Optimization and Scaling of CMOS-Bipolar Drivers for VLSI Interconnects

HECTOR J. DE LOS SANTOS, MEMBER, IEEE, AND BERND HOEFFLINGER

Abstract—**In this paper, rules are presented for the optimized design of CMOS-bipolar drivers for large capacitive loads typical of VLSI interconnects. Simulations and closed-form solutions show that the n-p-n bipolar transistors have to be operated in the high-level injection mode, and that their sizes have to be tailored to the two-thirds power of the load, and it scales with the two-thirds power of the base width of the n-p-n transistor and with the one-third power of the channel length of the MOS transistor. For comparison, the CMOS cascade with a tailored second stage is shown to have competitive potential at the expense of an area being approximately 2.5 times larger than that of a CMOS-bipolar stage.**

List of Symbols

β_0	DC current gain.
β	Effective dc current gain in high injection.
C_0'	Gate oxide capacitance per unit area.
C_{BC}'	Base–collector capacitance per unit emitter length.
C_{CS}	Collector–substrate capacitance.
C_{CS}'	Collector–substrate capacitance per unit emitter length.
C_{BC}	Base–collector capacitance.
D_{nB}	Diffusion constant of electrons in base.
D_{pE}	Diffusion constant of holes in emitter.
$E_C = V_L/\mu$	Hot-carrier critical field.
I_B	Bipolar transistor base current.
I_C	Bipolar transistor collector current.
$I_{D\infty}$	Maximum drain current under velocity saturation effects.
I_K	Upper knee current.
I_K'	Upper knee current per unit emitter length.
k_n, k_p	Short-channel ideality factors.
L	Channel length.
L_E	Emitter stripe length.
μ_n, μ_p	Low field electron and hole mobilities, respectively.
N_B	Base doping density.
N_{DE}	Emitter doping density.
q	Electron charge.
τ_F	Base transit time.
V_{dd}	Supply voltage.
V_d	Diode voltage drop.
$V_G = V_{GS} - V_T$	Effective gate voltage.
V_{GS}	Gate to source voltage.
V_{TN}, V_{TP}	NMOS and PMOS threshold voltages, respectively.
v_L	Saturation velocity.
W	Channel width.
W_B	Base width.
W_E	Emitter stripe width.

Manuscript received April 1, 1986; revised June 16, 1986. This work was supported in part by the Office of Naval Research under Contract N00014-84-K-0438 and by a grant from VTC Incorporated, Minneapolis, MN.

The authors are with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907.

IEEE Log Number 8610407.

I. Introduction

THE SIGNAL propagation delay due to large interconnect capacitances is a major factor determining the performance of VLSI circuits [1]. In order to drive these capacitances, several technologies have been developed that merge CMOS and bipolar devices [2]–[6]. Although the merged devices have been proposed and their characteristics reported [7], the approach did not progress until recently. In view of the limited driving capabilities of the MOS transistors in general, and in particular within the VLSI environment, the bi-CMOS combination, which theoretically provides for an increase of the transconductance by β_0, has become the focus of attention.

In this paper, the most fundamental circuit implemented with this structure, namely, the CMOS-bipolar buffer is analyzed. A closed form expression of its pull-up time is derived together with a simple rule for scaling the size of the bipolar transistor as a function of load capacitance. In addition, a comparison based on SPICE simulations [8] is presented between optimal CMOS and CMOS-bipolar buffers.

It will be pointed out that the CMOS-bipolar buffer improves the delay time of the common CMOS buffer by much less than a factor of β_0, because a smaller effective β results as the bipolar transistor preferably operates in the high injection regime, and because it presents a significant capacitive load to the MOS drivers. However, moderate savings in area and favorable scaling rules are also demonstrated.

II. CMOS-Bipolar Buffer Optimization

A. DC Analysis

A circuit diagram of the basic CMOS-bipolar buffer is shown in Fig. 1. Figs. 2 and 3 show the models used in

0018-9383/86/1100-1722$01.00 © 1986 IEEE

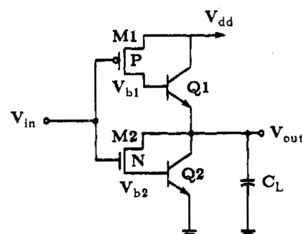


Fig. 1. CMOS-bipolar buffer.

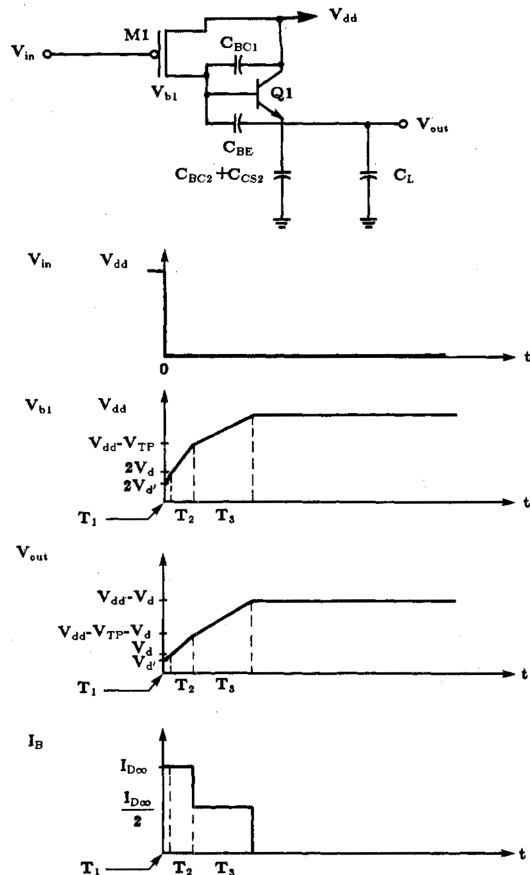
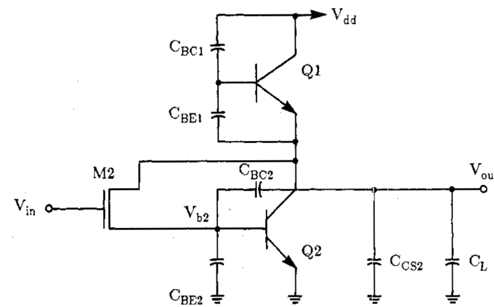


Fig. 2. Circuit model and waveforms for pull-up time analysis.

analyzing the pull-up and pull-down transients. For the dc analysis it is assumed that V_{in} has been at the specified level for a period of time much longer than that at which the circuit is clocked. The dc analysis results in the node voltages shown in Table I.

