NOTICE

AVIS

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

NEW BICMOS DRIVER CIRCUIT WITH IMPROVED ANALYTICAL DELAY MODEL

by

Z.X. Yan

B.Sc., Shanghai University of Science and Technology, 1963

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED SCIENCE

in the school

of

Engineering Science

© Zhixin Yan 1991

Simon Fraser University

November 1991

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
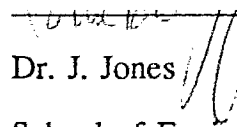
Canada

APPROVAL

Name: Zhixin Yan

Degree: Master of Applied Science (Electrical Engineering)

Thesis Title: New BiCMOS Driver Circuit with Improved Analytical Delay Model

Examining Committee:

Committee Chairman:

Dr. J. Jones
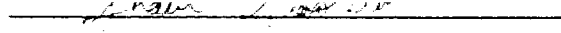
School of Engineering Science, SFU

Senior Supervisor:

Dr. M. J. Deen

School of Engineering Science, SFU

Committee Member:

Dr. S. Stapleton

School of Engineering Science, SFU

External Examiner:

Colin Harris

Pacific Microelectronics Centre, MPR

Date Approved: 1/12/91

Title of Thesis/Project/Extended Essay

New BiCMOS Driver Circuit with Improved Analytical

Delay Model

_____

Author: _____

(signature)

Zhixin Yan

(name)

12 December 1991

(date)

ABSTRACT

BiCMOS technology which combines Bipolar and CMOS transistors in a single integrated circuit is now proposed as a very promising candidate for VLSI circuits with speed-power-density performance that was previously unattainable with either technology individually.

A new BiCMOS driver circuit is proposed in this thesis. This new drive circuit uses a "dynamic-resistor" device to replace the original "on-resistor" device in conventional BiCMOS drivers. This dynamic-resistor-behavior was achieved using depletion mode PMOS devices and making appropriate feedback connections to control its switching current properly. A comparison of the proposed and two typical BiCMOS driver circuits using SPICE shows that the major features for this new BiCMOS driver circuit are: (1) better noise margin for DC characteristics; (2) faster falling edge for transient response; (3) more compact structure with merged PMOS-NPN devices; and (4) larger tolerance for additional process step control. The detailed DC and transient analysis of this new BiCMOS driver circuit is also presented and discussed.

In addition, an improved analytical model for the propagation delay τ_d suitable for BiCMOS driver with large capacitance load was also developed in this thesis. This new model explicitly incorporates the high-level injection effect (so called the Kirk effect) into Ebers-Moll bipolar transistor model, but still maintains its simple formulation. By comparison with the SPICE simulation results, this new model has the following advantages: (1) it has better accuracy (less than 10%) with respect to those important parameter variations of β_F (the forward current gain), I_{KF} (the Knee current), τ_F (the forward transit time) and I_{DS} (the drain current); (2) it is simple, easy to use and requires no empirical constants; (3) the effect of τ_F on τ_d is inherent in this model.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

INTRODUCTION

BiCMOS technology combines both bipolar and CMOS transistors in a single integrated circuit for improved performance. It is now becoming an increasingly attractive technology for VLSI as shown in Figure 1.1. This is mainly because BiCMOS-based VLSI circuits have speed-power-density performance that was previously unattainable with either technology individually [1.1-1.3].
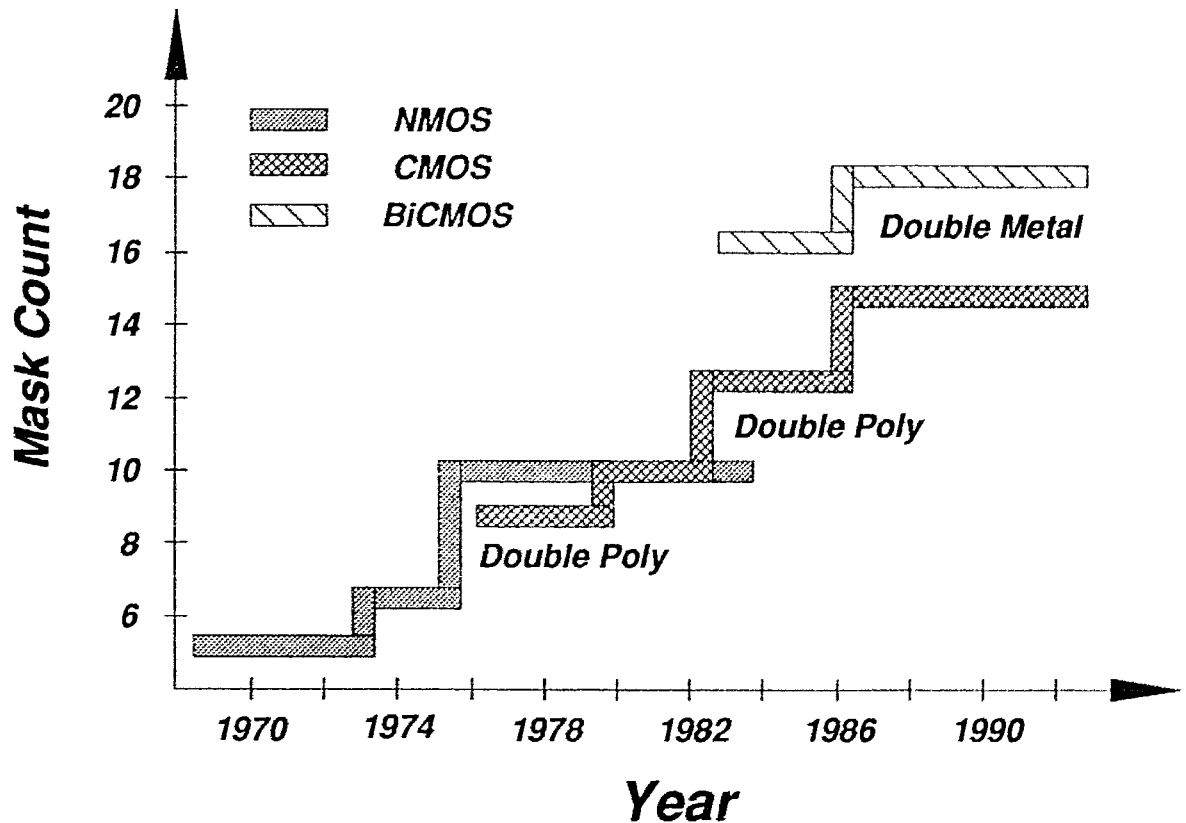
**Figure 1.1 Evolution of digital MOSFET technology
from NMOS to CMOS to BiCMOS**

1.1 Main features of BiCMOS

1.1.1 Comparison between Bipolar and MOS transistors

In order to understand the main features of BiCMOS technology, it is necessary to be knowledgeable of those important differences between bipolar transistors (BJT) and MOS transistors (MOSFET) [1.1] as shown in Table 1.1:

CHARACTERISTICS	BJT	MOSFET
1. Type of signal	Current	Voltage or Charge
2. Transconductance (g_m)	I_C/V_T ($V_T = k_B T/q$)	$\sqrt{2kI_D}$ ($k = \mu C_{ox} W/L$)
3. Output resistance (r_o)	V_A/I_C	$(\lambda I_D)^{-1}$
4. Input resistance (r_i)	V_T/I_B	∞
5. Intrinsic Gain (A_o)	$\approx V_A/V_T$	$\sqrt{2k/I_D}/\lambda$
6. Frequency Response (ω_T)	$2D_n/W_B^2$	$3\sqrt{I_D/2k}/L$
7. Noise ($v_{eq}^2/\delta f$)	$4k_B T r_b + 2qI_C/g_m^2$	$4k_B T/(1.5g_m) + K/(WLC_{ox}f)$
8. Mismatch tolerance	δV_{BE} (smaller)	δV_{TH} (larger)
9. Device size	larger	smaller
10. Major process steps	≥ 17	≥ 18
11. Latch-up susceptibility	no	yes
12. Temperature Sensitivity	good	poor

Table 1.1 Comparisons between BJT and MOSFET characteristics

Based on the above tabulation, it can be appreciated why bipolar transistors are usually used for high speed, large current drive, low noise, and high performance analog circuits. On the other hand, MOS devices are much more suitable for VLSI system with low power dissipation, high packing density, especially for a CMOS circuit (combining both NMOS and PMOS devices) which has extremely low quiescent power consumption.

1.1.2 Advantages and disadvantages of BiCMOS

Compared with either CMOS or bipolar technology that are now used for VLSI, BiCMOS technology offers the following advantages:

    i) higher speed over CMOS;

    ii) lower power dissipation than Bipolar;

    iii) flexible I/O's (TTL, CMOS, ECL);

    iv) high performance analog circuits; and

    v) immunity to latch-up.

On the other hand, due to the additional process steps required to build both CMOS and bipolar devices, the main drawbacks for BiCMOS technology are:

    i) higher costs; and

    ii) longer fabrication cycle time.

Although the above disadvantages have limited the success of BiCMOS's at present, it could be improved in the near future because both the technology and CAD tools associated with BiCMOS would be improved and become comparable to that available for either CMOS or Bipolar technologies. Note however that the advantages with BiCMOS are unique and significant, especially for the higher performance required in advanced VLSI systems. Therefore, it is expected that BiCMOS technology would be of ever-growing practical importance.

1.2 Advanced BiCMOS technology

1.2.1 P-well BiCMOS technology
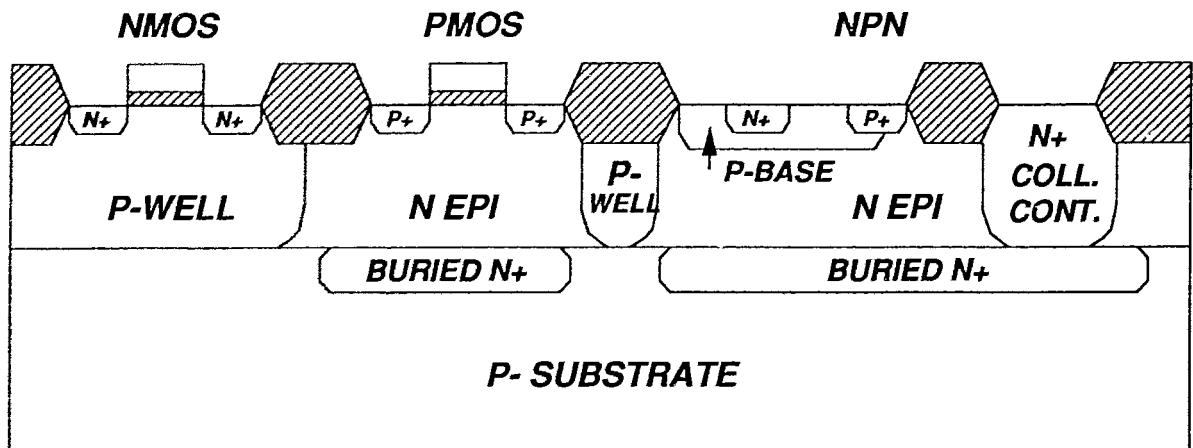
Figure 1.2 BiCMOS device structure with P-well CMOS process

Figure 1.2 schematically shows the typical BiCMOS structure with P-well CMOS process. Compared to typical CMOS technology, the following four process steps can be shared for both CMOS and NPN devices:

i) the PMOS N-epi process and the NPN collector;

ii) the NMOS P-well process and the NPN isolation steps;

iii) the NMOS N+ source/drain, the NPN emitter and collector contact steps; and

iv) the PMOS P+ source/drain and the NPN base contact steps.

In addition, three extra process steps are needed in order to form good BiCMOS circuits:

i) heavily-doped buried N+ layer for minimizing NPN collector resistance and also reducing the susceptibility to latch-up of CMOS;

ii) P-base region required for NPN transistor;

iii) Deep N+ connection to the buried N+ subcollector added for further reduction of collector resistance.

4

Therefore this P-well BiCMOS approach merges the process steps needed to achieve low bipolar collector resistance with those required to reduce CMOS latch-up susceptibility.

1.2.2 Twin-well BiCMOS technology
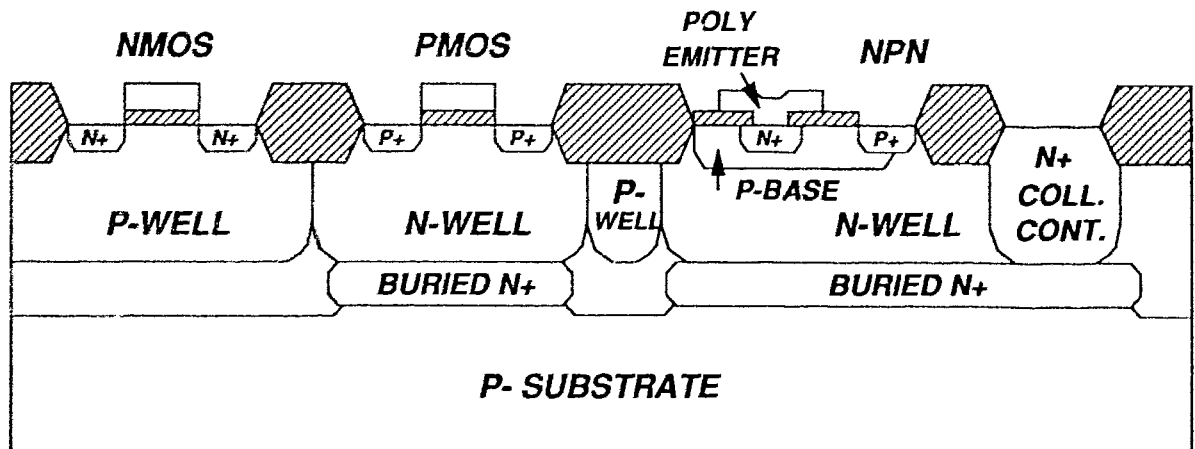
Figure 1.3 BiCMOS device structure with twin-well CMOS process

Figure 1.3 shows an optimized BiCMOS structure with both P and N-wells (twin-well CMOS). As compared with above single well structure, a key feature included in this design is the self-aligned P and N+ buried layers for improved packing density, separately optimized N and P-wells formed in an epitaxial layer with intrinsic background doping, and a polysilicon emitter for improved bipolar performance. In total four additional mask levels (buried N+, deep N+ contact, P-base and poly-emitter) are required to merge this higher performance BiCMOS process with a baseline CMOS process flow.

1.2.3 State-of-the-art BiCMOS technology

In Table 1.2 below, a listing of some of the state-of-the-art BiCMOS technology that have been recently reported are presented to give an idea about the processing features and design details:

Manufactures [Ref.]	Process feature	L (μm)	f_t (GHz)	t_d (ps)	Source
Siemens AG [1.4]	MCSL: merged CMOS/bipolar logic (adder).	1.5	9	150(car.) 700(sum)	1989 ISSCC
AT&T Bell [1.5]	NOVA: Non-OVerlapping super self-Aligned Structure.	1.5	12	87(ECL) 128(CMOS)	1988 IEDM
Siemens [1.6]	Buried layer, LDD NMOST, Implanted E/B (LC) or Polysilicon self-aligned E/B (HP)	1.2	10	ECL: 300(LC) 65(HP)	1989 IEEE
Fujitsu [1.7]	DOPOS: Doped polysilicon emitter structure.	1.0	6	500(CMOS)	1988 Fujitsu J.
Hitachi [1.8]	BiCMOS Macrocell library	1.0	9	70MHZ/ 32bit μP	1989 ISSCC
NTT LSI [1.9]	SST: Super Self-aligned process technology.	1.0	13(N) 5(P)	230(BiCMOS) 39/76(N/P)	1988 IEDM
LSI [1.10]	Bipolar-PMOS merged structure	0.8	8	250(BiCMOS)	1990 IEEE
Fairchild [1.11]	Optical polysilicon buried-contacts to both N+ and P+, Walled polysilicon emitter.	0.8	9	150(ECL) 75(CMOS)	1989 IEEE-ED
Signetics [1.12]	HS4: A single polysilicon layer advanced super high speed.	0.8	13	no date	1989 IEEE
NS [1.13]	ASPECT III: with self-aligned integrated well taps, silicided local interconnect,	0.8	14	50(ECL) 90(CMOS) 150(BiCMOS)	1990 IEEE

Tektronix [1.14]	Self-aligned double-polysilicon bipolar device.	0.8	16	65(ECL)	1988 IEDM
Toshiba [1.15]	BiPMOS: A base and drain merged bipolar-PMOSFET structure	0.7	10(P)	no data	1990 IEDM
Motorola [1.16]	A triple polysilicon process	0.5	14	115(ECL)	1990 IEDM
IBM [1.17]	Merged complementary BiCMOS: merging the NPN base with PMOS source, and PNP base with NMOS drain.	0.5	17	250(BiCMOS)	1990 IEDM
Philips [1.18]	SABR: using ultra-thin epitaxy and polysilican buffer LOCOS.	0.5	18	57(CMOS) 182(BiCMOS)	1990 IEDM
NTT LSI [1.19]	HSST: self-aligned double-poly bipolar process using e-beam direct writing tech.	0.2	22	25(ECL) 44(CMOS) 66(BiCMOS)	1990 IEDM

Table 1.2 The state-of-the-art BiCMOS technology

1.3 Typical applications of BiCMOS

Based on the advanced BiCMOS technology which combines both bipolar and CMOS devices, it is clear BiCMOS fills the market niche between very high speed, but power hungry bipolar ECL, and the very high density, medium speed CMOS. It basically covers the following area:

1.3.1 Digital systems

BiCMOS can provide memories with faster access times at the same power, or the same access times at lower power. Since the bipolar transistor provide large output drive current, BiCMOS logic gates are effectively applied at high capacitive nodes like decoders, wordline drivers, write drivers and output buffers in memories. BiCMOS memories now covers the full density spectrum from 1K to 1M at access time below 10ns. A typical example is the ~8ns 256K and 1M bit BiCMOS ECL SRAMs were successfully demonstrated in 1988-1989 [1.20-1.21].

In the area of semi-custom chip, BiCMOS offers performance and I/O flexibility advantages over CMOS, or lower power and higher density at nearly the same speed level compared with bipolar. Initially, the internal core section was CMOS-only with the I/O employing the bipolar devices as required. The current trend focuses on the use of bipolar devices in large macro functions (for example as adder, register, ST flip-flops) to make more efficient use of the bipolar devices in an average option and will be used in very high (30K-100K gates) density arrays.

Microprocessor (μP) is an another important area for using BiCMOS to enhance the performance by reducing delay through critical paths, reducing clock skews, and increasing I/O throughput. For example, BiCMOS can utilize CMOS design methodologies used to design previous generations of μP's while also providing ECL data paths as necessary, or integrate large and fast on-chip RAMs, or keep the power manageable by focusing it where required to optimize the speed-power tradeoff. Recently, a test chip of a BiCMOS μP with 25ns (worst case) machine cycle time was demonstrated by Hitachi [1.22].

Fig. 1.4 shows a digital circuit example of BiCMOS adder cell with a new MCSL logic (merged CMOS/Bipolar logic) [1.4]. The circuit for carry and sum is realized using three-level series gating. In the lower two levels, the MOS inputs control the current path. In the third level, only the bipolar transistors are used to provide fast propagation of the carry bit to the next adder cell. The ECL sum output communicates directly with fast bipolar logic. An optional ECL/CMOS level converter makes it possible to work in a CMOS environment.
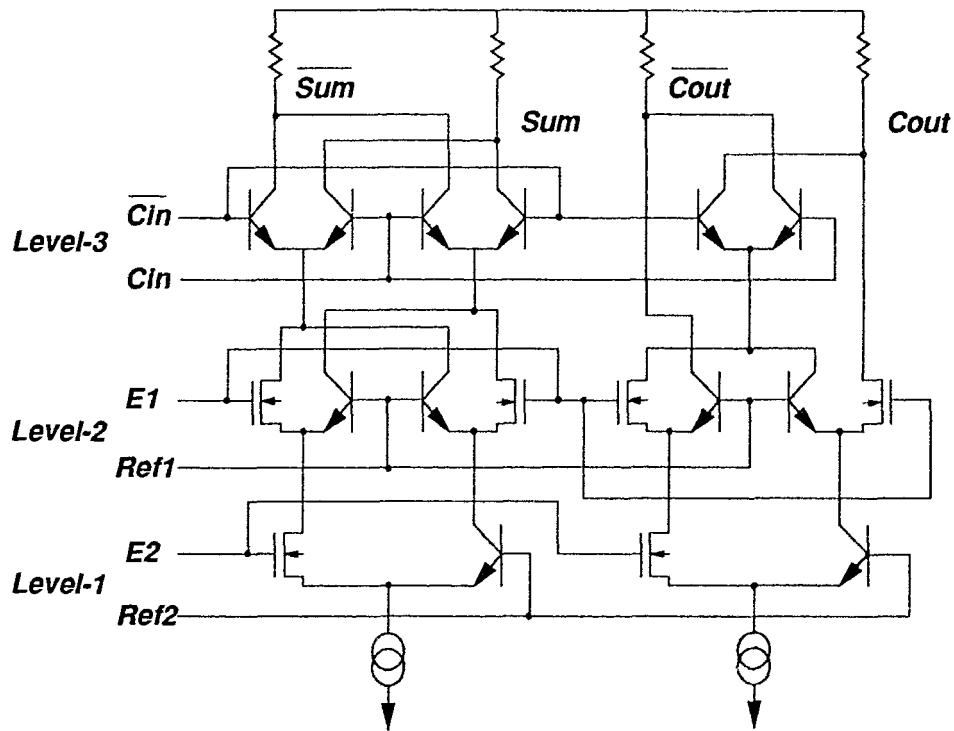
Figure 1.4 BiCMOS adder cell

1.3.2 Analog/Digital systems

In addition, it is expected that BiCMOS technology will play an increasingly important role for analog/digital VLSI systems in future. This is simply because BiCMOS can offer combined advantages of both bipolar and CMOS circuit design like

i) high gain of the bipolar;

ii) >1GHz toggle frequency;

iii) lower 1/f noise;

iv) high impedance inputs (FETs) for sample and hold applications;

v) low input offsets voltage for differential pairs;

vi) "zero offset" voltage analog switches;

vii) good voltage references; and

viii) minimal current drain, automatically scalable for low and medium speed digital control logic.

