, P. K. Ang, K. P. Loh, and B. Özyilmaz, "Graphene transport at high carrier densities using a polymer electrolyte gate," *http://arxiv.org/pdf/1009.3367.pdf*.

[13] *International Technology Roadmap*, Process Integration, Devices, and Structures (PIDS) Chapter, Semiconductor Industry Association, 2013.

2168-6734 (c) 2013 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.