When the input voltage is zero, $M2$ is OFF and $M1$ is ON. Because at the end of the transient that leads to this state, $Q1$ has $I_{B1} = 0$, $V_{BC1} = 0$, and $V_{BE1} = 0.7$ V, the minority-carrier concentrations at the base–emitter and base–collector junctions are $n_{BO} \exp(0.7/V_t)$ and zero, respectively [9]. As a result, there is an emitter–collector current that leads to find $Q1$ with V_{BE1} below 0.7 V. $f_0 V_d$ is the amount by which V_{BE1} is below 0.7 V. $Q2$ is OFF with V_{BE2} assumed to be equal to V_{BE1}. f_0 has values betwen 0 and 1 and ultimately depends upon the relative magnitude of $C_{BC1} + C_{BE1} + C_{CS2}$ with respect to C_L.

When the input is V_{dd}, $M1$ is OFF and $M2$ is ON. $Q2$ assumes a state identical to that of $Q1$ above. $Q1$ is OFF with V_{BE1} assumed equal to V_{BE2}. Ultimately f_0 depends



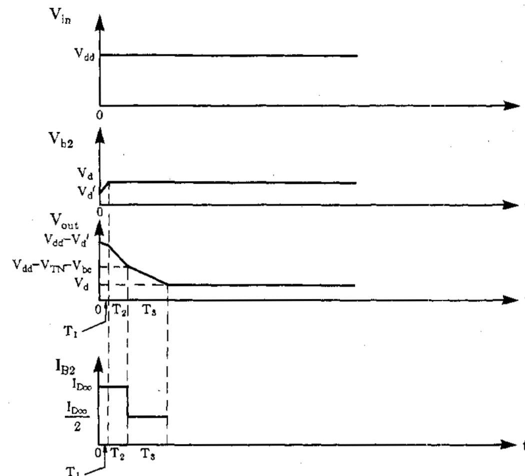


Fig. 3. Circuit model and waveforms for pull-down time analysis.

TABLE I
NODE VOLTAGES

V_{in}	V_{b1}	V_{b2}	V_{out}
0	V_{dd}	$(1 - f_0)V_d$	$V_{dd} - (1 - f_0)V_d$
V_{dd}	$2(1 - f_0)V_d$	$(1 - f_0)V_d$	$(1 - f_0)V_d$

upon the relative magnitude of $C_{BE1} + C_{BC2} + C_{CS2}$ with respect to C_L.

B. Pull-Up Time Analysis

During T_1, the PMOS current charges C_{BC1} and C_{BE1} by a voltage $f_0 V_d$. If C_L is comparable to $C_{BC2} + C_{CS2}$, then the output voltage also increases by $f_0 V_d$. But if C_L is much greater than $C_{BC2} + C_{CS2}$, then the change in output voltage is negligible. To characterize the output voltage change for intermediate values of C_L, two factors are introduced, namely, f_1 and f_2. $(1 - f_1)V_d$ denotes the voltage change experienced by $C_{BC2} + C_{CS2}$. f_2 denotes the extent to which C_L takes part in the charging period T_1 and like f_1 has values between 0 and 1. For values of C_L comparable to $C_{BC2} + C_{CS2}$, $f_2 \simeq 1$, whereas for values of C_L much greater than $C_{BC2} + C_{CS2}$, $f_2 \simeq 0$. At the end of T_1, $Q1$ enters into the active region and a current $\beta I_{D\infty}$ charges C_{BC2}, C_{BC1}, C_{CS2}, and C_L until the PMOS transistor comes out of saturation. At the end of T_2, the PMOS transistor enters into the ohmic region. Here, an average current $\beta I_{D\infty}/2$ continues the charging process further until the output is one diode drop below V_{dd}.

The charging current equals βI_{DS}, where $I_{DS} = I_{D\infty}$ and β is the current gain. $I_{D\infty}$ models the MOS transistor cur-

rent in saturation according to (1) [10], and β models the current gain including high-level injection effects as given in (2).

$$I_{D\infty} = kv_L C_0' W V_G$$
$$k \simeq \frac{1}{4}\sqrt{\frac{V_G}{E_c L}} \tag{1}$$

$$\beta = \sqrt{\frac{\beta_0 I_K}{I_B}}$$
$$I_K = \frac{qD_{nB} N_B L_E W_E}{W_B}. \tag{2}$$

The time intervals T_1, T_2, and T_3 are given by (3), (4), and (5), respectively.

$$T_1 = \frac{(C_{BC1} + C_{BE1} + C_{BC2} + C_{CS2} + f_2 C_L) f_0 V_d}{I_{D\infty}} \tag{3}$$

$$T_2 = \frac{(C_{BC1} + C_{BC2} + C_{CS2} + C_L)(V_{dd} - V_{TP} - V_d(1 + f_1))}{\beta I_{D\infty}} \tag{4}$$

$$T_3 = \frac{2(C_{BC1} + C_{BC2} + C_{CS2} + C_L) V_{TP}}{\beta I_{D\infty}}. \tag{5}$$

The total pull-up time is

$$T_d = T_1 + T_2 + T_3 = \frac{(C_{BC1} + C_{BE1} + C_{BC2} + C_{CS2} + f_2 C_L) f_0 V_d}{I_{D\infty}} + \frac{(C_{BC1} + C_{BC2} + C_{CS2} + C_L)(V_{dd} + V_{TP} - V_d(1 + f_1))}{\beta I_{D\infty}}. \tag{6}$$

By a similar analysis, the time intervals T_1, T_2, and T_3 of Fig. 3 and the total pull-down time are given by (A1)–(A4) in Appendix II. The circuit setup simulated using SPICE2.G is shown in Fig. 4.