These design advantages can be used in communication circuits, such as a single chip direct divide synthesizer. A further example of BiCMOS's flexibility is in the fast growing realm of oversample noise shaping and predictive coders. The circuit requirements for the coder itself demands low-noise high-speed linear devices, but these must be followed by a complex digital filter. Thus advanced coders of this type would reside in two packages unless BiCMOS is employed.

Fig. 1.5 shows an analog circuit example of BiCMOS operational amplifier [1.1]. The PMOS input stage is used to provide infinite input resistance, while the bipolar second stage with high transconductance offers higher frequency response. The input resistance of the second stage amplifier is $r_b + r_\pi$ (on the order of a few kilo ohms), which is much smaller than the output resistance of the first stage. To reduce this loading, an emitter-follower is added. The frequency response of this circuit should be better than the source-follower buffered bipolar design due to the absence of the potentially low frequency parasitic pole. The emitter degeneration resistors in the current mirror are added to reduce the noise contribution from the bipolar current mirror. The PMOS input differential pair also improves slew rate compared with a bipolar differential pair.

10
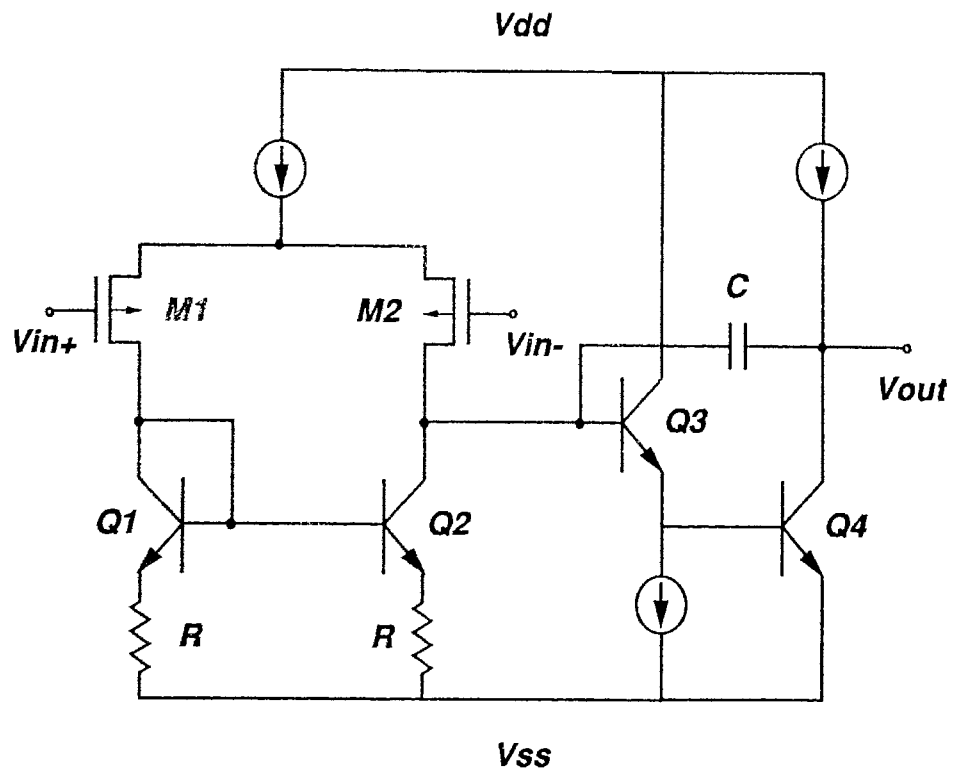
Figure 1.5 BiCMOS operational amplifier

CHAPTER 2:

NEW BiCMOS DRIVER CIRCUIT

2.1 Introduction

The new BiCMOS driver circuit, proposed as part of this research, is described in detail in this chapter. After reviewing the basic features of the typical BiCMOS driver circuits (which are widely used at present) and the improved BiCMOS driver circuit (which was designed by one company recently), this new BiCMOS driver circuit is introduced in section 2.2. This new driver circuit uses a "dynamic-resistor-behavior" device to replace the original "on-resistor-behavior" device in the improved BiCMOS driver. This dynamic-resistor-behavior was achieved using depletion mode PMOS devices and making appropriate feedback connections to control its switching current properly. SPICE simulation results show that this new BiCMOS driver circuit has distinct advantages compared with the other two BiCMOS drivers: (1) better noise margin for DC characteristics; (2) faster falling edge for transient output response; (3) more compact structure with merged PMOS-NPN devices; and (4) larger tolerance for additional process step control. In the following sections 2.3 - 2.6, these features will be qualitatively discussed.

2.2 Schematic circuit diagram

2.2.1 Typical BiCMOS driver circuit

Figure 2.1 shows the typical BiCMOS driver circuit (to be referenced to as a BCD-4 in future) that is currently widely used in VLSI digital systems [2.1-2.3]. It consists of four enhancement MOSFETs (three NMOS and one PMOS devices) and two NPN BJTs.

Basically this driver uses the CMOS inverter to perform the logic function, and the bipolar transistors to drive the output load. The logic operation principal for this BiCMOS driver can be described simply as follows:
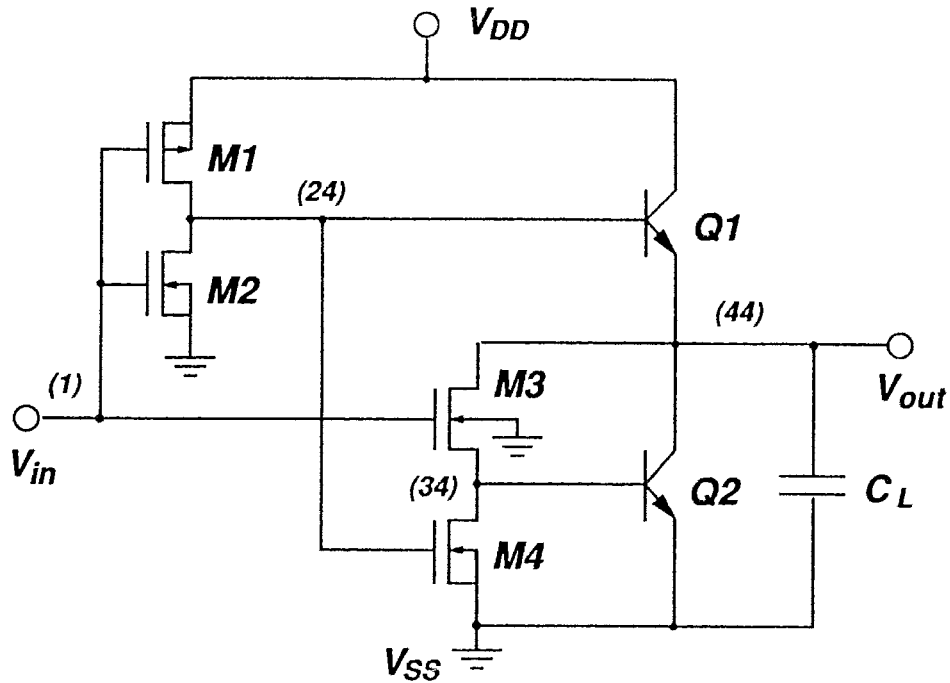
Figure 2.1 Typical BiCMOS driver circuit (BCD-4)

A. When V_{in} is low (~0V), M1 (PMOS) is on, M2 and M3 (NMOS) are off, so the base voltage of Q1 is high (= V_{DD}) due to CMOS inverter function. This high voltage then turns Q1 and M4 (NMOS) on. Meanwhile, because M3 is off and M4 is on, the base voltage of Q2 is pulled down to ground and this forces Q2 to be off. It results in no noticeable current flowing through Q1, but some inevitable leakage current. Therefore the logic "1" output voltage V_{OH} is equal to $V_{DD} - V_{BEO1}$, where V_{BEO1} represents the turn-on base-emitter voltage (about 0.5-0.6V) for Q1. This V_{OH} value is obviously lower than the corresponding output voltage of $V_{OH} = V_{DD}$ for a typical CMOS inverter.

B. When V_{in} is high (~5V), M1 is off and M2, M3 are on, the base voltage of Q1 is pulled down to V_{SS} (0V) and it turns both Q1 and M4 off. Since M3 is on and M4 is off, Q2 is now in the active region, but also with only unavoidable leakage current flow because Q1 is off. Therefore the logic "0" output voltage V_{OL} would approach to V_{BE02} above ground, here V_{BE02} represents the turn-on base-emitter voltage (about 0.5-0.6V) for Q2 required to flow its leakage current. The V_{OL} value here is higher than the $V_{OL} = 0$V for a typical CMOS inverter.

C. When the input step waveform $V_{in}(t)$ goes from high to low, the node voltage $V(24)$ of the CMOS inverter output will correspondingly change from low to high. This results in turning on both Q1 and M4, and cutting off Q2. The output voltage $V_{out}(t)$ would then rise up by charging the load capacitor C_L through Q1. The charging current for C_L is equal to I_{E1} ($= I_{C1} + I_{B1}$), which is approximately β_F times of the drain current I_{DS1} (because I_{DS1} is equal to I_{B1}). It is obvious that the rising response of $V_{out}(t)$ in this BiCMOS driver should be much faster than the corresponding one with only the CMOS inverter output.

D. Similarly, when the input step waveform $V_{in}(t)$ from low to high, the node voltage $V(24)$ would correspondingly change from high to low. It turns off Q1 and M4, and also turns on Q2 (because M3 is already on). Then the output voltage $V_{out}(t)$ falls by discharging the load capacitor C_L through Q2. This discharge current ($= I_{C2} + I_{DS3} \approx \beta_F I_{DS3}$) is of course much larger than the drain current I_{DS2} which would be the discharge current of the CMOS (M1-M2) inverter.

Based on the description above, the main features for this typical BiCMOS driver compared with typical CMOS driver is: the static logic swing ($=V_{OH}-V_{OL}$) and related

DC noise margin is somewhat reduced, but the transient behavior is improved significantly for larger capacitance loads. That is why this BiCMOS circuit shown in Fig. 2.1 is also called a driver.

2.2.2 Improved BiCMOS Driver circuit
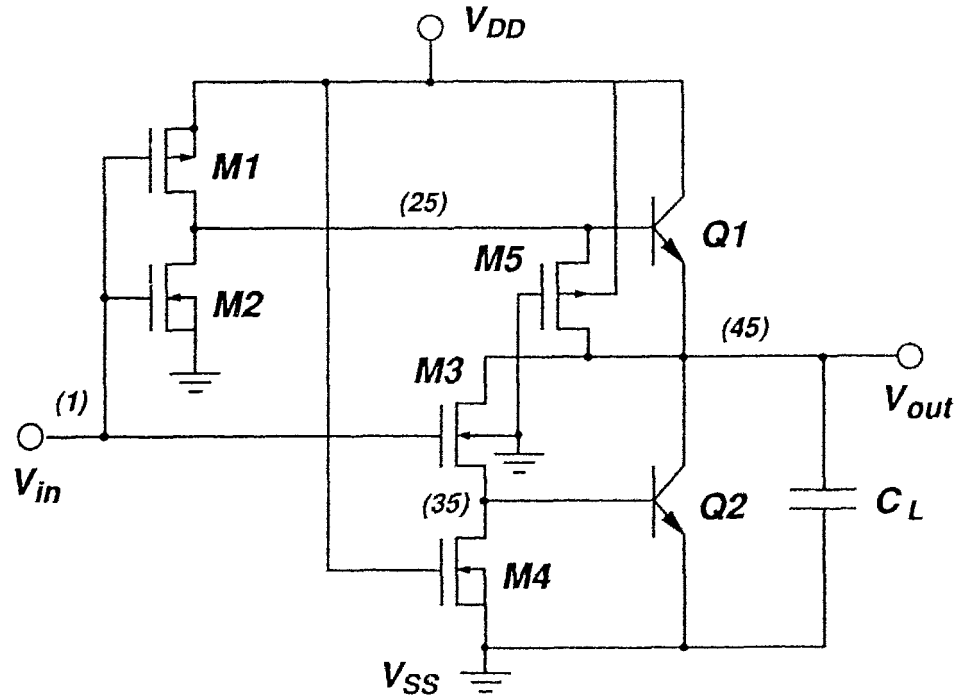
Figure 2.2 Improved BiCMOS driver circuit (BCD-5)

In order to achieve full rail-to-rail (V_{SS} to V_{DD}) static logic swing as typical CMOS inverter does, an improved BiCMOS driver was proposed by an industrial organization and is shown in Fig. 2.2 (short for BCD-5). The differences between this driver and the typical one in Fig. 2.1 are:

A. The gate of M4 (NMOS) is connected to V_{DD} instead of to the output node (24) of CMOS (M1-M2) inverter in Fig. 2.1;

B. The additional enhancement mode PMOSFET M5 is placed between the base and the emitter terminals of Q1, with its gate terminal connected to ground V_{SS}.

By examining this improved BiCMOS circuit diagram, it is easy to understand that both M4 (NMOS) with gate connected to V_{DD} and M5 (PMOS) with gate connected to V_{SS} are always operating in the active region and act as "on-resistors". Since there is no DC current flowing through this circuit for the static status of either logic "0" or logic "1" output as described above in typical BiCMOS driver, these two "resistor-like" devices of M4 and M5 would pull the V_{BE} voltage of Q2 or Q1 down to zero respectively. Therefore the logic "0" or logic "1" output voltage will now approach to either $V_{OL} = V_{SS} = 0V$ or $V_{OH} = V_{DD} = 5V$, the same as the typical CMOS inverter output.

On the other hand, during the transient rising and falling response periods, these two "on-resistor-behavior" devices of M4 and M5 in BCD-5 would also take some portion of the drain current from I_{DS1} or I_{DS3} when Q1 or Q2 is on, because the voltage across the M4 or M5 is equal to $V_{BE} \approx 0.8V$. Therefore during the rising time of output voltage response, the base current I_{B1} $(= I_{DS1}-I_{DS5})$ and its collector current I_{C1} for Q1 are reduced somewhat. The same argument is also true for the falling time of output voltage response, with reduced base current I_{B2} $(= I_{DS3}-I_{DS4})$ and related collector current I_{C2} for Q2. That means the transient response rise and fall-times in BCD-5 would not be as fast as that with typical BiCMOS driver of BCD-4. Nevertheless, with careful design of making I_{DS4} and I_{DS5} much smaller than I_{DS1} and I_{DS2}, and also taking high-level injection effect of NPN BJT into account (which means the current gain β_F would be degraded after I_C increases over the Knee current I_{KF}), this negative effect on transient response in BCD-5 might not be so significant, especially for larger capacitance loads.

In summary, this improved BiCMOS driver shown in Fig. 2.2 indeed has full rail-to-rail static logic output of V_{SS} to V_{DD}, and still shows much larger current drive capability for fast switching response.

2.2.3 New BiCMOS driver circuit

In terms of the above discussion, the function of M4 in the BiCMOS driver of BCD-5 is actually two-fold: on the one hand, M4 should be on when the output response $V_{out}(t)$ goes towards logic "0" voltage, in order to reach the lowest output value of $V_{OL} = V_{SS}$; on the other hand, M4 should be off during the transition time period of output response $V_{out}(t)$ from high to low, in order to get the largest base current for Q2. There is a similar situation for M5. By carefully examining these two opposite-like requirements for M4 (or M5), it was found that these two conflicting requirements actually take place in different time periods of the output voltage response. The question "is it possible to design M4 (or M5) like a dynamic resistor to meet the above contradiction demands?" can then be raised. To solve this problem, the new BiCMOS driver developed in this thesis was proposed.
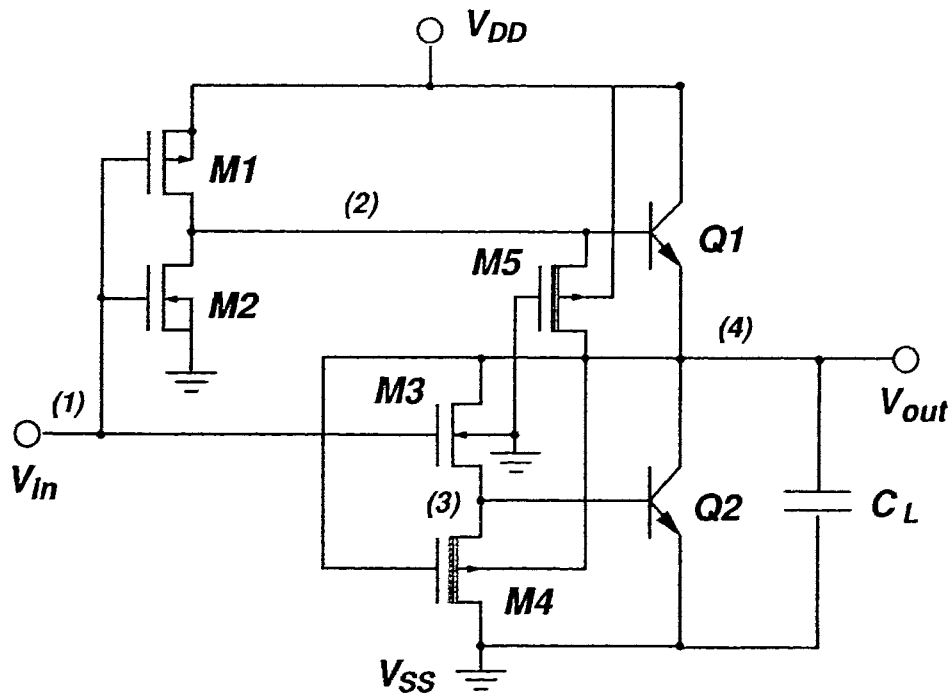
Figure 2.3 New BiCMOS driver circuit (BCD-N)

Figure 2.3 shows this new BiCMOS driver circuit (short for BCD-N). It also consists of five MOS devices and two NPN BJTs, but using two depletion mode PMOS devices to replace original enhancement mode M4 and M5 in BCD-5. The key points for this new driver circuit are:

A. By connecting the gate of M4 to the output node, this depletion mode PMOS device is now controlled by the output voltage to be turned either on with lower $V_{out}(t)$, or off with higher $V_{out}(t)$. The specific range of the output voltage for distinguishing these two cases are determined by the threshold voltage V_{TD} of depletion mode PMOS device M4;

B. With the connection of the substrate of M4 also to the output node, and substituting the original enhancement mode M5 with the same depletion mode PMOS device as M4, both M4-Q2 and M5-Q1 devices show the same connection and can be merged together for more compact structure.

A point to note here is that this proposed BiCMOS driver circuit solves the problem related to M4 in BCD-5, and it has almost nothing to do with M5 in BCD-5. In other words, for this new BiCMOS driver shown in Fig. 2.3, M4 now functions as a "dynamic-resistor-behavior" device, but M5 is still an "on-resistor-behavior" device. The reason for not solving the problem related to M5 in BCD-N is twofold.

First, it is not as easy as being done for M4 in BCD-N because its gate terminal can not be connected directly to the output node (it should be inversely controlled by $V_{out}(t)$ according to its logic function);

Second, the negative effect of M5 on the rising output response in BCD-5 is less serious as compared with M4, and this will be explained more detail in the following section 2.4.