The total pull-up time T_d can be minimized for any given load by tailoring the emitter stripe length L_E. Since $C_{BC} = C_{BC}' L_E$ and $I_K = I_K' L_E$ we can express (6) as

$$T_d = \frac{f_0 V_d (C_{BC1}' + C_{BE1}' + C_{BC2}' + C_{CS2}) L_E}{I_{D\infty}} + \frac{f_2 f_0 C_L V_d}{I_{D\infty}} + \frac{(V_{dd} + V_{TP} - V_d(1 + f_1))}{\sqrt{\beta_0 I_K' I_{D\infty}}} \left(\frac{C_L}{\sqrt{L_E}} + (C_{BC1}' + C_{BC2}' + C_{CS2}')\sqrt{L_E}\right). \tag{7}$$

Then, neglecting the fourth term and differentiating with respect to L_E, a minimum is obtained for any load C_L if

$$L_{E\text{opt}} = \left(\frac{VC_L}{2 f_0 V_d (C_{BE1}' + C_{BC2}' + C_{CS2}' + f_2 C_L')}\right)^{2/3} \cdot \left(\frac{I_{D\infty}}{\beta_0 I_K'}\right)^{1/3} \tag{8}$$

where $V = V_{dd} + V_{TP} - V_d(1 + f_1)$.

The minimum pull-up time can be approximated by

$$T_{d,\min} \simeq 2V\left(\frac{I_{D\infty} f_0 V_d (C_{BE1}' + C_{BC2}' + C_{CS2}' + f_2 C_L')}{V \beta_0 I_k'}\right)^{1/3} \cdot \frac{C_L^{2/3}}{I_{D\infty}} \tag{9}$$

and is proportional to the two-thirds power of C_L. The proportionality factor is a constant typical of a given CMOS-bipolar technology. The minimum pull-up time is dictated by $I_{D\infty}$, which scales with $L^{1/3}$ according to (1). It is also limited by the product $\beta_0 I_K'$. This product is a constant for a given technology, as it is given by

$$\beta_0 I_K' = \frac{qD_{nB}^2 N_{DE} L_{pE} W_E}{D_{pE} W_B^2}. \tag{10}$$

Hence, by decreasing the base width W_B, this lower bound on pull-up time can be improved with the two-thirds power of W_B.

The assumption that the bipolar transistors operate under high-level injection conditions was verified by observing the collector and base currents during both pull-up and pull-down transients (Fig. 5). This situation has been investigated further by varying the base impurity N_B, and hence the low-level current gain β_0 according to

$$\beta_0 = \frac{D_{nB} N_{DE} L_{pE}}{D_{pE} N_B W_B}. \tag{11}$$

Results are shown in Fig. 6 for $C_L = 10$ pF. It can be seen that the pull-up time decreases with increasing β_0. The dashed curve represents a closed-form solution, where

$$\beta_0 I_K = \frac{qD_n^2 N_{DE} L_{pE} W_E}{D_{pE} W_B^2}. \tag{12}$$

C. Limitations of Fundamental CMOS-Bipolar Buffer

Ideally, the bias conditions of $Q1/Q2$ following a pull-up/pull-down transient are those discussed in Section II-A with $f_0 = 0$. Whenever the buffer is driven, a portion of the initial output charging current is fed to the base of $Q1$ via C_{BC1} if the output is to go from HI to LO, or to the

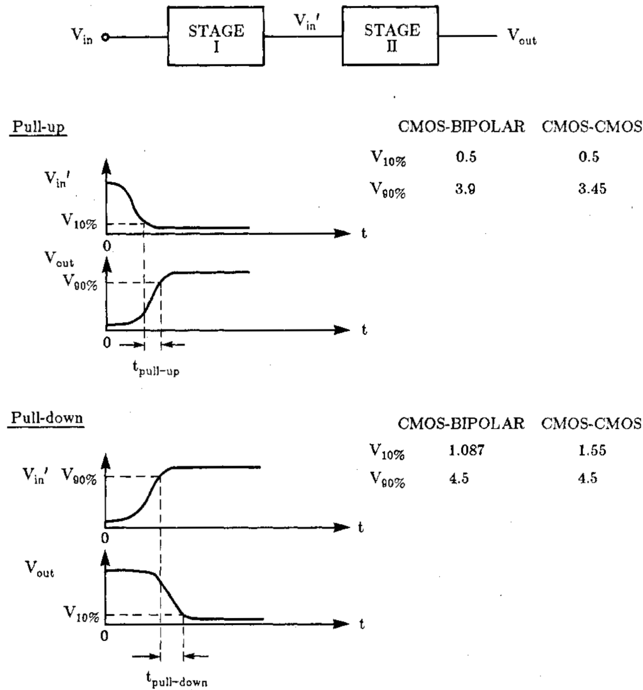


Fig. 4. Setup simulated using SPICE2.G.

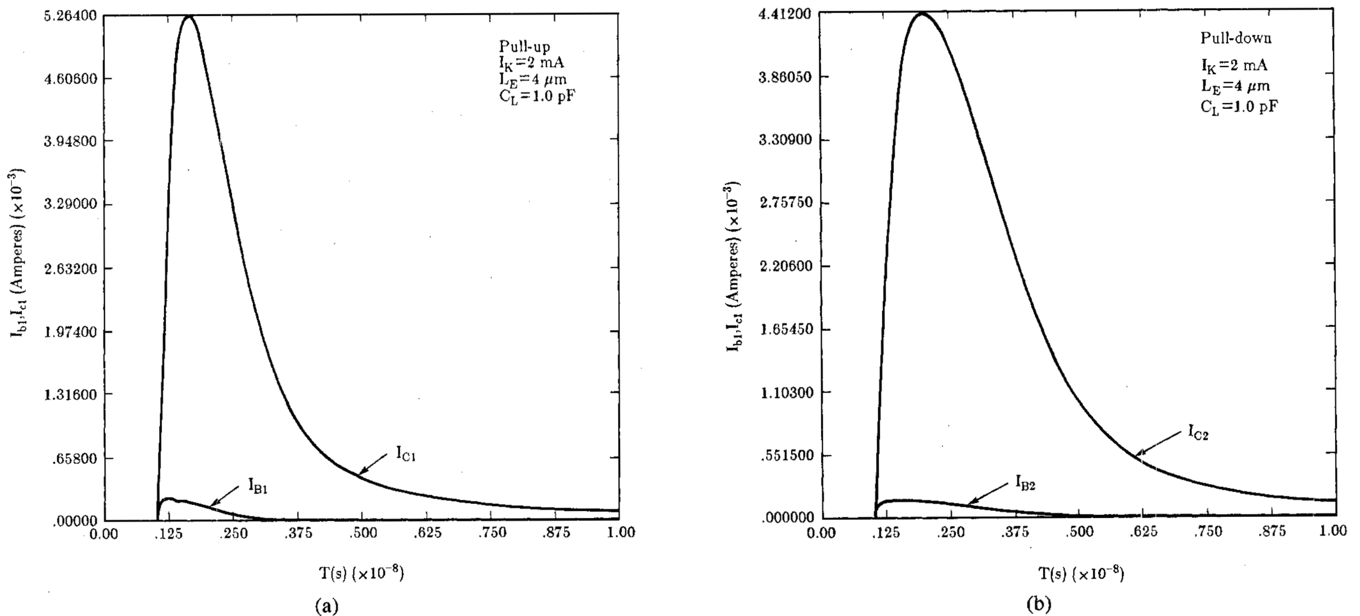


Fig. 5. n-p-n transistors base and collector currents (a) Pull-up. (b) Pull-down.

base of $Q2$ via C_{BC2} if the output is to go from LO to HI. Because of the charge already present in the bases, this portion of the output charging current applied to the bases causes $Q1$ or $Q2$ to quickly supply considerably large emitter or collector currents that reduce the total output current and thus limit the maximum achievable speed. Another limitation of the circuit is that if equal pull-up and pull-down times are persued, the size of $M2$ needed is larger than usual due to it having its effective gate voltage reduced by the base–emitter drop of $Q2$.

D. Improved CMOS-Bipolar Buffer

To avoid the undesirable currents subtracted by $Q1$ and $Q2$ from the total output current some means must be introduced to quickly remove their base charge as well as the currents fed through C_{BC1} and C_{BC2} in response to a change in V_{in}. This is accomplished by $M3$ and $M4$ in the circuit shown in Fig. 7. Upon a 0 to V_{dd} change in V_{in}, $M3$ removes the base drive from $Q1$. Upon a V_{dd} to 0 change in V_{in}, $M4$ removes the base drive from $Q2$.

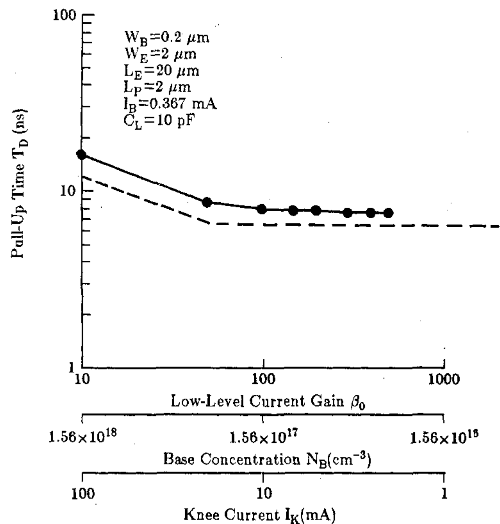


Fig. 6. Pull-up time versus n-p-n transistor current gain (β_0), base concentration (N_B), and knee current (I_K).

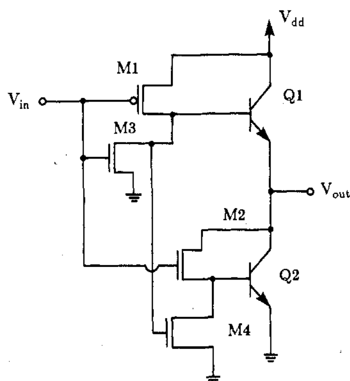


Fig. 7. Improved CMOS-bipolar buffer.

The dimensions of $M3$ and $M4$ are a function of the details of the input waveform (rise/fall times, duty cycle, etc.) and the load capacitance (which also determines the dimensions of $Q1$ and $Q2$). By setting an upper limit to the undesirable currents supplied by $Q1$ and $Q2$ as compared to the initial output current at the beginning of a transient, the widths of $M3$ and $M4$ can be adjusted until the desired limit is met.

At this point a procedure for the optimal design of the CMOS-bipolar buffer can be formulated as follows:

1) Obtain the appropriate parameters of the technology as they appear in Appendix I.
2) Obtain a T_d versus C_L curve.
3) Obtain the factors f_0, f_1, and f_2 by fitting (6) to the characteristic obtained in (2).
4) Calculate the optimum emitter length with (8).
5) Optimize the widths of $M3$ and $M4$.

This procedure has been followed starting with a non-optimal basic CMOS-bipolar buffer $L_E = 4$ μm. After obtaining a match between (6) and the simulated characteristic to within 10 percent, the optimal emitter lengths were obtained. Two cases have been evaluated, namely, the basic CMOS-bipolar buffer shown in Fig. 1 and the complete optimal buffer shown in Fig. 7. A comparison of their T_d versus C_L characteristics including estimated layout area is given in Fig. 8. The pull-up and pull-down times were made equal to within 1 ns by adjusting the size of $M2$. While the minimum emitter size is optimal for loads $C_L < 5$ pF, delay times for larger loads can be lowered by using larger emitters. The lower bound on the simulated T_d follows closely the two-thirds power dependence on C_L as suggested by the closed-form solution (see (9)).