In summary, the major advantages of this new BiCMOS driver shown in Fig. 2.3 are:

i) better noise margin for DC characteristics;

ii) faster falling edge for transient response;

iii) more compact structure for merged PMOS-NPN devices; and

iv) larger tolerance for V_{TD} process control.

To verify the above feature, SPICE [2.4] was used to simulate all DC and transient characteristics for this new BiCMOS driver circuit, together with those of BCD-4 and BCD-5 for comparison. Table 2.1 lists the input data file with all parameters for MOSFET and NPN devices required in SPICE simulation, which were either chosen from the typical values or referenced from the industrial company. For example, the threshold voltage value of V_{TE} used in the simulation was slightly different for NMOSFET (V_{TEN} = 0.8V) and for PMOSFET (V_{TEP} = -0.85V).

Input BiCMOS data file for SPICE Simulation

***** PARAMETER *****

.STEP PARAM VTD 0 2 1

.PARAM CLOAD=10PF VTEN=0.8V VTEP=-0.85V VTD=1.V

.PARAM L1=0.8U L2=1.6U W1=4U W2=1.4U

***** BIAS *****

VCC 5 0 5V

VIN 1 0 PULSE 5 0 2NS 0N 0N 20NS 40NS

```
*****           BCD-N              *****

M1-N 2 1 5 5    PMOSE L={L1} W={W1}

M2-N 2 1 0 0    NMOSE L={L1} W={W1}

M3-N 4 1 3 0    NMOSE L={L1} W={W1}

M4-N 0 4 3 4    PMOSD L={L2} W={W2}

M5-N 4 0 2 5    PMOSD L={L2} W={W2}

Q1-N 5 2 4      NPN

Q2-N 4 3 0      NPN

CL-N 4 0        {CLOAD}


*****           BCD-5              *****

M1-5 25 1  5  5  PMOSE L={L1} W={W1}

M2-5 25 1  0  0  NMOSE L={L1} W={W1}

M3-5 45 1  35 0  NMOSE L={L1} W={W1}

M4-5 35 5  0  0  NMOSE L={L2} W={W2}

M5-5 45 0  25 5  PMOSE L={L2} W={W2}

Q1-5 5  25 45   NPN

Q2-5 45 35 0    NPN

CL-5 45 0        {CLOAD}


*****           BCD-4              *****

M1-4 24 1  5  5  PMOSE L={L1} W={W1}

M2-4 24 1  0  0  NMOSE L={L1} W={W1}

M3-4 44 1  34 0  NMOSE L={L1} W={W1}

M4-4 34 24 0  0  NMOSE L={L2} W={W2}
```

Q1-4 5 24 44   NPN

Q2-4 44 34 0   NPN

CL-4 44 0   {CLOAD}

*****   MODEL   *****

.MODEL NPN NPN BF=100 BR=1 VAF=10 RB=100 RC=50 RE=20

+   IS=1E-16 NF=1 NR=1 NE=1.5 NC=2 EG=1.11

+   IKF=3E-3 VJE=0.75 MJE=0.33 VJC=0.75 MJC=0.33

+   CJE=40F CJC=33F CJS=0.05P TF=40P TR=800P

.MODEL NMOSE NMOS LEVEL=3 VTO={VTEN} TOX=1.75E-8 TPG=1

+   NSUB=3.8E16 XJ=.15U UO=550 THETA=0.05

.MODEL PMOSE PMOS LEVEL=3 VTO={VTEP} TOX=1.75E-8 TPG=1

+   NSUB=4.5E16 XJ=.15U UO=200 THETA=0.05

.MODEL PMOSD PMOS LEVEL=3 VTO={VTD} TOX=1.75E-8 TPG=1

+   NSUB=2E17 XJ=.15U UO=200 THETA=0.05

*****   ANALYSIS   *****

.DC VIN 0 5 0.025

.TRAN .1NS 40NS

.PROBE

.END

Table 2.1 Input BiCMOS data file for SPICE simulation

21

2.3 Noise margin for DC characteristics
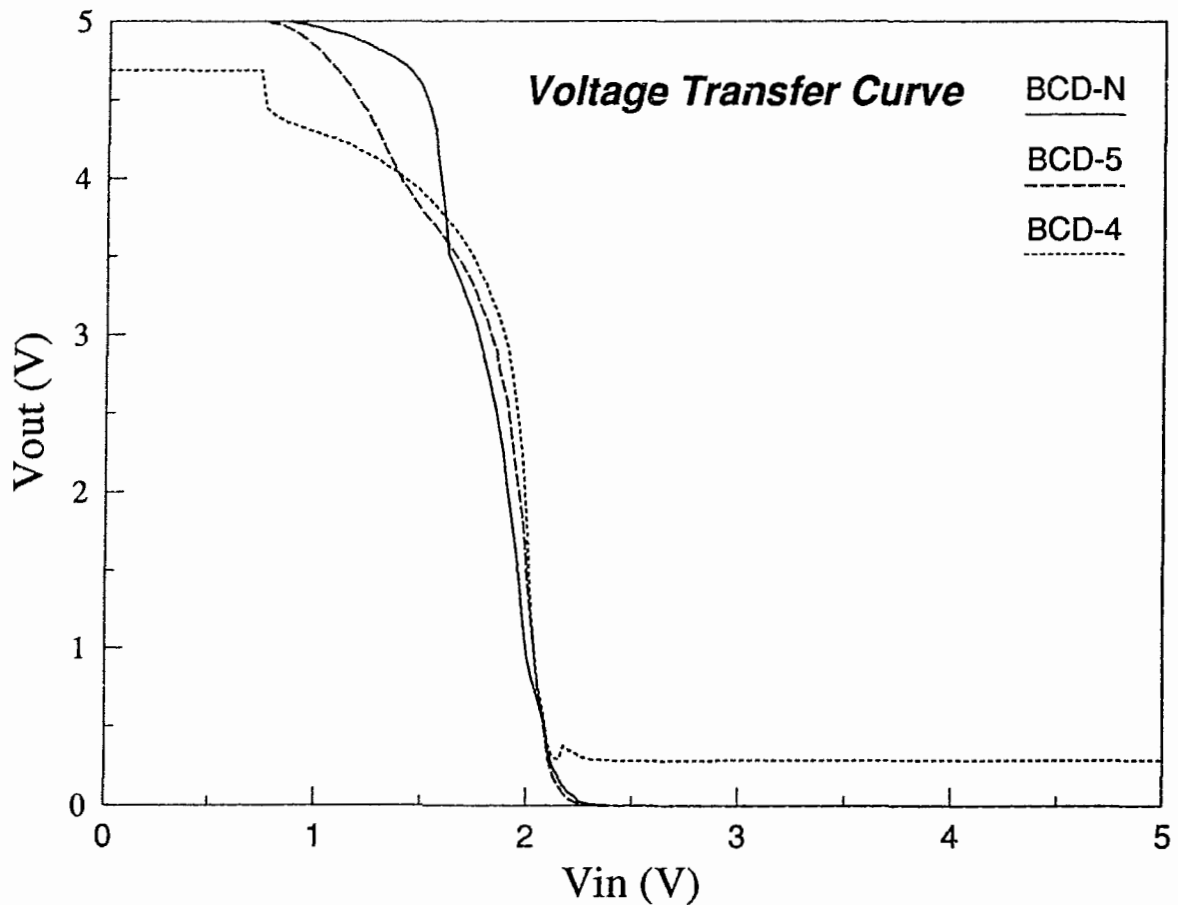
Figure 2.4 The DC voltage transfer curve for BCD-N, BCD-5 and BCD-4

Fig. 2.4 shows the simulated DC characteristics of the voltage transfer curves (VTC) for the above three BiCMOS drivers. The differences between these three drivers circuit are now described.

i) Both BCD-5 and BCD-N have full rail-to-rail static logic swing of $V_{OL} = V_{SS} = 0V$ to $V_{OH} = V_{DD} = 5V$, but BCD-4 has smaller logic swing of $V_{OL} = V_{BE0} = \sim 0.5V$ to $V_{OH} = V_{DD} - V_{BE0} = \sim 4.5V$ (the value of V_{BE0} is strongly related to the leakage current flowing through the BJT and is approximately 0.5V);

ii) BCD-N has the best low noise margin NM_L $(= V_{IL} - V_{OL} = 1.52V)$ of the three

22

circuits. Here V_{IL} is defined as the input voltage which corresponds to the output voltage of 90% of $(V_{OH}-V_{OL})$ in VTC. By comparison, the NM_L value is 1.18V for BCD-5 (with smaller V_{IL}) or 0.95V for BCD-4 (with larger V_{OL});

iii) BCD-N has almost the same high noise margin NM_H $(= V_{OH} - V_{IH} = 2.9V)$ as BCD-5. Here V_{IH} is defined as the input voltage which corresponds to the output voltage of $10\%(V_{OH}-V_{OL})$ in VTC. For BCD-4, its NM_H value is 2.5V (with lower V_{OH}).

The reason for this improvement of NM_L in BCD-N can be explained clearly with the help of Fig. 2.5 showing the simulation results to compare the related drain or base currents between BCD-N and BCD-5. When the input voltage V_{in} is in the range of 0.8-1.45V, due to its threshold voltage V_{TE} of 0.8V for enhancement mode NMOS device M3 and turn-on voltage V_{BE0} of ~0.65V for bipolar NPN transistor Q2, M3 is on and Q2 should be off. Whether there is a current flowing through M3 or not is now determined by the state of M4 since it is in parallel with Q2. In BCD-5, both M4 and M5 are designed as "on-resistor-behavior" devices, so as long as M3 is on $(V_{in} > 0.8V)$, there would be a certain amount of drain current I_{DS5} flowing through $M5 \rightarrow M3 \rightarrow M4$ as shown in Fig. 2.5. This causes an increase of the drain-source voltage V_{DS5}. In addition, this conductive path is connected in parallel with M2, so it would also increase the total drain current I_{DS1} $(= I_{DS2} + I_{DS5})$, or its associated drain-source voltage V_{DS1} for M1. That causes the output voltage V_{out} $(= V_{DD}-V_{DS1}-V_{DS5})$ to decrease rather steeply as the input voltage increases in this region. On the other hand, in BCD-N, the threshold voltage V_{TD} of depletion mode M4 is set to 1V in the simulation. The gate voltage $(= V_{out})$ and the source voltage $(= V_{BE2})$ for M4 are > 4V and < 1V respectively in this region. So M4 in BCD-N should be operating in the cutoff region due to its V_{GS4} being larger than V_{TD} + V_{BE2} (notice this is a depletion mode PMOS device). It then follows that no current is flowing through $M5 \rightarrow M3 \rightarrow M4$ due to its series connections, as shown in Fig. 2.5

with I_{DS5} = 0. Therefore the output voltage V_{out} in this region is obviously higher than that in BCD-5 because V_{DS5} is zero and also V_{DS1} is smaller, as compared with the corresponding ones in BCD-5. That is why the noise margin NM_L in BCD-N is better than that in BCD-5.
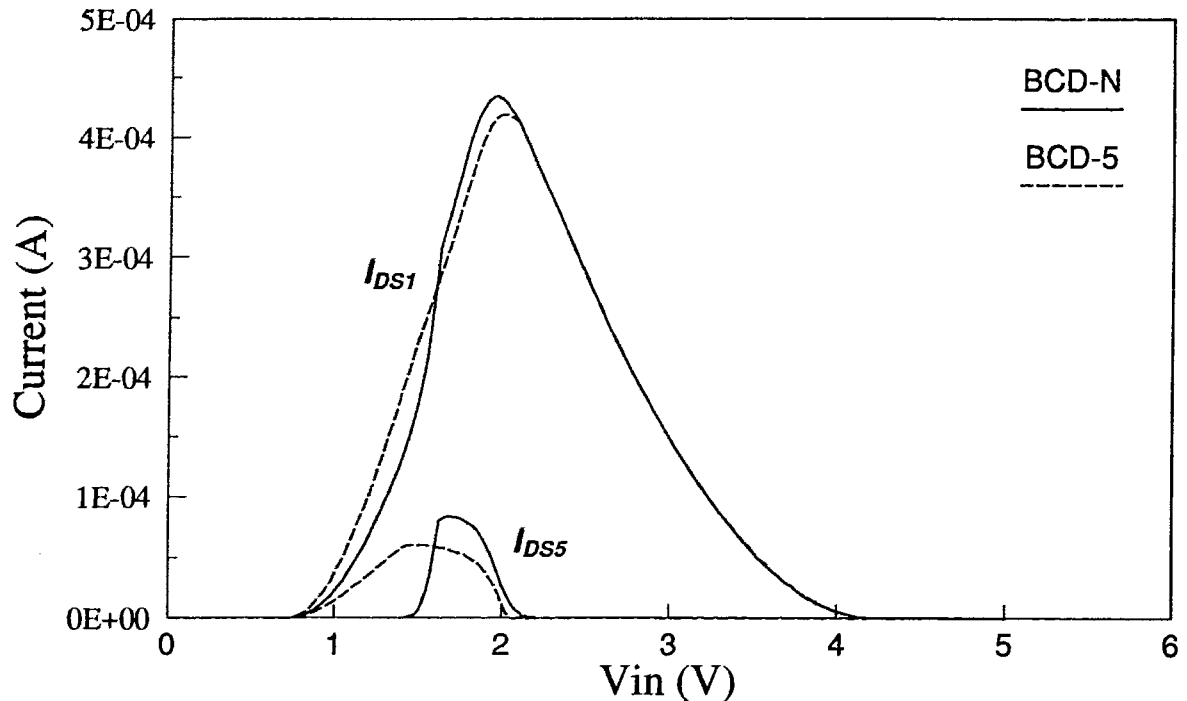
Figure 2.5 The drain current comparisons of I_{DS1} **and** I_{DS5}

between BCD-N and BCD-5

As to the noise margin NM_H, there is almost no difference between BCD-5 and BCD-N as shown in Fig. 2.4. This is mainly because both M4 and M5 in BCD-N are on when $V_{out} \approx V_{OL}$ and act similarly to the corresponding "on-resistor" of M4 and M5 in BCD-5.

The input voltage which corresponds to the $V_{out} = V_{in}$ in the middle of the voltage transfer curve, is slightly smaller (about 0.05V less) for BCD-N compared with BCD-5.

The reason can be explained briefly also from the Fig. 2.5. When the input voltage further increases into this region, the larger drain current of I_{DS5} in BCD-5 would turn on Q1 for smaller V_{in} and then diminish the drain current I_{DS5} dramatically as shown in Fig. 2.5. On the other hand, due to the off state of M4 in BCD-N, a similar process would take place for larger V_{in} after first turning on Q2, which results in increasing I_{DS5}, and turning on Q1 which results in diminishing I_{DS5} as also illustrated in Fig. 2.5. Therefore in this V_{in} region, I_{DS5} in BCD-5 is less than I_{DS5} in BCD-N as also shown in this figure. Since V_{DS5} is equal to V_{BE1} which is the same for both BCD-5 and BCD-N, the output voltage V_{out} (= V_{DD}-V_{DS1}-V_{DS5}) is a little lower for BCD-N than for BCD-5.

Based on the above discussion, it is clear to see that compared with BCD-5, this new BiCMOS driver keeps the same rail-to-rail static logic swing, but has better noise margin NM_L and almost the same noise margin NM_H.

2.4 Transient output response characteristics

Fig. 2.6a and 2.6b show the transient waveforms of both falling output response or rising output response for these three circuits with a step function input respectively. The transient output response for a cycle time of the pulse period in the real application case is also shown in Fig. 2.6c. The load capacitances C_L for these three circuit are fixed at 10pF. The simulation results clearly show that

i) BCD-N has the faster falling output response of $\tau_{d,f}$(50%) = 1.41ns and approximately the same rising output response of $\tau_{d,r}$(50%) = 1.79ns, as compared with BCD-5 of $\tau_{d,f}$(50%) = 1.55ns and $\tau_{d,r}$(50%) = 1.76ns respectively. Here $\tau_{d,f}$(50%) and $\tau_{d,r}$(50%) are defined as the delay time corresponding to 50% output voltage swing for the falling output response and the rising output response respectively;

ii) due to its lower static logic swing of BCD-4, it shows the fastest falling output

response of $\tau_{d,f}(50\%)$ = 1.31ns and also the fastest rising output response of $\tau_{d,r}(50\%)$ = 1.65ns in Fig. 2.6a and 2.6b. But as far as the slope of the transient output response is concerned, BCD-4 has the same slope for falling output response with BCD-N, and a little steeper slope for rising output response, when compared to BCD-5 and BCD-N;

iii) the transient logic swing for both BCD-N and BCD-5 are lower than its static logic swing (about V_{BE} to $V_{DD} - V_{BE}$) as shown in Fig. 2.4, and are approximately the same as BCD-4, except the transient logic "1" voltage for BCD-N is even lower (~4V);

iv) for the simulated transient output waveform after the first input pulse cycle (which is closer to the real application situation) as shown in Fig. 2.6c, BCD-N demonstrates the fastest falling output response and also the same rising output response compared to both BCD-5 and BCD-4. This is partly due to the contribution of the lower transient logic swing for BCD-N (which is also true for BCD-5).