III. Optimized CMOS Cascade

The CMOS cascade, Fig. 9(a), used as vehicle of comparison with the CMOS-bipolar buffer, was optimized for minimum pull-up time (for details of the derivation, please refer to Appendix III). The pull-up time model equation is given in (13), where $V = 3.45$ V, for proper comparison with the CMOS-bipolar buffer.

$$T_{d,c\min} = \sqrt{2}V\sqrt{\frac{C_0'L\left(1 + \frac{\mu_n}{\mu_p}\right)}{I_{D\infty}'I_{D\infty}}}\sqrt{C_L}. \quad (13)$$

The scaling law for the gate width of the pull-up transistor of the second stage W_p^1 for a minimum pull-up time as a function of the load capacitance is given by

$$W_p^1 = \sqrt{\frac{C_L k_n W_n^0}{C_0' k_p L\left(1 + \frac{\mu_p}{\mu_n}\right)}}. \quad (14)$$

The simulation results and the closed-form expression for the pull-up time versus load capacitance for the CMOS cascade together with estimates of its layout area are shown in Fig. 10. For comparison, the corresponding simulation results for the improved CMOS-bipolar buffer are included. A planar oxide-isolated twin-tub CMOS-bipolar technology has been assumed. The device areas have been assigned, including isolation, as shown in Table II. The areas quoted in Figs. 8 and 10 are related to a trench-isolated technology with $L = W_E = 2$ μm. The CMOS cascade requires an area approximately 2.5 times larger than the CMOS-bipolar stage. The speed of the CMOS-bipolar stage is not higher than 2.5 times that of the CMOS cascade.

IV. Conclusion

The design of a CMOS-bipolar buffer for minimum delay time will be a function of the intended capacitive load or range of loads. For a given technology, the emitter area of the bipolar transistor is the design parameter, assuming a minimum size PMOS transistor is used. Once the expression for pull-up time, (7), is fitted to the T_d versus C_L characteristic of an inverter with arbitrary emitter size, the optimum emitter size will be given by (8). The corresponding optimized pull-up time for the basic CMOS-bipolar buffer (Fig. 1) will be given by (9). Further speed

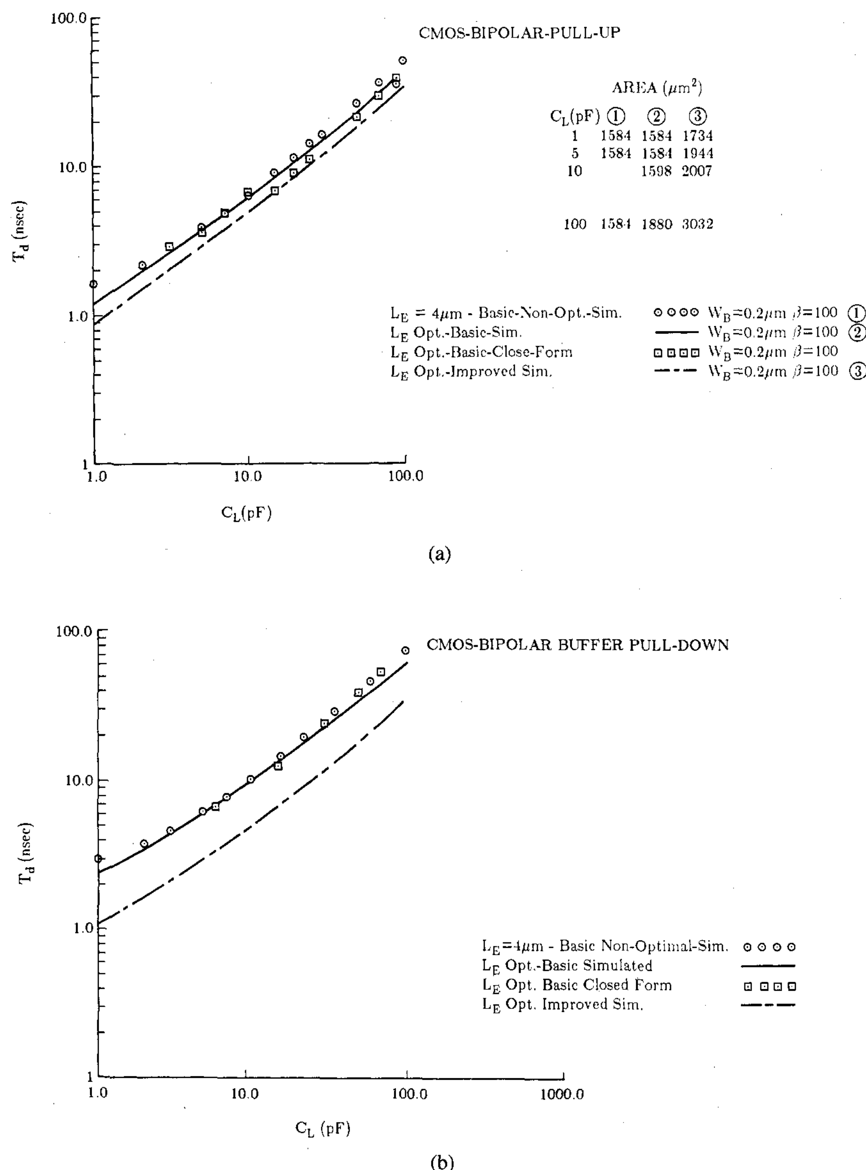


Fig. 8. T_d versus C_L: optimized CMOS-bipolar buffers.

improvement can be obtained by incorporating base charge removing transistors $M3$ and $M4$ to obtain the complete optimal buffer (Fig. 7).