The simple explanation for above transient output response comparison is now presented.
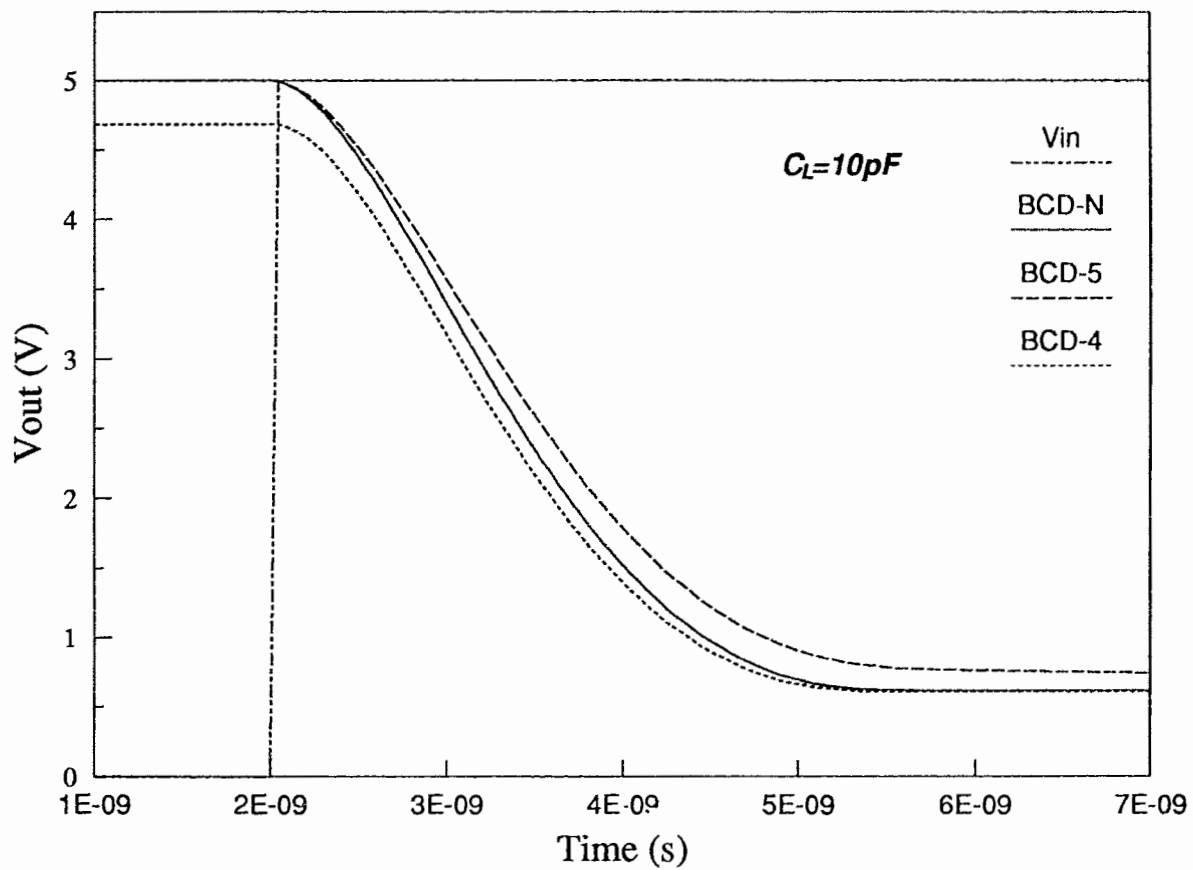
Figure 2.6a. Falling output response waveform comparisons
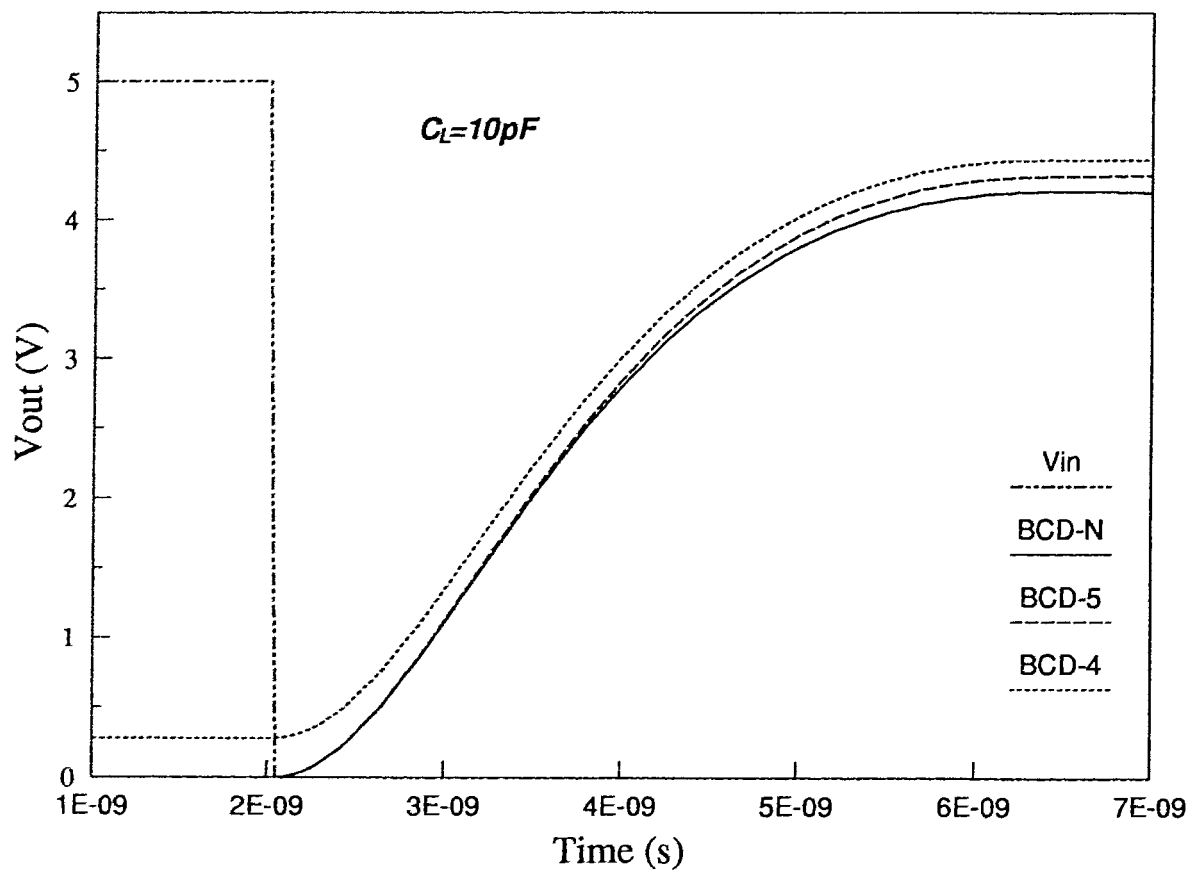
Figure 2.6b. Rising output response waveform comparisons
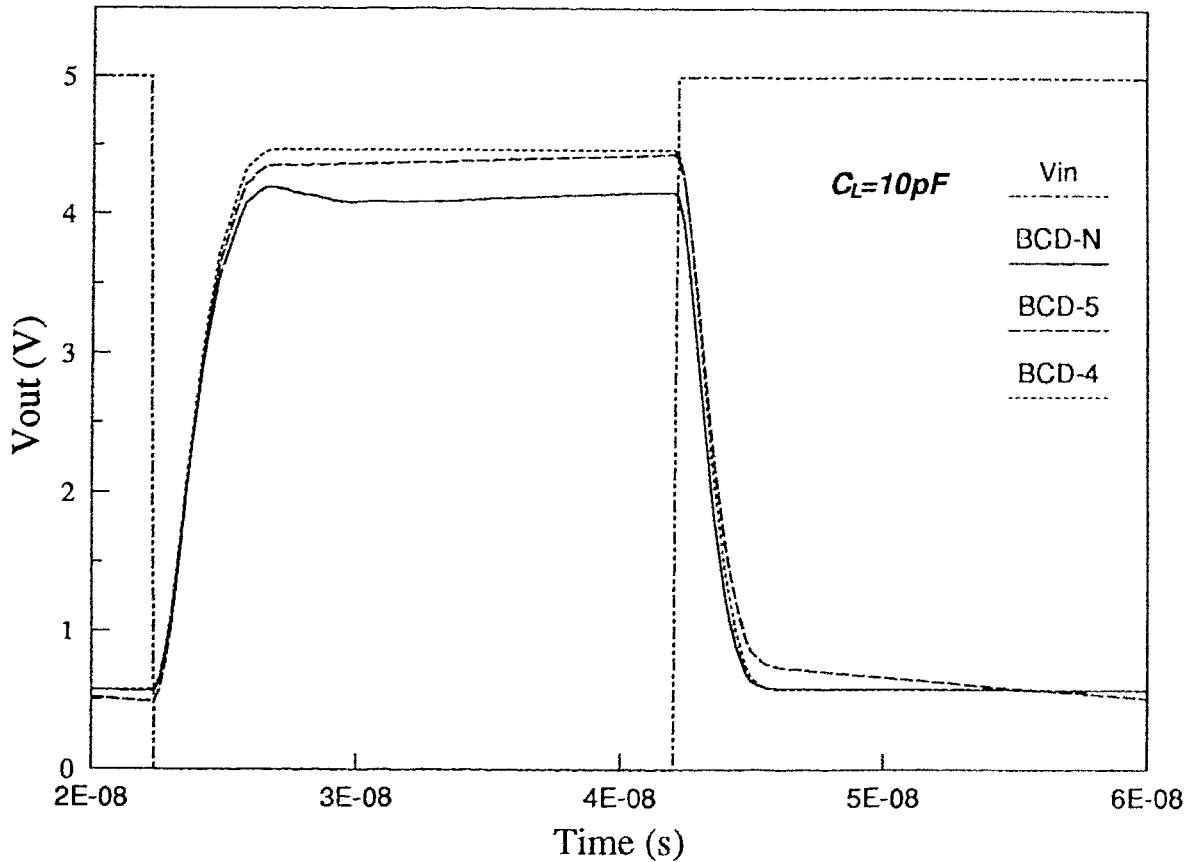
Figure 2.6c. The transient output response waveform comparisons

2.4.1 Falling output response comparison

As mentioned in section 2.2, the falling output response difference between BCD-5 and BCD-N is mainly due to the different type of MOS devices M4 are used. Fig. 2.7 further shows the simulated current response of I_{DS3}, I_{DS4} and I_{B2} (which are related to the falling output response process) for both BCD-N and BCD-5. It clearly shows the enhancement mode NMOS device M4 in BCD-5 is operating in the active region with an almost constant drain current $I_{DS4} = 0.365mA$ flowing through it during this time period. This results in a decrease in about 20% base current I_{B2} from the total drain current I_{DS3}. Meanwhile the depletion mode PMOS device M4 in BCD-N is operating in the cutoff region with no current taken by M4 in the same time period. So all the

drain current I_{DS3} flows as the base current I_{B2} for Q2. Obviously the discharge current of C_L through Q2 (= I_{C2}) for βCD-N should be larger. That is why the falling response rate of $\delta V_{out}(t) / \delta t$ for BCD-N is larger (absolute value) than that for BCD-5.
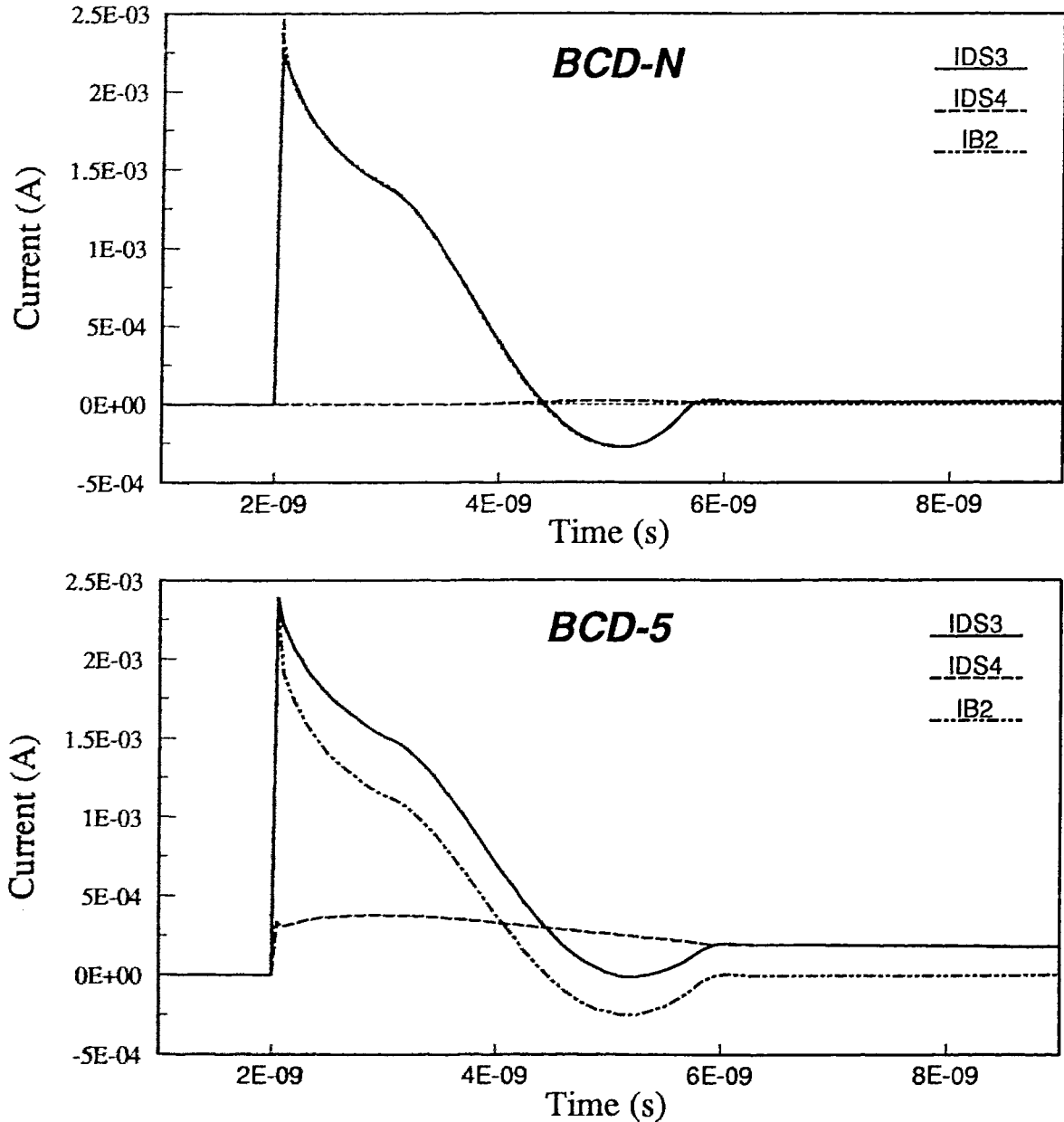
Figure 2.7 The comparisons of discharge current I_{DS3}, I_{DS4}, and I_{B2} between BCD-N and BCD-5

Theoretically, the falling output response for BCD-N and BCD-4 should be the same because M4 in BCD-4 is also off during this time period. This can also be verified by almost the same falling output response rate of $\delta V_{out}(t) / \delta t$ as shown in Fig. 2.6a.

2.4.2 Rising output response comparison

For the rising output response waveforms of $V_{out}(t)$ shown in Fig. 2.6b, there are some differences between these three circuits, especially for the higher output voltage part of the rising output waveform. That is BCD-4 is a little faster than BCD-5, and BCD-5 is also a little faster than BCD-N. The reason for these differences is that during the time period of rising output response, the depletion mode PMOS device M5 in BCD-N would take away the most current from I_{DS1}, as compared with less current taken from the enhancement mode PMOS device M5 in BCD-5 (with the same bias voltage and device size), and with no such current taken in BCD-4, (since there is no M5 device in BCD-4). But in reality, as shown in Fig. 2.8 with the simulated current response of I_{DS1}, I_{DS5} and I_{B1} (which are related to the rising output response process), the drain current I_{DS5} in both BCD-5 and BCD-N increases very slowly from zero to the maximum value of ~0.1mA (BCD-5) or ~0.135mA (BCD-N). The significant current sunk by M5 from drain current I_{DS1} only takes place at the time of t > ~4ns (as shown in Fig. 2.8), or the output voltage $V_{out}(t)$ goes higher than 50% of the maximum output voltage (as shown in Fig. 2.6b). The initial maximum charge current of I_{DS1} = ~1.35mA flows during the most leading edge of the rising response time. That is why using additional M5 in either BCD-5 or BCD-N does not affect this rising output response seriously at least for the leading edge, as it does with M4 for the falling output response just discussed above. It is also clear that by substituting the enhancement mode M5 in BCD-5 with the depletion mode M5 in BCD-N, the negative effect on this rising output response could be neglected because the maximum base current difference for this substitution is quite

small (about 0.035mA or < 3% of total I_{DS1} current), and only appear at the higher output response region of the rise-time period. This outcome in BCD-N actually can be further diminished if M5 is designed with smaller channel width-to-length ratio W/L.
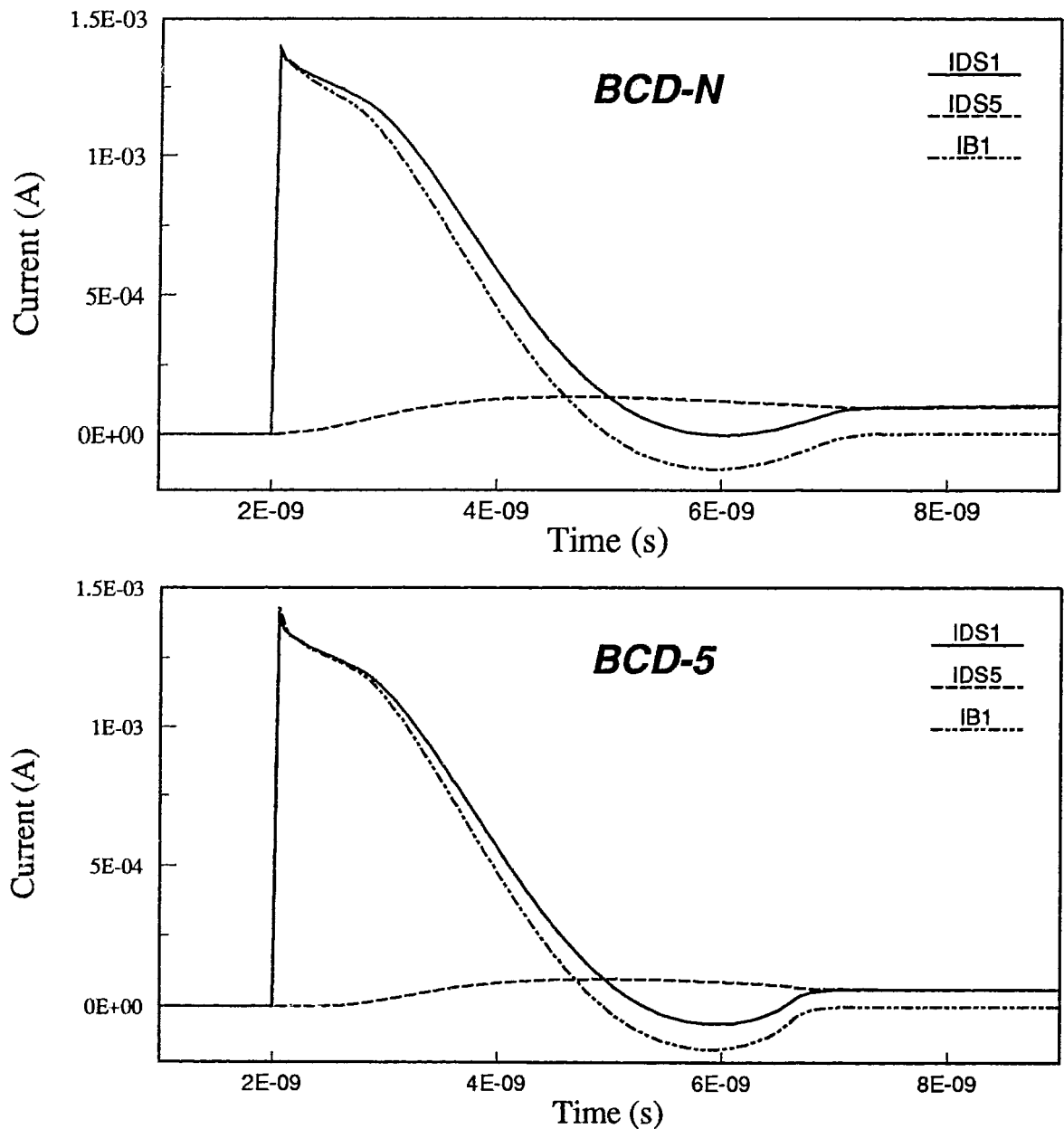
Figure 2.8 The comparisons of charge current I_{DS1}, I_{DS5}, and I_{B1} between BCD-N and BCD-5

2.4.3 Transient output logic swing comparison

The transient output logic swing in Fig. 2.6 shows BCD-N and BCD-5 do not have larger logic swing as compared with BCD-4. This result seems to contradict the static voltage transfer characteristics shown previously in Fig. 2.4, where both BCD-N and BCD-5 have the full V_{SS} to V_{DD} static logic swing.

The reason for this difference is the transient output response for BCD-5 and BCD-N are divided into two sections. The first section is referred to charging (or discharging) load capacitor C_L process through Q1 (or Q2) until its base-source voltage is reduced to approximately the turn-on value of $V_{BE} \approx 0.6V$. During this process, because Q1 (or Q2) is operating in the active region with large current flowing, its associated charging (or discharging) time constant is quite small (that is why BiCMOS offers fast speed). But the transient logic swing which could be reached in this process is obviously lain between $V_{OL} = V_{BE2}$ to $V_{OH} = (V_{DD} - V_{BE1})$, about the same as that for BCD-4 as shown in Fig. 2.6. When this first section process is almost ended, the drain-source voltage V_{DS1} for M1 (or V_{DS3} for M3) are about zero, and Q1 (or Q2) is now operating in the cutoff region. Then the second section following should be the charging (or discharging) C_L process through $M1 \rightarrow M5$ charge path (or $M3 \rightarrow M4$ discharge path). Clearly these charging or discharging currents are very small and result in a very large time constant in this section, as observed in Fig. 2.6. The simulation result showed that it would take more than 20mS to reach the final maximum logic swing of $(V_{DD} - V_{SS})$.

As to the lowest transient logic "1" output for BCD-N just mentioned at the beginning in this section, it is mainly due to the off state of M4 when $V_{out}(t)$ goes higher, and will be discussed in detail in chapter 5.

2.5 Compact structure with merged PMOS-NPN devices

(a)

(b)
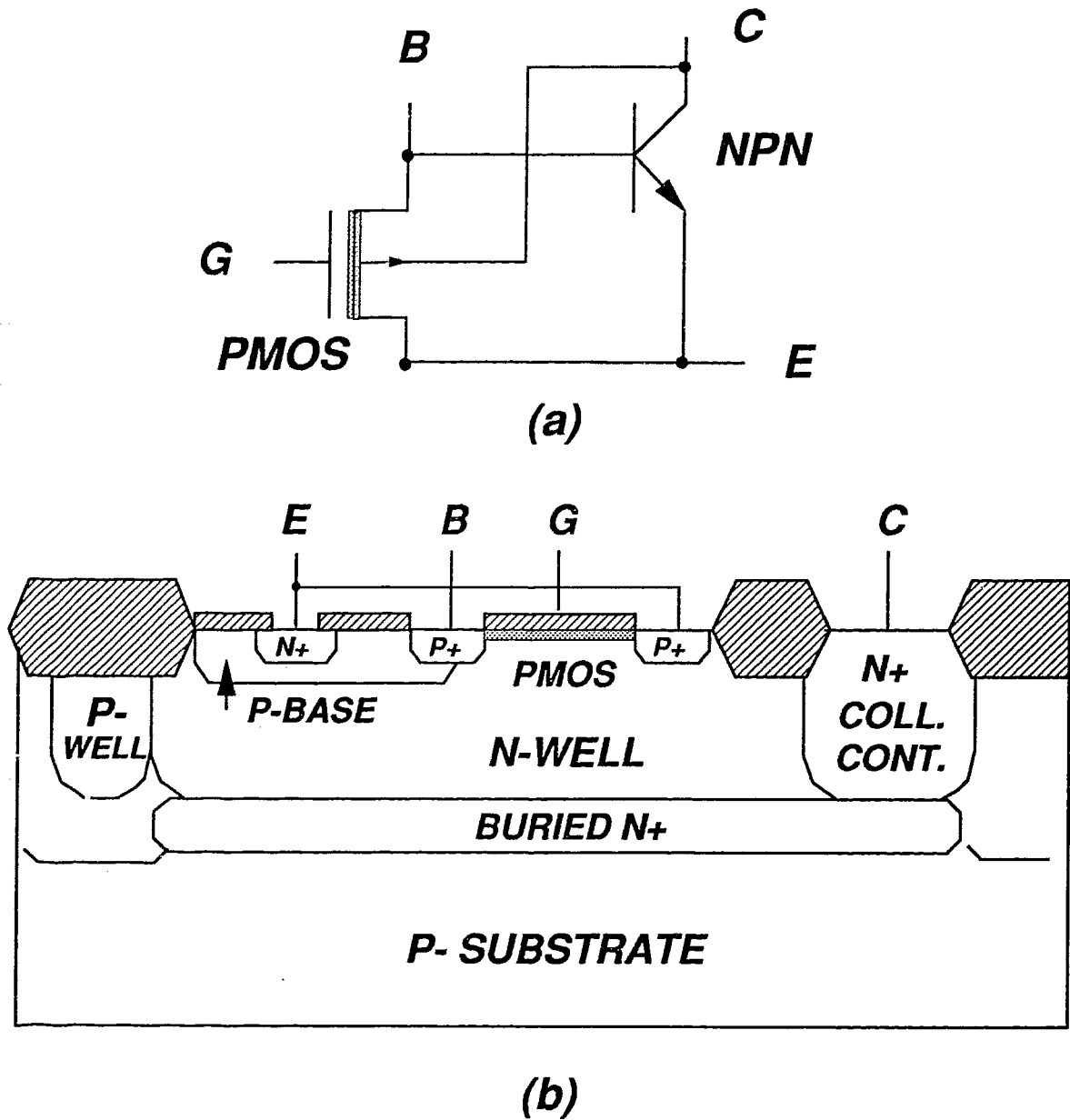
Figure 2.9 (a). Merged PMOS-NPN device diagram;
(b). Merged PMOS-NPN device structure

With the BCD-N circuit diagram shown in Fig. 2.3, it is easy to see that due to the replacement of the enhancement mode PMOS device M5 in BCD-5 with the depletion mode PMOS device M5 in BCD-N (which is the same as M4), both M4(PMOS)-Q2(NPN) and M5(PMOS)-Q1(NPN) pair devices in BCD-N have the same connection and can be merged together. Fig. 2.9a shows the schematic diagram for this merged PMOS-NPN devices which share the same P-type region for both drain (PMOS) and base (NPN), and the same N-type region for both substrate (PMOS) and collector (NPN). These two merged devices of M4-Q2 and M5-Q1 are entirely symmetrical. So the layout design for these merged devices could be quite compacted. Fig 2.9b shows one possible structure for the design of these merged PMOS/NPN devices.

2.6 Better tolerance for V_{TD} process control

From a processing point of view, one problem with this new BiCMOS driver is an additional requirement for fabricating the depletion mode PMOS device together with the enhancement mode PMOS and NMOS devices which are typically used in BiCMOS technology. However, our simulation results show that the threshold voltage V_{TD} for this depletion mode PMOS device is not a critical parameter for BCD-N circuit performance. Figs. 2.10 and 2.11 show both the DC and transient characteristics as a function of threshold voltage V_{TD} variation from 0 to 2V for this new BiCMOS driver. From these two figures, we make the following observations.
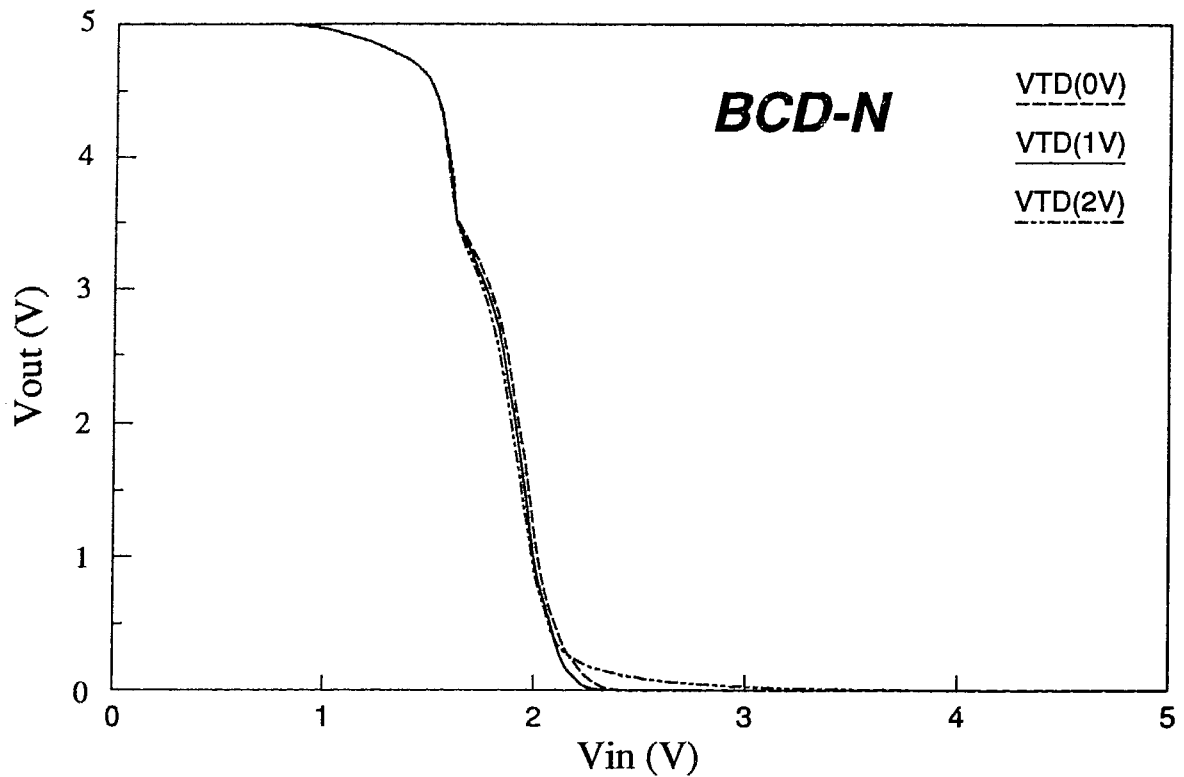
Figure 2.10 The DC voltage transfer curve for V_{TD} **varied from 0 to 2V**

A. For the DC characteristics, the V_{TD} variations of 0 to 2V have no effect on static logic swing of $V_{DD} - V_{SS}$, and also no effect on NM_L, but with slight decrease for NM_H. Since NM_H is usually larger than NM_L for BCD-N, then this decrease of NM_H does not seriously affect the overall DC characteristics;

B. For the transient output response, similarly this V_{TD} variations of 0 to 2V have almost no effect on the falling output response, and would increase a little bit for the rising output response. The reasons for these effects are:

i) for the fall-time period, because the gate voltage of M4 (same as the output voltage) is higher than or equal to the source voltage of M4 (same as the base voltage of Q2), this depletion mode device is in the cutoff region, so the V_{TD} variations of 0-2V for M4 should have no effect on this falling output response;

ii) for the rise-time period, because the drain current taken by M5 would increase as its threshold voltage V_{TD} increases, this results in the decrease of the base current I_{B1} accordingly. But as mentioned above, this drain current I_{DSS} itself does not significantly affect the rising output response, so its effect due to the V_{TD} variation of 0-2V could also be neglected as shown in Fig. 2.11.
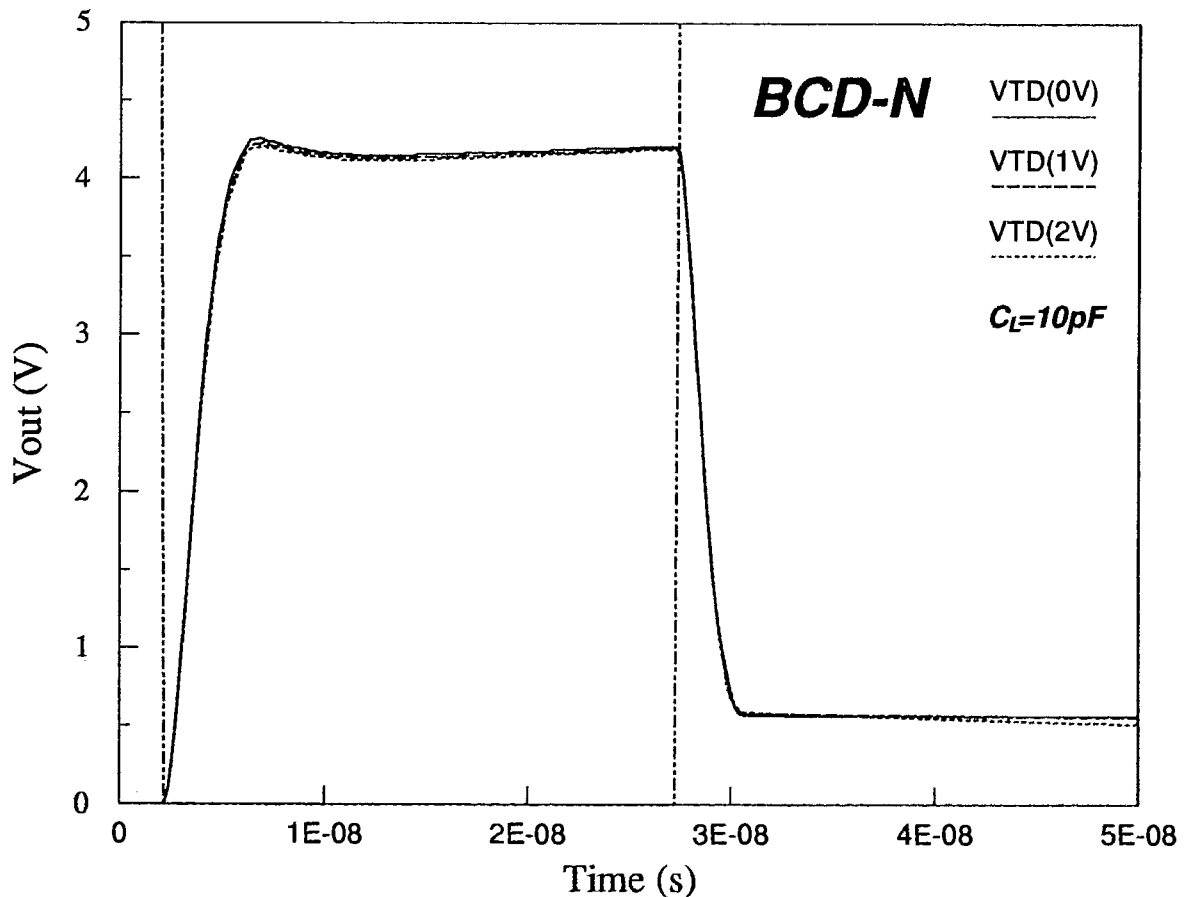
Figure 2.11 The transient output response waveform for V_{TD} varied from 0 to 2V

Therefore the requirement for the process control of the threshold voltage V_{TD} for this depletion mode PMOS device is not critical. It can be implemented either using additional ion-implant step, or even sharing with other possible step in the typical BiCMOS technology.

CHAPTER 3:

DETAILED DC ANALYSIS

3.1 Introduction

This chapter will derive the analytical expression of the DC characteristics for BCD-N and also discuss the effects of some important device parameters on this voltage transfer curve (VTC). It is well know that due to the highly nonlinear function of the large signal I-V characteristics for both NPN BJT and MOSFET, it is very difficult to derive the entire DC curves (as shown in Fig. 3.1 with solid line) analytically for this new BiCMOS driver circuit based on each individual device model. In order to implement the quantitative analysis, the piece-wise linear approximation model [3.1] is used here to analyze the voltage transfer curve (VTC) for BCD-N. In the following sections, simplified I-V models for both bipolar and MOSFET device in each operation region are listed first, which form the basis of the piecewise linear function approximation model. Based on the region of operation for each device in BCD-N and the associated equivalent circuit, the analytical expressions for the slope of the voltage transfer curve, denoted as S = δV_{out} / δV_{in} in this thesis, are derived individually for each segment in VTC. The calculated values S for each segment from this model are in very good agreement with the simulated results by SPICE. In the final section, some important effects on this DC curve with device parameter variations like threshold voltage V_{TD} for depletion mode MOSFET and V_{TE} for enhancement mode MOSFET are briefly discussed.
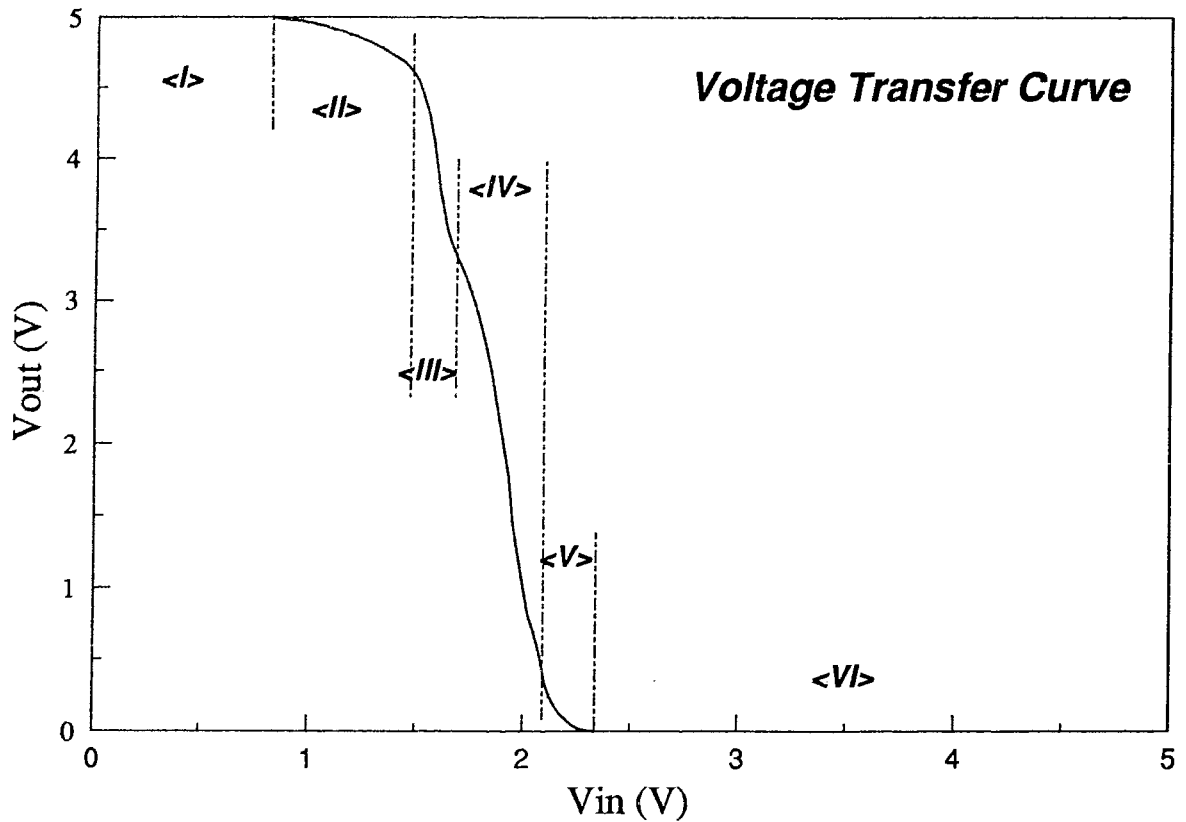
**Figure 3.1 The DC voltage transfer curve for BCD-N with six-segment
piecewise linear function approximation model**

3.2 Piece-wise linear approximation model

From a close examination of the voltage transfer curve (VTC) shown in Fig. 3.1, we
can approximate it by a six-segment piecewise linear function, as also illustrated in this
figure by the dash line. Therefore it is possible to use piece-wise linear function
approximation model to analyze this voltage transfer curve in terms of the linear I-V
approximation model for both BJT and MOSFET devices as described below.

3.2.1 Linear approximation for large signal I-V model of BJT

Taking NPN BJT as an example, linear approximation expressions of large signal I-V model for NPN BJT can be written for the following three regions of operation [3.2]:

(1). In the cutoff region:

When $V_{BE} < V_{BE0}$ (*turn − on voltage*), $\qquad\qquad\qquad\qquad$ (3.1)

$$I_B = I_C = 0. \qquad\qquad\qquad\qquad (3.2)$$

(2). In the forward active region:

When $V_{BE} \geq V_{BE0}$ and $V_{BC} < 0$, $\qquad\qquad\qquad\qquad$ (3.3)

$$I_C = \beta_F I_B \qquad\qquad\qquad\qquad (3.4)$$

with

$$\beta_F(I_C) \approx \beta_{F0}\left(1 + \frac{I_C}{I_{KF}}\right)^{-1}. \qquad\qquad\qquad\qquad (3.5)$$

The above β_F expression includes both low-level and high-level injection effects for the BJT. β_{F0} is its ideal maximum forward current gain. I_{KF} is the Knee current which represents the corner for β_F high-current roll-off.

(3). In the saturation region:

When $V_{BE} \geq V_{BE0}$ and $V_{BC} \geq 0V$, $\qquad\qquad\qquad\qquad$ (3.6)

$$V_{CE(sat)} = V_{CE0(sat)} + I_C R_C \quad (\textit{with } V_{CE0(sat)} \approx 0.1V) \qquad\qquad (3.7)$$

3.2.2 Linear approximation for large-signal I-V model of MOSFET

Taking enhancement mode NMOSFET as an example, linear approximation expressions of large-signal I-V model for NMOSFET can be written for the following three regions [3.3]:

(1). In the cutoff region:

When $V_{GS} < V_{TE}$ (threshold voltage), $\qquad\qquad$ (3.8)

$I_{DS} = 0.$ $\qquad\qquad$ (3.9)

(2). In the linear (or ohmic) region:

When $V_{GS} > V_{TE}$ and $V_{DS} < (V_{GS} - V_{TE})$, $\qquad\qquad$ (3.10)

$I_{DS} = k(V_{GS} - V_{TE} - 0.5V_{DS})V_{DS}$ $\qquad\qquad$ (3.11)

with

$k = \mu_n C_{OX} W/L.$ $\qquad\qquad$ (3.12)

Here, k is the transconductance parameter, μ_n is the channel electron mobility, C_{OX} is the gate capacitance per unit area, W and L are the channel width and length respectively.

The equivalent resistance $R_{eq}(lin)$ for NMOSFET operating in this region with $V_{DS} \to 0$ is approximately equal to

$$R_{eq}(lin) = \frac{\delta V_{DS}}{\delta I_{DS}} \approx \{k(V_{GS} - V_{TE})\}^{-1}. \qquad (3.13)$$

(3). In the saturation region:

When $V_{GS} > V_{TE}$ and $V_{DS} \geq (V_{GS} - V_{TE})$, $\qquad\qquad$ (3.14)

$I_{DS}(sat) = I_{DSS} = 0.5k(V_{GS} - V_{TE})^2.$ $\qquad\qquad$ (3.15)

The transconductance $g_m(sat)$ for NMOSFET operating in this region can be

expressed as

$$g_m(sat) \ = \ \frac{\delta I_{DS}}{\delta V_{GS}} \ = \ k(V_{GS} - V_{TE}).$$ (3.16)

Notice that this $g_m(sat)$ expression in the saturation region is happened to be the reciprocal expression of $R_{eq}(lin)$ in the linear region.