Important parameters influencing the performance of the CMOS-bipolar buffer are the base doping density and the base width. For a given bipolar transistor size and load capacitance, Fig. 6 suggests that by decreasing N_B (i.e., increasing the dc current gain), the pull-up time decreases. Further improvements can be obtained by decreasing the base width, which will improve the lower bound on delay time with the two-thirds power of W_B. For MOS transistor channel lengths $L \leq 2$ μm, the optimized cascade becomes quite competitive, given the simpler technology, at the expense of an area approximately 2.5 times larger than that of a CMOS-bipolar stage. The area of the CMOS cascade will be halved with the arrival of three-dimensional CMOS technologies [11]. Similar progress is to be expected for the CMOS-bipolar circuits with the development of three-dimensionally merged CMOS and bipolar transistors.

Appendix I
Transistors Characteristics

A set of devices with parameters typical of an advanced CMOS-bipolar technology have been assembled and used in SPICE2.G simulations as a vehicle in studying the trade-offs present in the optimization of the CMOS-bipolar buffer. The resulting transistor characteristics for minimum dimension devices are shown in Tables III and IV, and in Figs. 11 and 12.

Appendix II
CMOS-Bipolar Pull-Down Transient Analysis

The pull-down time is modeled by three time intervals as noted in Fig. 3. During T_1 the NMOS current discharges C_{BC1}, C_{BE1}, and C_{BC2} by a voltage $f_0 V_d$; and

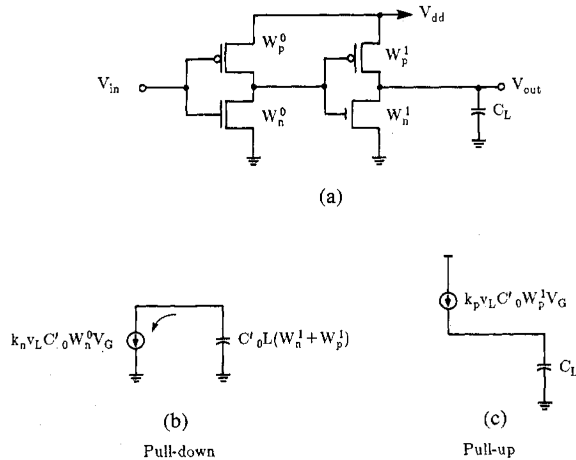


Fig. 9. Circuit models (and waveforms) for CMOS-cascade transient analysis.

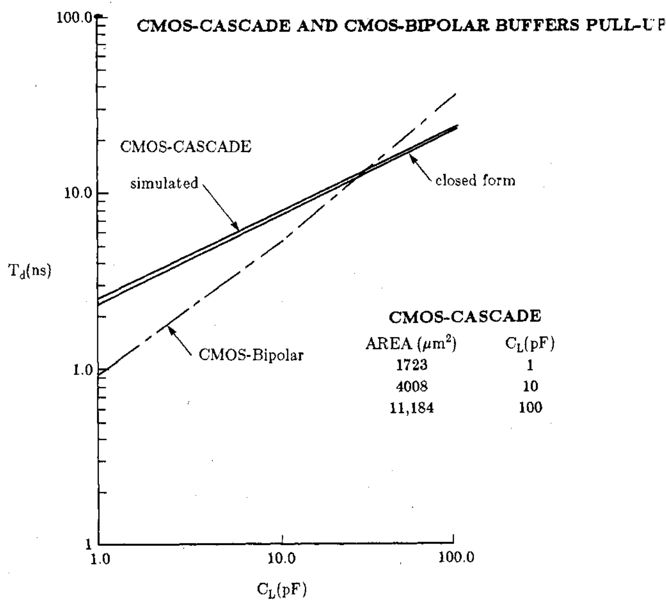


Fig. 10. T_d versus C_L: optimized CMOS-bipolar and CMOS-cascade buffers.

C_{CS2} and C_L, by a voltage $f_1 V_d$, where f_1 has values between 0 and 1. The current that discharges the output node, charges C_{BE2} simultaneously. At the end of T_1, $Q2$ enters into the active region and a current $(\beta + 1)I_{D\infty}$ discharges the output node until the NMOS transistor comes out of saturation. At the end of T_2, the NMOS transistor enters into the ohmic region. Here, an average current $((\beta + 1)I_{D\infty})/2$ continues the discharging process until the output is one diode drop above ground.

The three time intervals are given by (A1)–(A3).

$$T_1 = \frac{(C_{BC1} + C_{BE1} + C_{BC2} + C_{CS2} + f_2 C_L) f_0 V_d}{I_{D\infty}} \tag{A1}$$

$$T_2 = \frac{(C_{BC1} + C_{BC2} + C_{CS2} + C_L)(V_{Tn} - f_1 V_d)}{(\beta + 1) I_{D\infty}} \tag{A2}$$

$$T_3 = \frac{2(C_{BC1} + C_{BC2} + C_{CS2} + C_L)(V_{dd} - V_{Tn} - 2V_d)}{(\beta + 1) I_{D\infty}}. \tag{A3}$$

TABLE II
DEVICE AREAS

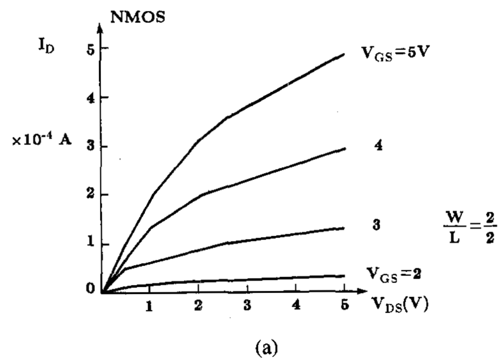
Isolation	npn Bipolar Transistor	CMOS inverter pair
Jct. Isolated	$22W_E(L_E+10W_E)$	$14L(W_p+W_n+7L)$
Trench Isolated	$18W_E(L_E+6W_E)$	$12L(W_p+W_n+5L)$

TABLE III
(An effective channel length of 2.0 μm is used. The widths of the NMOS and PMOS devices are 2 and 6 μm, respectively. The base width of the bipolar transistors is 0.2 μm.)