The dynamic output resistance $R_{eq}(sat)$ in the saturation region is theoretically infinity (for a large MOSFET), as implied in equation (3.15). However, due to channel length modulation effect, this $R_{eq}(sat)$ value is reduced somewhat and can be described by the parameter λ in typical MOSFET model as

$$I_{DS}(sat) \ = \ I_{DSS}(1 + \lambda V_{DS})$$ (3.17)

$$R_{eq}(sat) \ = \ \frac{\delta V_{DS}}{\delta I_{DS}} \ = \ (\lambda I_{DSS})^{-1}.$$ (3.18)

The difference between the NMOSFET and the PMOSFET for above models is distinguished by different sign for all the voltages and currents used in above models, that is, with plus signs for NMOSFET and minus signs for PMOSFET. The voltage values in equations (3.8), (3.10) and (3.14) should be taken as the absolute values in the PMOSFET case. The difference between the enhancement mode and the depletion mode PMOSFETs in the above models is simply to replace V_{TE} by V_{TD}.

3.3 Analytical expressions for different operating regions

Before deriving the analytical expressions for the slope of VTC, the regions of operation for all devices in BCD-N should be clarified first. Fig. 3.2 - 3.8 show the simulated currents of I_{DS} or I_B, and the simulated voltages of V_{GS}, V_{DS}, or V_{BE} for all seven devices in BCD-N as a function of V_{in} (varied from 0 to 5V). The detailed analyses for this six-segment linear approximation regions are now given.
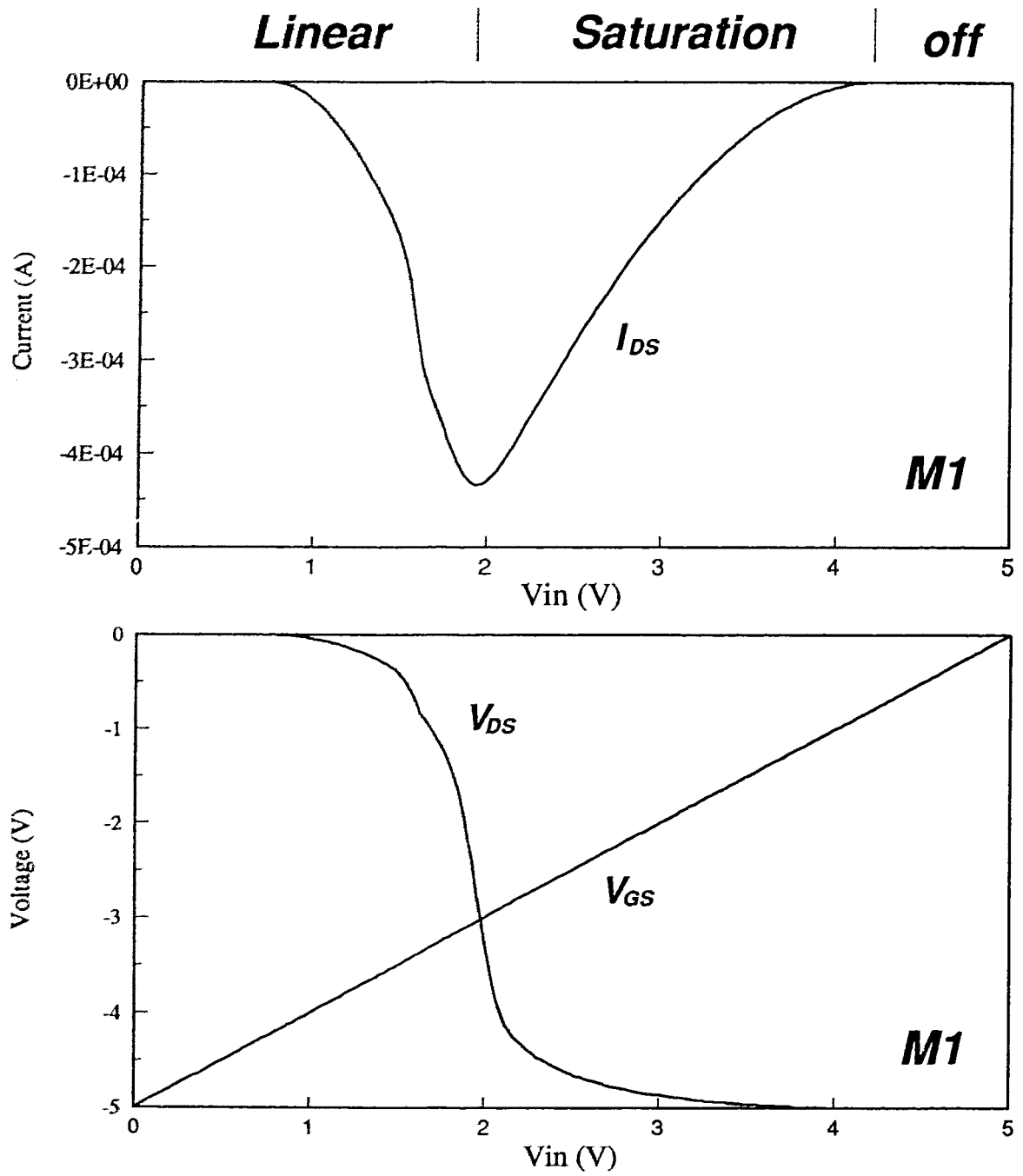
Figure 3.2 V_{GS}, V_{DS} and I_{DS} versus V_{in} for M1 device
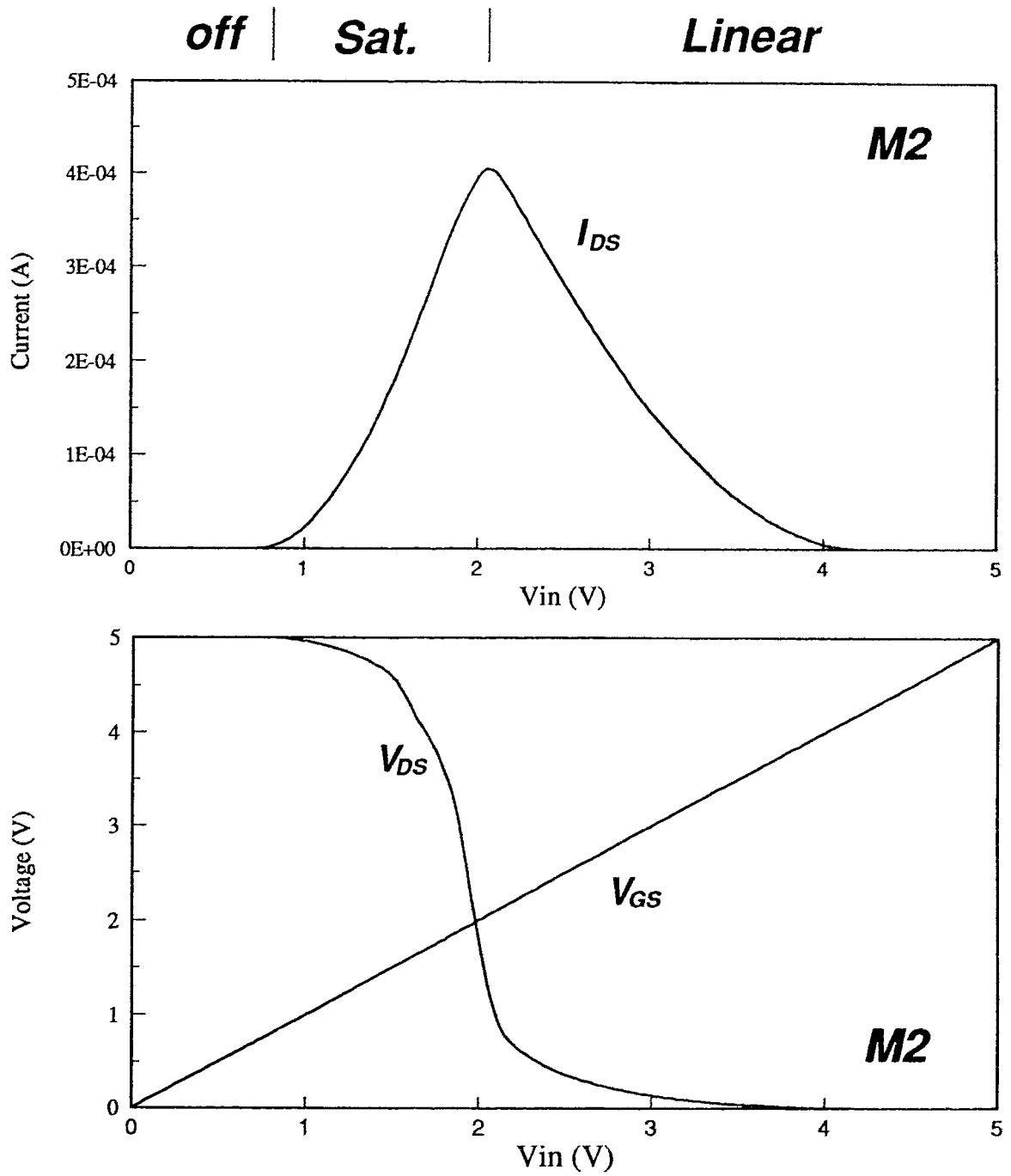
Figure 3.3 V_{GS}, V_{DS} and I_{DS} versus V_{in} for M2 device
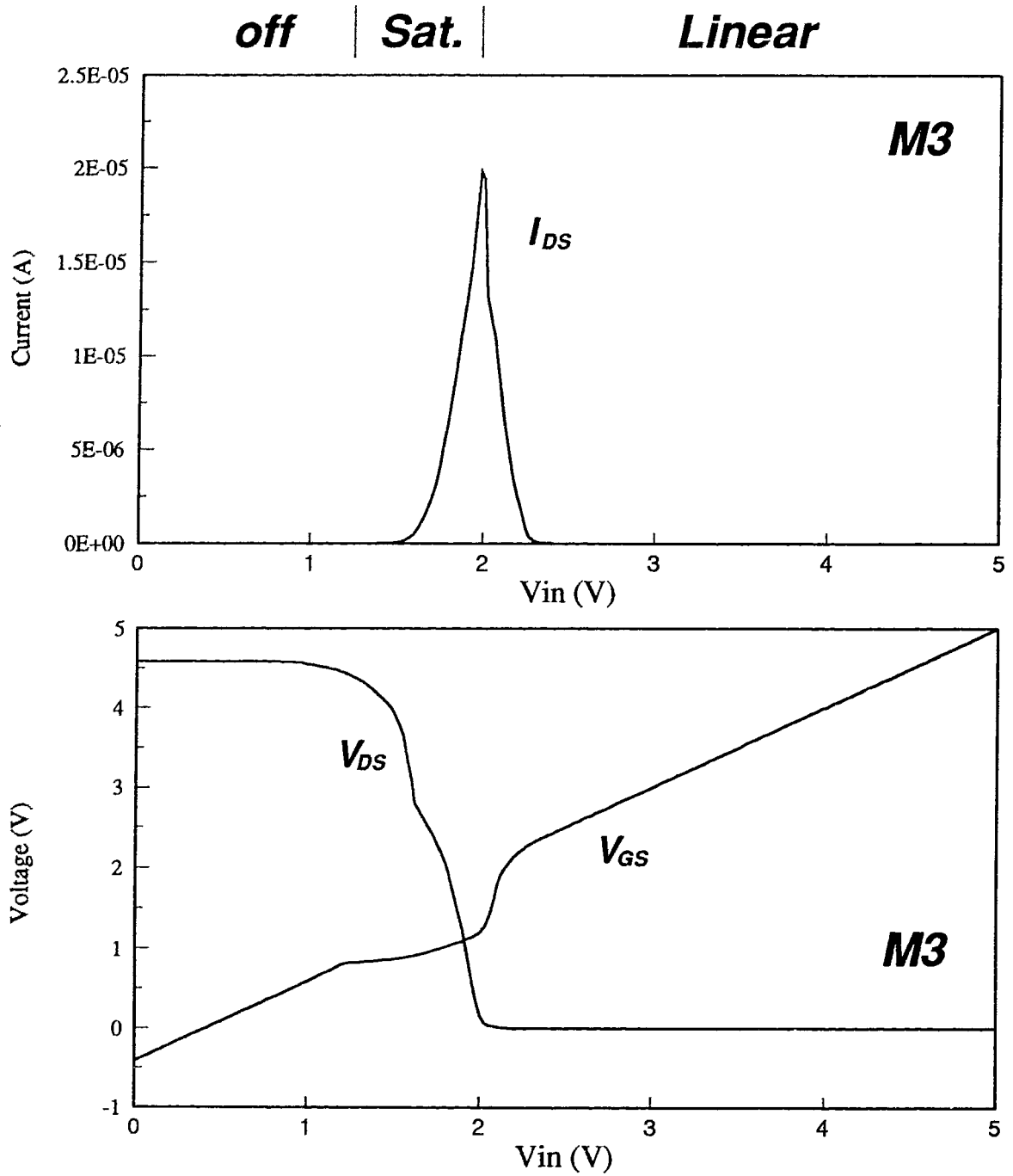
Figure 3.4 V_{GS}, V_{DS} and I_{DS} versus V_{in} for M3 device
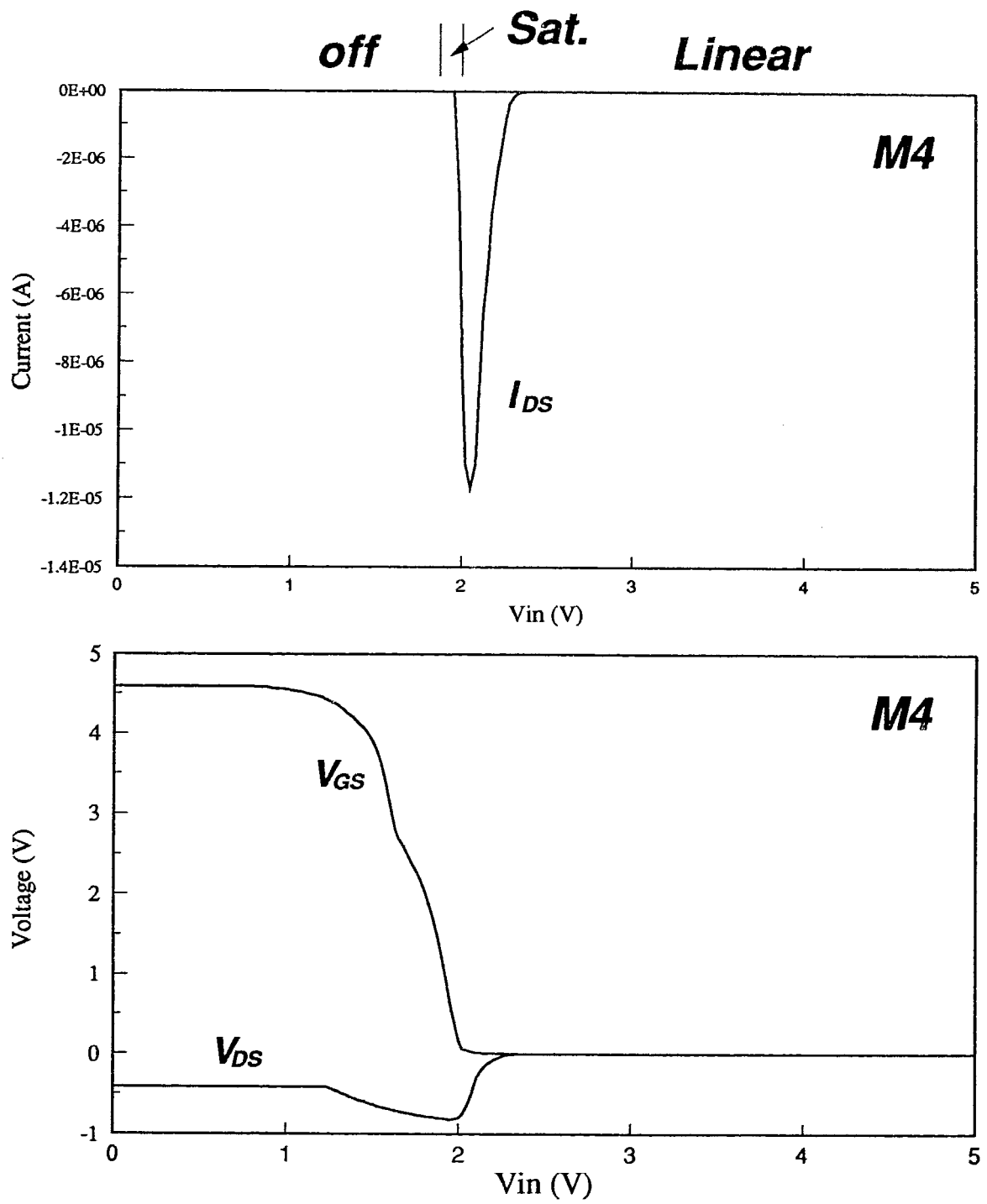
Figure 3.5 V_{GS}, V_{DS} and I_{DS} versus V_{in} for M4 device

47
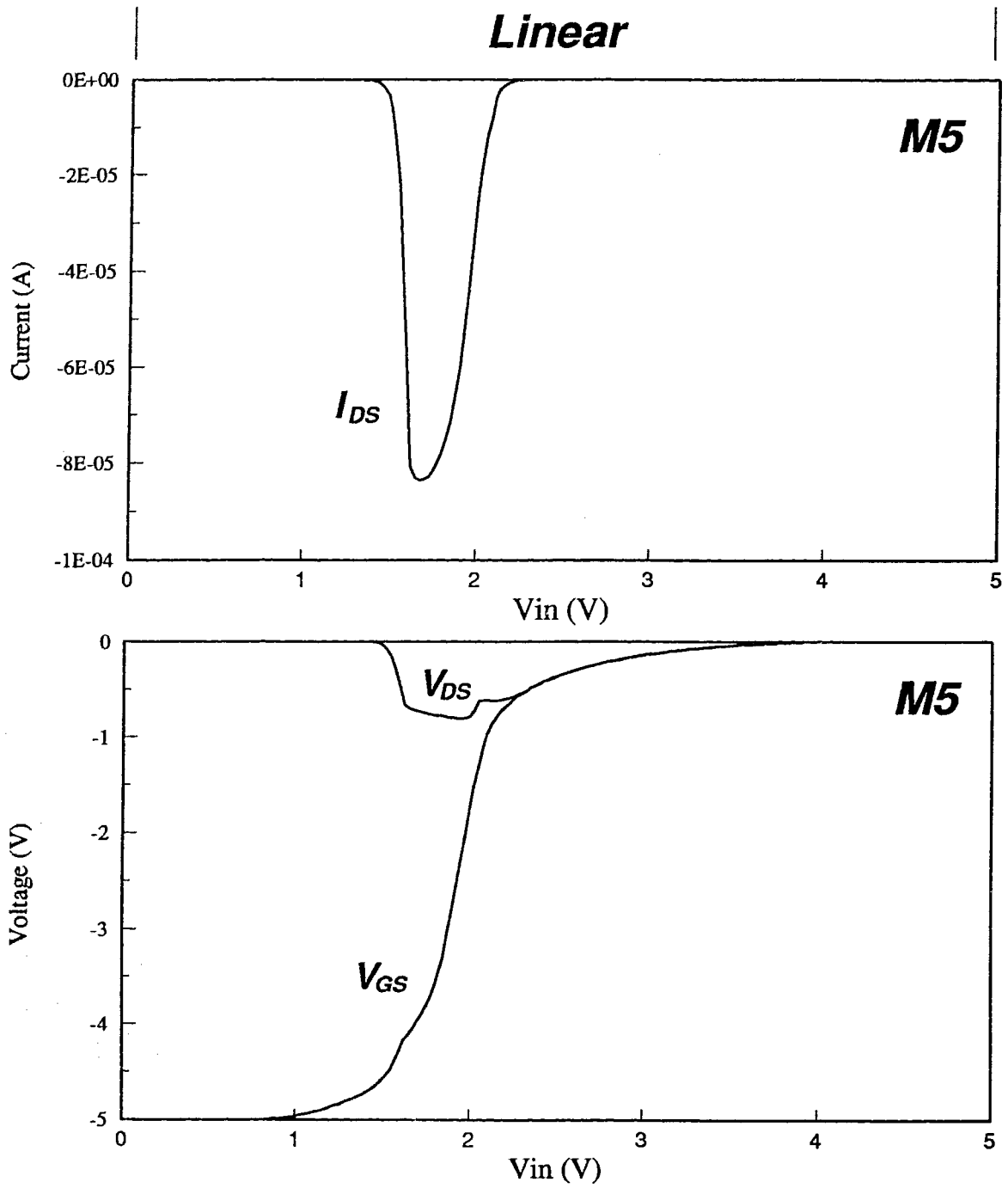
Figure 3.6 V_{GS}, V_{DS} and I_{DS} versus V_{in} for M5 device
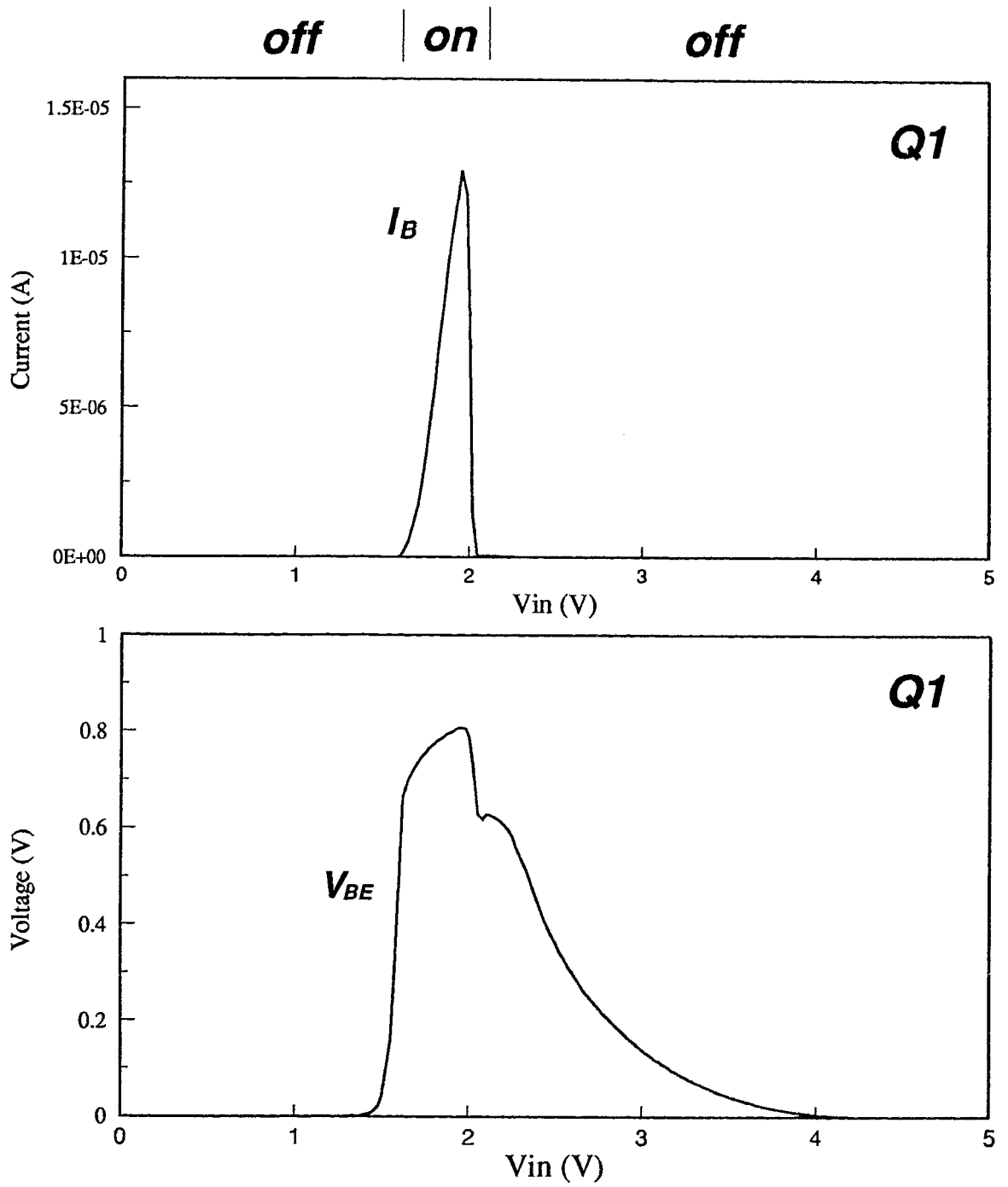
Figure 3.7 V_{BE} and I_B versus V_{in} for Q1 device

49
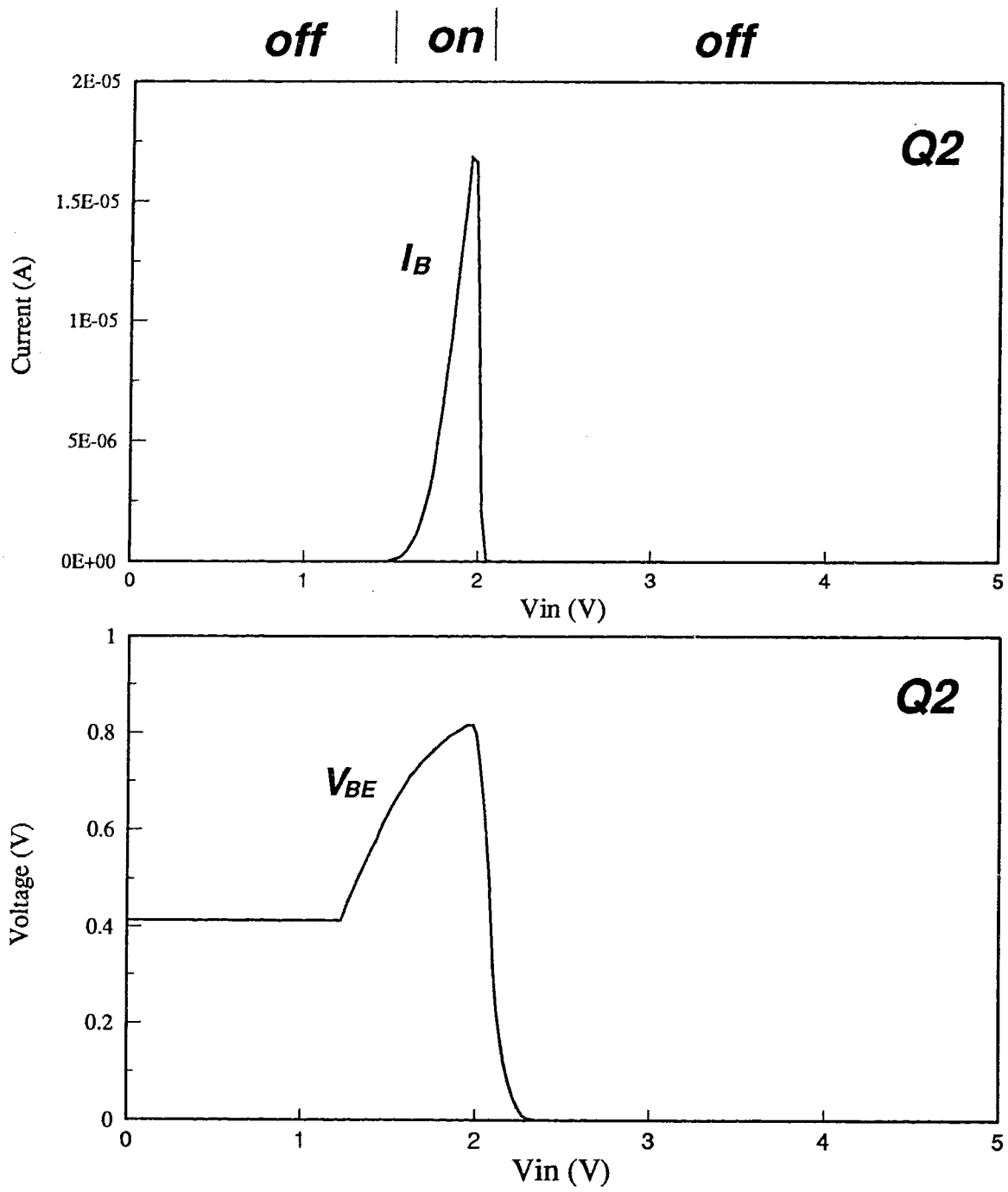
Figure 3.8 V_{BE} and I_B versus V_{in} for Q2 device

3.3.1 Region <I> ($0V < V_{in} < 0.8V$)

This region is referred to the input voltage V_{in} from 0 to 0.8V, which is equal to threshold voltage V_{TEN} of M2 (NMOSFET).
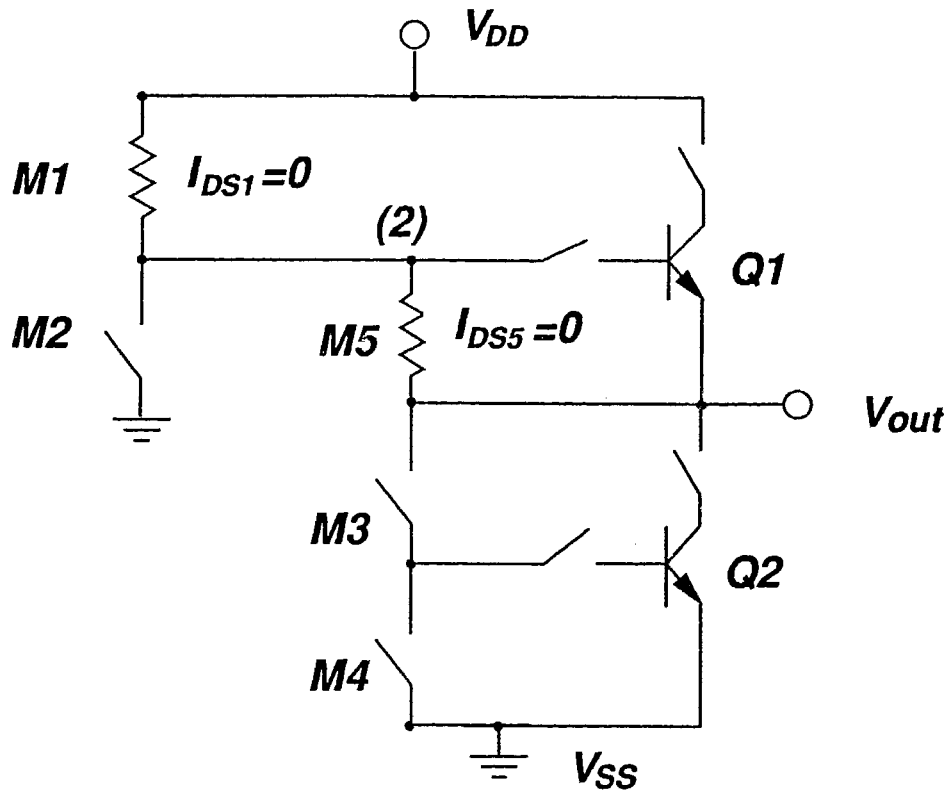
Figure 3.9 The equivalent circuit diagram of BCD-N in region <I>

In this region, it is easy to see that because V_{in} is less than V_{TEN}, M2, M3 and Q2 are off; M1, Q1, M5 are on, but with no DC current path or $I_{DS1} = I_{DS5} = 0$; Q1 is off due to $V_{BE1} = 0$; M4 is off because its V_{GS4} ($\approx V_{DD} - V_{BE2}$) is larger than V_{TD} (1V) for this depletion mode PMOSFET. Therefore the schematic equivalent circuit for BCD-N in this region is shown in Fig. 3.9. Since the output voltage V_{out} in this region <I> is equal to V_{DD}, a constant value, so its slope of $\delta V_{out} / \delta V_{in}$ is simply equal zero:

$$S_{(j)} = 0. \tag{3.19}$$

3.3.2 Region <II> ($0.8V < V_{in} < 1.45V$)

This region is referred to the input voltage V_{in} from 0.8V (V_{TEN}) to 1.45V, which is equal to threshold voltage V_{TEN} (0.8V) of M3 plus the turn-on base-emitter voltage V_{BE0} of Q2 (~0.65V).
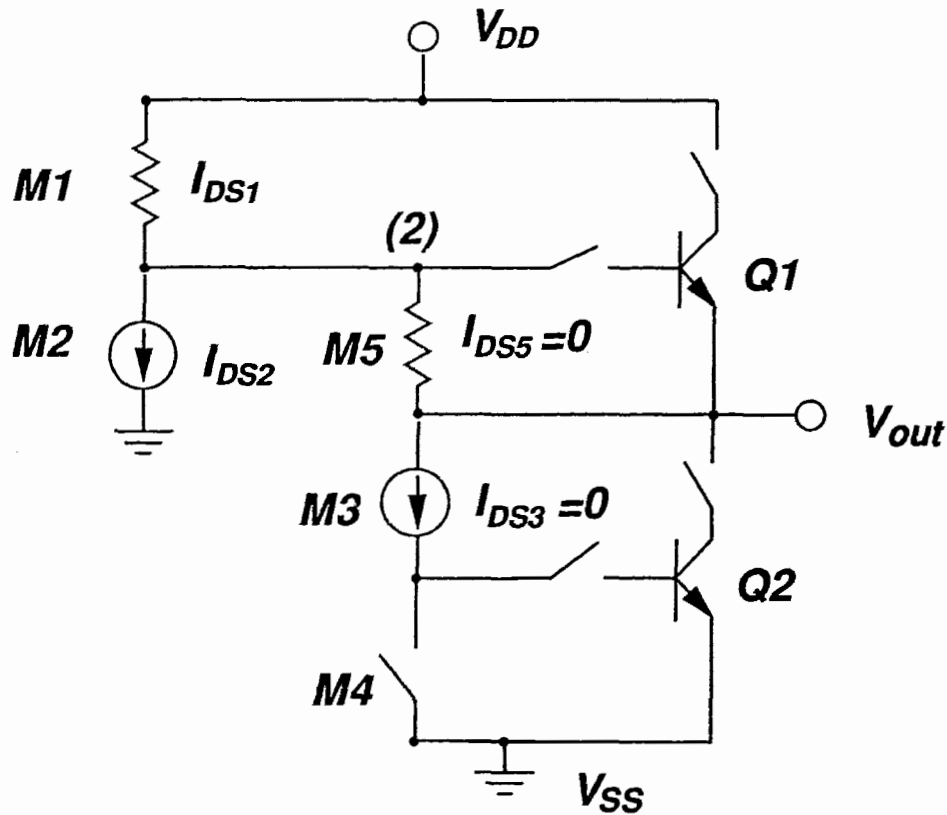
Figure 3.10 The equivalent circuit diagram of BCD-N in region <II>

First, in this region both M2 and M3 are turned on because of $V_{in} > 0.8V$. But it is also seen from Fig. 3.4 that there is no DC current flowing through M3. The reason for $I_{DS3} = 0$ here is because both M4 and Q2 are still off. For M4, similar to region <I> its V_{GS4} is still larger than V_{TD}, so M4 remains in the cutoff region as seen in Fig. 3.5. For

Q2, because its base-emitter junction is connected in serious with M3, the input voltage V_{in} is not lager enough to turn on both M3 and Q2 until it reaches to ~1.45V (V_{TEN} + V_{BE0}). Therefore the schematic equivalent circuit of BCD-N for this region can be drawn in Fig. 3.10. The physical meaning for this region can be expressed as: **from M2 turned on (at $V_{in} = 0.8V$) to M3 (or Q2) turned on (at $V_{in} = 1.45V$).** Since no drain current flows through M5, then the output voltage V_{out} is equal to the node voltage $V(2)$ which is the output voltage of CMOS inverter (M1-M2).

According to Figs. 3.2 and 3.3, in region <II>, M1 is operating in the linear region with V_{DS1} closer to zero, and M2 is operating in the saturation region due to its V_{DS2} (~5V) being larger than $V_{GS2} - V_{TEN}$ (<0.7V). Also these expressions of $\delta I_{DS1} = \delta I_{DS2}$, $\delta V_{out} = \delta V(2) = \delta V_{DS1}$, $\delta V_{in} = \delta V_{GS2}$ are valid from the equivalent circuit in Fig. 3.10. So the analytical expression for the slope of VTC can be expressed directly as:

$$S_{(I)} = \frac{\delta V_{out}}{\delta V_{in}} = \frac{\delta V_{DS1}}{\delta I_{DS1}} \frac{\delta I_{DS2}}{\delta V_{GS2}} = R_{eq1}(lin)g_{m2}(sat)$$

$$\approx \frac{k_2(V_{GS} - V_{TEN})}{k_1(V_{GS} - V_{TEP})} = \frac{k_2(V_{in} - V_{TEN})}{k_1(V_{in} - V_{DD} - V_{TEP})}. \tag{3.20}$$

Substituting the parameter values used in the simulation into the above equation, one can get the slope value of VTC at $V_{in} = 1.3V$ as:

$$S_{(I)} = \frac{550(1.3 - 0.8)}{200(1.3 - 5 + 0.85)} \approx 0.48. \tag{3.21}$$

The simulated value for this $S_{(I)}$ from Fig. 3.1 is ~0.485, which is in very good agreement with the value derived in above model.

53

3.3.3 Region <III> (1.45V < V_{in} < 1.65V)

This region is referred to the input voltage V_{in} from 1.45V (V_{TEN} + V_{BE0}) to 1.65V, which corresponds to Q1 being just turned on as shown in Fig. 3.7.
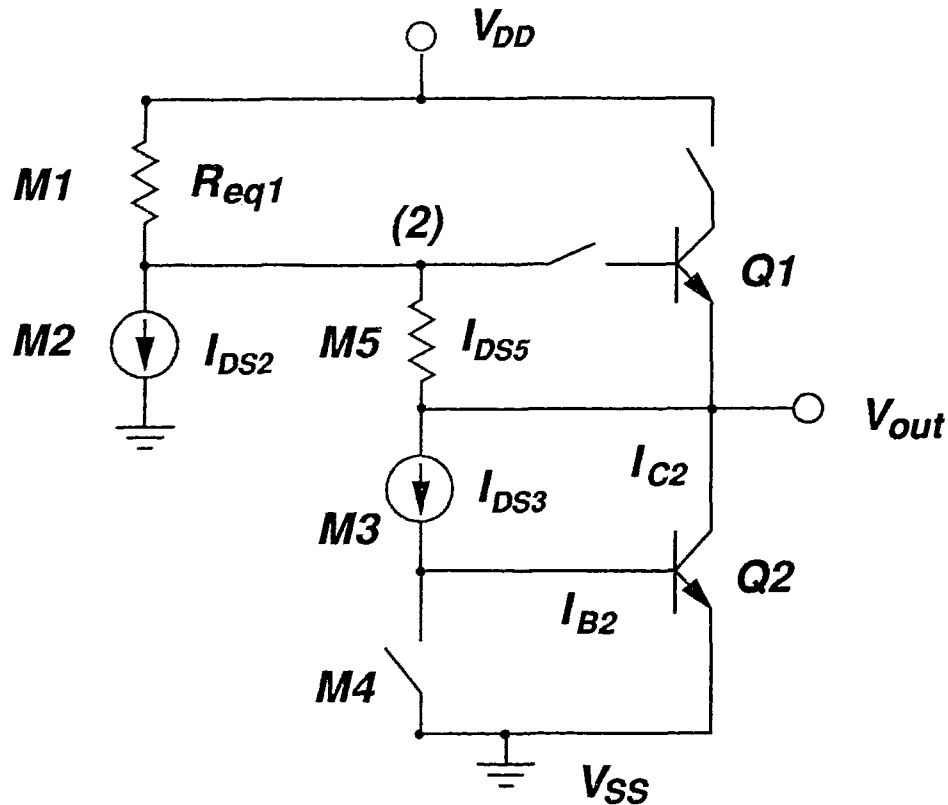
Figure 3.11 The equivalent circuit diagram of BCD-N in region <III>

In this region, because V_{in} is larger than 1.45V, both M3 and Q2 are now on. Also

M5 is a resistor-like device, so there is a DC current flowing through $M5 \rightarrow M3 \rightarrow Q2$. This conduction path is in parallel with M2. Since the current through Q2 (= I_{B2}) is exponentially proportional to the voltage across its base-emitter junction which is also related to V_{in}, it is easy to understand the slope of VTC in this region is much larger than that in region <II>. Fig. 3.11 shows the schematic equivalent circuit related to this

region. Clearly as V_{in} increases, the DC current flowing through M5 is also increased so that its drain-source voltage V_{DS5} would reach to V_{BE0}, and results in turning on Q1. Therefore the physical meaning for this region is: **from Q2 turned on to Q1 turned on.**

The analytical expression for the slope of VTC in this region can be derived based on the Fig. 3.11 with the following steps:

First, the output voltage variation can be expressed directly as

$$\delta V_{out} = \delta V_{DS1} + \delta V_{DS5}$$

$$= R_{eq1}(\delta I_{DS2} + \delta I_{DS5}) + R_{eq5}\delta I_{DS5}. \tag{3.22}$$

Second, since M1, M5 are in the linear region, and M2, M3 are in the saturation region, to the first order approximation one can get

$$R_{eq1}(lin) \approx g_{m1}^{-1}(sat),$$

$$R_{eq5}(lin) \approx g_{m5}^{-1}(sat),$$

$$\delta I_{DS2} \approx g_{m2}\delta V_{in},$$

$$\delta I_{DS5} = \delta I_{DS3} + \delta I_{C2} = \delta I_{DS3} + \beta_{F2}\delta I_{B2}$$

$$= (1 + \beta_{F2})\delta I_{DS3} \approx \beta_{F2}g_{m3}\delta V_{GS3}. \tag{3.23}$$

Because V_{GS3} is equal to $V_{in} - V_{BE2}$ (as shown in Fig. 2.3), and Q2 is now in weak conduction mode (with I_{B2} varied from ~50nA to ~1μA from Fig. 3.8). So its V_{BE2} value is related to V_{in} in a complicated way and can not be simply assumed as constant (say ~0.8V), as also shown in Fig. 3.8. Based on the simulation result in Fig. 3.4 or Fig. 3.8, one could assume the following empirical relationship between V_{GS3} and V_{in} for simplicity as:

$$\delta V_{GS3} = \delta(V_{in} - V_{BE2}) \approx 0.35\delta V_{in}. \tag{3.24}$$

Substituting above expression into equations (3.23) and (3.22), one can finally get the slope expression as:

$$S_{(II)} = \frac{\delta V_{out}}{\delta V_{in}} = \frac{g_{m2}}{g_{m1}} + 0.35\beta_{F2}(\frac{g_{m3}}{g_{m1}} + \frac{g_{m3}}{g_{m5}}).$$

(3.25)

To estimate this value around the $V_{in} = 1.6V$, the parameters used in simulation are substituted into above equation:

$$\frac{g_{m2}}{g_{m1}} = \frac{k_2(V_{GS2} - V_{TEN})}{k_1(V_{GS1} - V_{TEP})} = \frac{550(1.6 - 0.8)}{200(1.6 - 5 + 0.85)} \approx 0.86,$$

(3.26)

$$\frac{g_{m3}}{g_{m1}} = \frac{k_3(V_{GS3} - V_{TEN})}{k_1(V_{GS1} - V_{TEP})} = \frac{500(1.6 - 0.7 - 0.8)}{200(1.6 - 5 + 0.85)} \approx 0.108,$$

(3.27)

$$\frac{g_{m3}}{g_{m5}} = \frac{k_3(V_{GS3} - V_{TEN})}{k_5(V_{GS5} - V_{TD})} = \frac{550x5(1.6 - 0.7 - 0.8)}{200x0.875(0 - 4 - 1.0)} \approx 0.314.$$

(3.28)

So the slope value of VTC at $V_{in} = 1.6V$ in this region can be found as:

$$S_{(II)} = \frac{g_{m2}}{g_{m1}} + 0.35\beta_{F2}(\frac{g_{m3}}{g_{m1}} + \frac{g_{m3}}{g_{m5}})$$

$$= 0.86 + 0.35x100x(0.108 + 0.314) \approx 15.6.$$

(3.29)

The simulated value of $S_{(II)}$ from the Fig. 3.1 at $V_{in} = 1.6V$ is ~14.7, in fairly good agreement to the value of 15.6 given above.