PARAMETERS	NMOS	PMOS
VTO (V)	1.2	-1.2
KP (AV^{-2})	5.75e-05	2.3e-05
GAMMA ($V^{0.5}$)	.354	.354
PHI (V)	.66	.66
PB (V)	.8	.8
CJ (Fm^{-2})	1.94e-04	1.94e-04
MJ	.5	.5
CJSW (Fm^{-1})	4.0e-11	3.5e-10
MJSW	.44	.44
JS (Am^{-2})	1.0e-09	1.0e-09
TOX (nm)	30	30
NSUB (cm^{-3})	5.0e15	5.0e15
XJ (um)	.3	.3
UO (cm^2/V-s)	500	200
UCRT (kV/cm)	20	50
VMAX (kms^{-1})	100	100

TABLE IV

PARAMETERS	NPN
IS	2e-17A
β_F	100
NF	1
IKF	2mA
NE	1.5
β_R	1
NR	1
NC	1.5
R_B	50 ohms
R_E	10 ohms
R_C	20 ohms
VJE	.7V
MJE	.44
TF	10 ps
VJC	.75V
MJC	.5
VJS	.7V
MJS	.5
EG	1.11eV
L_E	4μm
W_E	2μm
C_{BC}	0.04pF
C_{BE}	0.04pF
C_{CS}	0.05pF



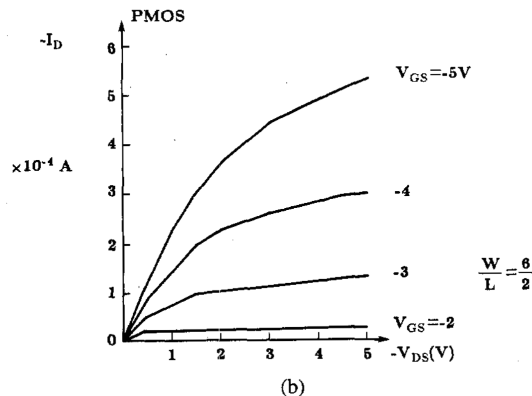


Fig. 11. Simulated output characteristics of MOS transistors.

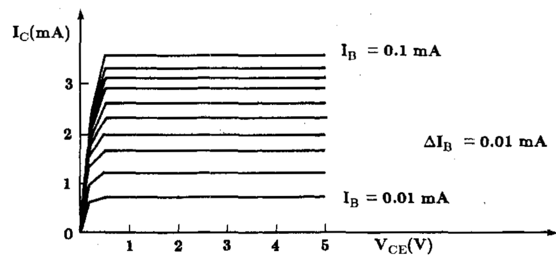


Fig. 12. Simulated output characteristics of n-p-n transistor.

The total pull-down time is

$$T_{d,PD} = T_1 + T_2 + T_3 = \frac{(C_{BC1} + C_{BE1} + C_{BC2} + C_{CS2} + f_2 C_L) f_0 V_d}{I_{D\infty}} + \frac{(C_{BC1} + C_{BC2} + C_{CS2} + C_L)(2V_{dd} - V_{Tn} - V_d(4 + f_1))}{(\beta + 1) I_{D\infty}}. \quad \text{(A4)}$$

The curve fitting factors for a 10-percent match between closed-form and simulation results for the pull-up transient are given in Table V. The device dimensions, in micrometers, for the optimized CMOS-bipolar buffer are given in Table VI.

TABLE V

C_L(pF)	f_0	f_1	f_2
1	0.8	1.0	0.7
3	0.8	1.0	0.3
5	0.8	1.0	0.15
7	0.8	1.0	0.15
15	0.5	0.75	0.10
.	.	.	.
.	.	.	.
.	.	.	.
100	0.5	0.75	0.10

TABLE VI

C_L(pF)	L_E	W_{M1}	W_{M2}	$W_{M3} = W_{M4}$
1	1.42	6	2	2
3	2.56	6	3	2
5	4.0	6	3	2
10	4.2	6	3	3
20	7.8	6	3	4
30	7.9	6	3	9
50	8.0	6	4	16
100	8.1	6	4	18

Appendix III
CMOS Cascade Transient Analysis

The pull-dowm and pull-up transients of the CMOS cascade were modelled by the circuits in Fig. 9(b) and (c), respectively. Assuming the gates of the PMOS and NMOS devices are related by $W_p = (\mu_n W_n)/\mu_p$. The pull-down and pull-up times of the first and second stages, respectively, are

$$T_1 = \frac{VLW_P^1\left(1 + \frac{\mu_p}{\mu_n}\right)}{k_n V_L W_n^0 V_G} \quad \text{(A5)}$$

$$T_2 = \frac{VC_L V_{dd}}{k_p V_L C_0' W_p^1 V_G}. \quad \text{(A6)}$$

Following Elmore's approach [12] the average pull-up time is given by $T_{d,c} = \sqrt{T_1^2 + T_2^2}$, i.e.

$$T_{d,c} = \frac{V}{V_L V_G} \left[\left(\frac{LW_p^1\left(1 + \frac{\mu_p}{\mu_n}\right)}{k_n W_n^0}\right)^2 + \left(\frac{C_L}{k_p C_0' W_p^1}\right)^2\right]^{1/2}. \quad \text{(A7)}$$

By differentiating (A7) with respect to W_p^1, a minimum is found if

TABLE VII

C_L(pF)	$W^1_p(\mu m)$	$W^1_n(\mu m)$
1	31.3	12.5
3	54.2	22
5	70	28
10	99	40
15	121	48.5
25	157	63
50	221	88.5
70	262	105
100	313	125

$$W_p^1 = \sqrt{\frac{C_L k_n W_n^0}{C_0' k_p L\left(1 + \frac{\mu_p}{\mu_n}\right)}} \tag{A8}$$

and the minimum pull-up time is

$$T_{d,c\min} = \sqrt{2}V\sqrt{\frac{C_0' L\left(1 + \frac{\mu_n}{\mu_p}\right)}{I_{Dn\infty}' I_{D\infty}}}\sqrt{C_L} \tag{A9}$$

where $I_{Dn\infty}' = k_n v_L C_0' V_G$.

The Optimal CMOS-Cascade device dimensions are given in Table VII.

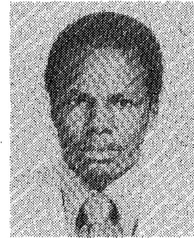
Acknowledgment

Discussions with S. T. (Mike) Liu, Honeywell, Plymouth, MN, are gratefully acknowledged. The authors thank the reviewers for many most helpful comments.

References

[1] H. B. Bakoglu and J. D. Meindl, "Optimal interconnect circuits for VLSI," in *1984 ISSCC Dig. Tech. Papers*, pp. 164-165.
[2] H. Higuchi *et al.*, "Performance and structures of scaled-down bipolar devices merged with CMOSFETs," in *IEDM Tech. Dig.*, pp. 694-697, 1984.
[3] J. Miyamoto *et al.*, "A 1.0 μm n-well CMOS/bipolar technology for VLSI circuits," in *IEDM Tech. Dig.*, pp. 6-66, 1983.
[4] F. Walczyk *et al.*, "A merged CMOS/bipolar VLSI process," in *IEDM Tech. Dig.*, pp. 59-62, 1983.
[5] A. R. Alvarez *et al.*, "2 micron merged bipolar-CMOS technology," in *IEDM Tech. Dig.*, pp. 761-764, 1984.
[6] G. Zimmer, B. Hoefflinger, and J. Schneider, "A fully implanted NMOS, CMOS, bipolar technology for VLSI of analog-digital systems," *IEEE J. Solid-State Circuits*, vol. SC-14, pp. 312-318, Apr. 1979.
[7] H. C. Lin *et al.*, "Complementary MOS-bipolar transistor structure," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 945-951, Nov. 1969.
[8] A. Vladimirescu, K. Zhang, A. R. Newton, D. O. Pederson, and A. Sangiovanni-Vincentelli, *SPICE Version 2G User's Guide*, Univ. of Calif., Berkeley, CA, Aug. 1981.
[9] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1981, pp. 138-148.
[10] B. Hoefflinger, "Output characteristics of short channel field effect transistors," *IEEE Trans. Electron Devices*, vol. ED-28, no. 28, pp. 971-976, Aug. 1981.
[11] B. Hoefflinger, S. T. Liu, and B. Vajdic, "A three-dimensional CMOS design methodology," *IEEE J. Solid-State Circuits*, vol. SC-19, pp. 37-39, Feb. 1984.
[12] W. C. Elmore and M. Sands, *Electronics: Experimental Techniques*, 1st ed. New York: McGraw-Hill, 1949, p. 139.

*

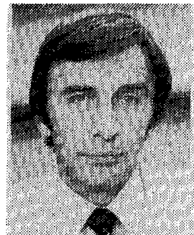


Hector J. De Los Santos (S'85-M'85) was born in Santo Domingo, Dominican Republic, in 1957. He received the B.S. degree in electrical engineering from the University of Puerto Rico, Mayaguez Campus, in 1979, the M.S.E. degree in electronic circuits from the University of California at Los Angeles in 1981, and the M.S.E.E. degree in solid-state devices from Purdue University, West Lafayette, IN, in 1985. He attended the University of California as Hughes Fellow. He is currently working toward the Ph.D. degree at Purdue.

While at Hughes, he performed analysis and design for the electronic circuitry of switching power supplies, worked on the implementation of a D/A converter test station including the design of a sampling voltage tracker, and designed a controller-motor interface. In 1981, he joined Intel Caribbean, Inc., Las Piedras, Puerto Rico, as a start-up Product Engineer. While at Intel Caribbean he was responsible for the test of SRAM's and switched capacitor filters and codecs. In 1984, he joined Purdue University, where he is a Fellow (GPOP) and a Research Assistant. His interests are analog and digital integrated circuit design, solid-state electronics, and instrumentation.

Mr. De Los Santos is a member of Tau Beta Pi.

*



Bernd Hoefflinger was born in 1939 in Bucharest, Romania. He studied physics in Goettingen and Munich and received the diploma degree from the University of Munich, West Germany, in 1964, and the Doctor rer. nat. degree from the Technical University of Munich in 1967.

From 1964 to 1967, he was associated with the Siemens Research Laboratory in Munich. From 1967 to 1970, he was Assistant Professor in the School of Electrical Engineering, Cornell University, Ithaca, NY. From 1970 to 1972, he was Manager of the MOS Integrated Circuits Department, Semiconductor Division, Siemens AG, Munich. In 1972, he became Professor of Electrical Engineering and the founder and first dean of the Electrical Engineering Department at the University of Dortmund, West Germany. There he was also director of the Electron Devices Laboratory. In 1979-1980, he spent a half-year sabbatical with the Electronics Research Laboratory at the University of California, Berkeley. From 1981 to 1983, he was head of the Department of Electrical Engineering and co-director of the Microelectronics and Information Sciences Center at the University of Minnesota, Minneapolis, MN. He became Head of the School of Electrical Engineering at Purdue University, West Lafayette, IN, in 1984. In the fall of 1985, he was appointed Director of the Institute for Microelectronics, a public research foundation, in Stuttgart, West Germany. He has authored or coauthored over 100 publications and has edited and coauthored two books on microelectronics.

Dr. Hoefflinger received the 1968 Award of the German Nachrichtentechnische Gesellschaft and the 1969 IEEE ISSCC Outstanding Paper Award. He is the co-recipient of the 1980 Darlington Prize of the IEEE Circuits and Systems Society and of the 1982 Electronics Letters Premium of the British IEE. He became a member of the Dusseldorf Academy, one of the five German Academies of Science, in 1981. From 1973 to 1977, he was an Associate Editor of the IEEE Transactions on Electron Devices.