3.3.4 Region <IV> (1.65V < V_{in} < 2.05V)

This region starts from $V_{in} = 1.65V$ to 2.05V, which is referred from the turn-on of Q1 to the turn-off of Q2.
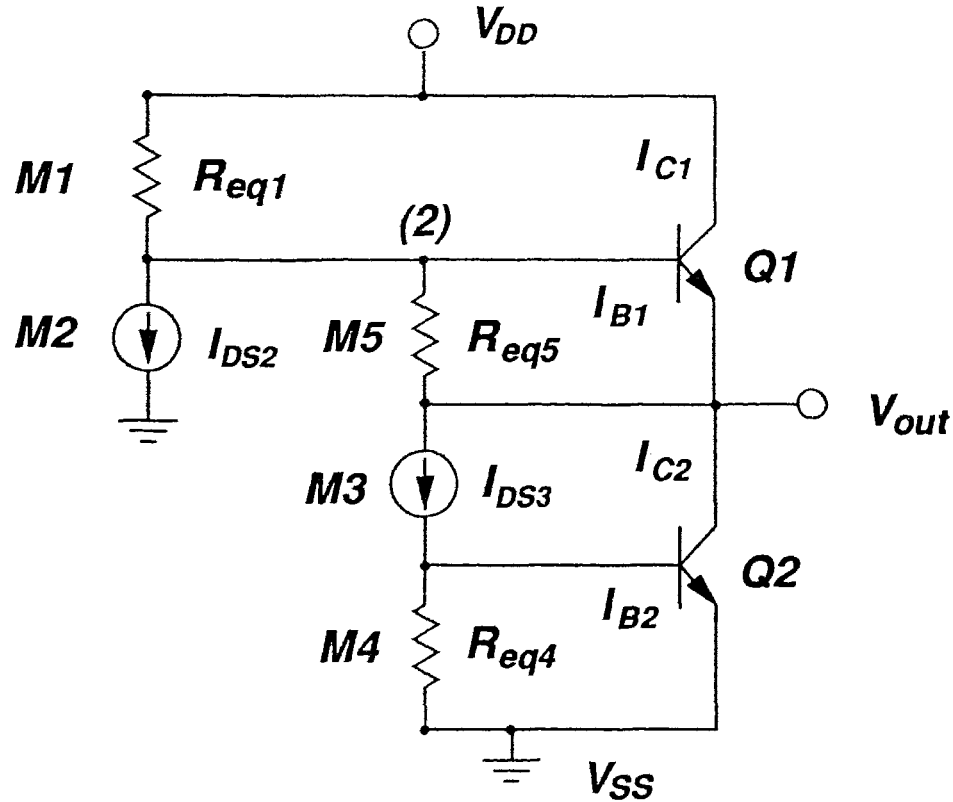
Figure 3.12 The equivalent circuit diagram of BCD-N in region <IV>

The output voltage V_{out} in this region decreases dramatically from ~3.5V down to

~0.8V. Since V_{in} is larger than 1.65V, so M1, M2, M3, Q1 and Q2 should be on as

already explained above. M5 is always on and behaviors as a resistor. But for M4, its

operation region is a little complicated as shown in Fig. 3.5. At the beginning of this

region, the higher output voltage V_{out}, which is connected to the gate of M4, causes M4

to still be in the cutoff region. When V_{out} goes below ~1.8V ($\approx V_{TD} + V_{BE2}$), the gate-source

voltage V_{GS4} ($= V_{out} - V_{BE2}$) is less than its threshold voltage value of V_{TD} (1V), so M4

begins to turn-on. The schematic equivalent circuit of BCD-N in this region is shown

in Fig. 3.12. As the output voltage V_{out} further decreases down to ~0.8V (V_{BE2}), the drain

voltage V_{DS3} and related I_{DS3} for M3 would goes towards zero, which means the base current I_{B2} is also about zero, or Q1 enters in the cutoff region. Therefore the physical meaning for this region is: **from Q1 turn-on to Q2 turn-off.**

This region actually can be further divided into two parts as shown in Fig. 3.2 and 3.3. The first part of V_{out} from ~3.5V down to ~2.0V (or V_{DS2} from ~4.3V to ~2.8V) is attributed to M1 in the linear region and M2 in the saturation region respectively. The second part of V_{out} from ~2.0V down to ~0.4V (or V_{DS2} from ~2.8V to ~1.2V) is referred to both M1 and M2 in the saturation region. As M1 is operating from the linear region to the saturation region, its dynamic output resistance increases significantly and would theoretically approach infinity. On the other hand, M2 is in the saturation region with larger drain current I_{DS2}, and also larger value of δI_{DS2} (~0.2mA) as shown in Fig. 3.3. Compared with this, the variation of the drain current I_{DS5} and the base current I_{B1} with δV_{in} are quite small (< 0.05mA) as shown in Figs. 3.6 and 3.7. This is mainly due to its almost fixed and small voltage values of $V_{DS5} = V_{BE1} = $ ~0.8V in this region. Therefore, to the first order approximation, it is reasonable to neglect both δI_{DS5} and δI_{B1} and use only δI_{DS2} to calculate the slope of VTC as

$$\delta V_{out} = R_{eq1}(\delta I_{DS2} + \delta I_{DS5} + \delta I_{B1})$$
$$\approx R_{eq1}\delta I_{DS2} = R_{eq1}g_{m2}\delta V_{in}. \tag{3.30}$$

Because M1 is in the saturation region, then from the equation (3.18), R_{eq1} can be expressed as

$$R_{eq1} = (\lambda I_{DSS1})^{-1} = \{0.5\lambda k_1(V_{in} - V_{DD} - V_{TEP})\}^{-1}, \tag{3.31}$$

so that

$$S_{(V)} = R_{eq1}g_{m2} = \frac{g_{m2}}{\lambda I_{DSS1}}$$

$$= \frac{k_2(V_{in} - V_{TEN})}{0.5\lambda k_1(V_{in} - V_{DD} - V_{TEP})}. \qquad (3.32)$$

The slope value at V_{in} = 1.9V can be estimated with the typical parameter value of λ = 0.05 as

$$S_{(V)} = \frac{k_2(V_{in} - V_{TEN})}{0.5\lambda k_1(V_{in} - V_{DD} - V_{TEP})}$$

$$= \frac{550x(1.9 - 0.8)}{0.5x0.05x200x(1.9 - 5 + 0.85)^2} \approx 23.9, \qquad (3.33)$$

which is fairly close to the simulation result of about 20 from Fig. 3.1.

As to the beginning part of this region, say from V_{out} = ~3.5V down to V_{out} = ~2.7V, because M2 is still in the linear region with relatively lower dynamic resistance than that in the saturation region, so the slope of VTC around this range is shown smaller as compared with both the previous region <III> or the following range of V_{out} below ~2.7V.

3.3.5 Region <V> (2.05V < V_{in} < 2.3V)

This region starts from V_{in} = ~2.05V, which refers to Q2 turned off, to V_{in} = ~2.3V, which refers to the drain voltage V_{DS4} of M4 and also the output voltage V_{out} goes down to zero.
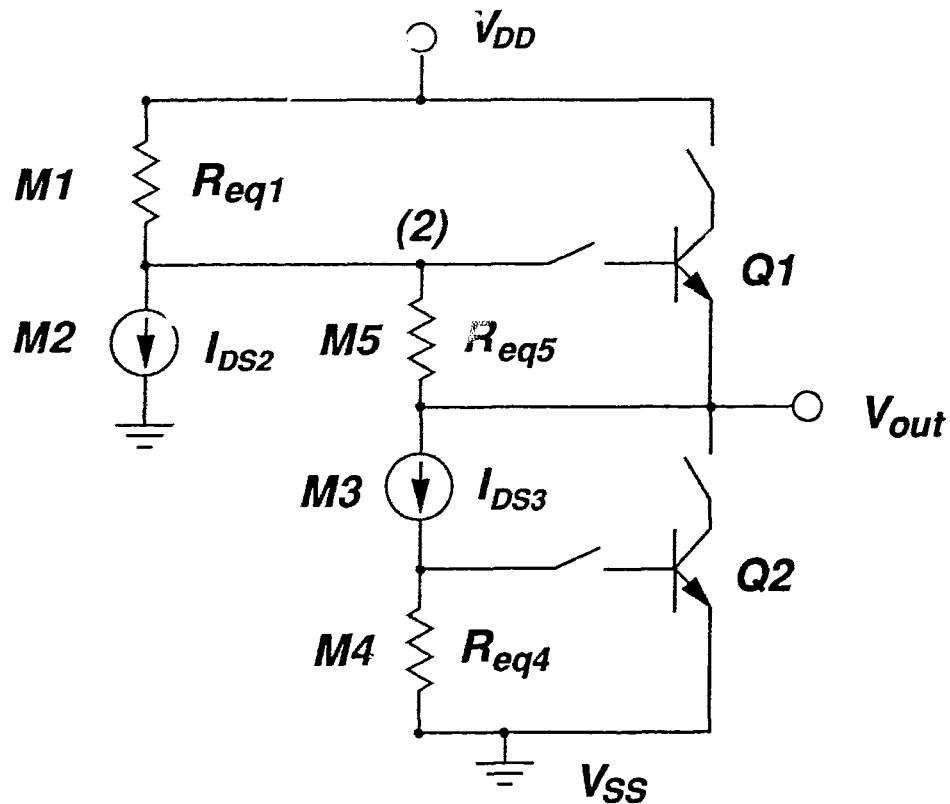
Figure 3.13 The equivalent circuit diagram of BCD-N in region <V>

In this region, Q2 is off due to its V_{BE2} being less than ~0.7V, and Q1 is also off because there is no DC current flowing through it. M1 and M2 are in the saturation and linear regions respectively. M5, M3 and M4 are all in the linear region and form a conduction path, which is in parallel with M2. But through this path there is only a few μA current flowing (see Fig. 3.4) because the drain-source voltage V_{DS3} is almost zero, as mentioned above. The schematic equivalent circuit for this region is shown in Fig. 3.13. This region would not be ended until the drain-source voltage V_{DS4} reduces to zero. So the physical meaning for this region is: **from Q1 turned off to $V_{DS4} = 0V$ of M4.** Knowing that the gate-source voltage V_{GS3} is larger than the threshold voltage V_{TEN}, its equivalent resistance is relatively small, as compared with the other two depletion mode

PMOS devices M4 and M5 and can be neglected. Also both PMOS devices of M4 and M5 have the same drain-source voltage of V_{BE0}, but the gate-source voltage V_{GS} is ~0V for M4, and ~-0.6V for M5, which means that the equivalent resistance of M4 is larger than that of M5. Therefore the variation of the output voltage δV_{out} would be mainly determined by the variation of the node voltage $\delta V(2)$ and can be expressed by:

$$\delta V_{out} = \frac{R_{eq4}}{R_{eq5} + R_{eq4}} \delta V(2) = \left\{ \frac{R_{eq4}}{R_{eq5}} + 1 \right\}^{-1} \delta V(2) \approx \delta V(2). \tag{3.34}$$

with

$$\delta V(2) = R_{eq2}(lin)\delta I_{DS1} = g_{m2}^{-1}(sat)g_{m1}(sat)\delta V_{in}. \tag{3.35}$$

The slope value of VTC at $V_{in} = 2.2$V can then be estimated as:

$$S_{(V)} = \frac{\delta V_{out}}{\delta V_{in}} \approx \frac{k_1(V_{in} - V_{DD} - V_{TEP})}{k_2(V_{in} - V_{TEN})}$$

$$= \frac{200x(2.2 - 0.8)}{550x(2.2 - 0.8)} \approx 0.5. \tag{3.36}$$

The simulated value of $S_{(V)}$ at $V_{in} = 2.2$V is about 0.33, in fair agreement with the above model.

3.3.6 Region <VI> (2.3V < V_{in} < 5V)

This region corresponds to the output voltage $V_{out} = 0$V, so the slope of VTC is equal to zero as:

$$S_{(VI)} = 0. \tag{3.37}$$
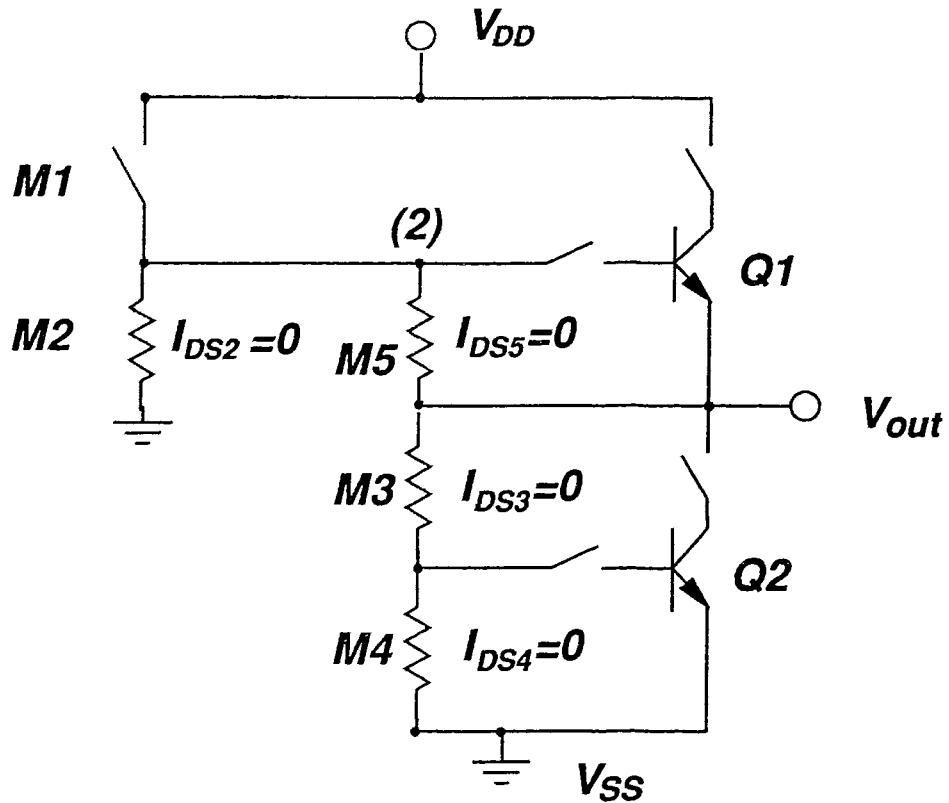
Figure 3.14 The equivalent circuit diagram of BCD-N in region <VI>

In this region, M1 is first in the saturation region, and then enter into the cutoff region when $V_{in} > 4.15V$ (or $|V_{GS1}| < |V_{TEP}|$). All other devices remain in the same operation region: M2 is in the linear region with $I_{DS2} \to 0$; M3, M4 and M5 are all in the linear region with no drain current; Q1 and Q2 are in the cutoff region. So the physical meaning for this region is: **from the $I_{DS4} = 0$ to M1 completely turned off.** Fig. 3.16 plots its schematic equivalent circuit for this region.

Finally Table 3.1 lists the summary of the regions of operation for all the devices in BCD-N corresponding to above six segments of VTC:

Region: V_{in}(V)	M1 (PMOS-E)	M2 (NMOS-E)	M3 (NMOS-E)	M4 (PMOS-D)	M5 (PMOS-D)	Q1 (NPN)	Q2 (NPN)
<I>: 0 -> 0.8	Linear	Cutoff	Cutoff	Cutoff	Linear	Off	Off
<II>: 0.8 -> 1.45	Linear	Saturation	Saturation	Cutoff	Linear	Off	Off
<III>: 1.45 -> 1.65	Linear	Saturation	Saturation	Cutoff	Linear	Off	On
<IV>: 1.65 -> 2.05	Lin -> Sat	Saturation	Saturation	Off -> Sat.	Linear	On	On
<V>: 2.05 -> 2.3	Saturation	Linear	Linear	Linear	Linear	Off	Off
<VI>: 2.3 -> 5	Sat -> Off	Linear	Linear	Linear	Linear	Off	Off

Table 3.1: The regions of operation for all the devices in BCD-N corresponding to six segments of VTC

3.4 Effects of some important parameter variations

3.4.1 The effect of V_{TD} variation on voltage transfer curve

It is very important to understand how the V_{TD} variation could affect the DC characteristics for this new BiCMOS driver circuit. In the previous section 2.6, the simulated voltage transfer curves for the threshold voltage variation of V_{TD} from 0 to 2V was demonstrated in Fig. 2.9. Based on the piecewise linear function approximation model developed in this chapter, the VTC variations as a function of V_{TD} shown in Fig. 2.9 can be explained physically as follows

(1) In region <IV>, the DC curves are shifted a little bit (about 0.02V) towards lower range of V_{in} as the V_{TD} varied from 0V to 2V. The reason for corresponding smaller V_{in} value when the V_{TD} value increases is mainly due to the lower equivalent resistance of M5. That is for the same voltage biases and device size, the larger positive V_{TD} value would reduce the equivalent resistor R_{eq} of the depletion mode PMOS device as implied in the model equation (3.13). This lower equivalent resistance results in a higher drain current flowing through $M1 \rightarrow M5 \rightarrow Q2$ (I_{C2}) as shown in Fig. 3.15, or the larger drain-source voltage V_{DS1} for the fixed input voltage V_{in}. Therefore the output voltage V_{out} is slightly reduced.

(2) In region <V>, the slopes of the VTC are changed with different V_{TD} values. This would enlarge the input voltage range for this region especially with a larger V_{TD} value (2V), as shown in Fig. 2.9. By looking at the equivalent circuit for this region <V> shown in Fig. 3.13, it is seen that the ratio of the equivalent resistance of M5 over M4 (both are PMOS depletion mode devices) would increase with increasing V_{TD} values. So

the output voltage variation δV_{out} becomes smaller as compared with the node voltage variation $\delta V(2)$ as given by the equation (3.34) above. That is why the slope of VTC in this region is also smaller with larger V_{TD} value.
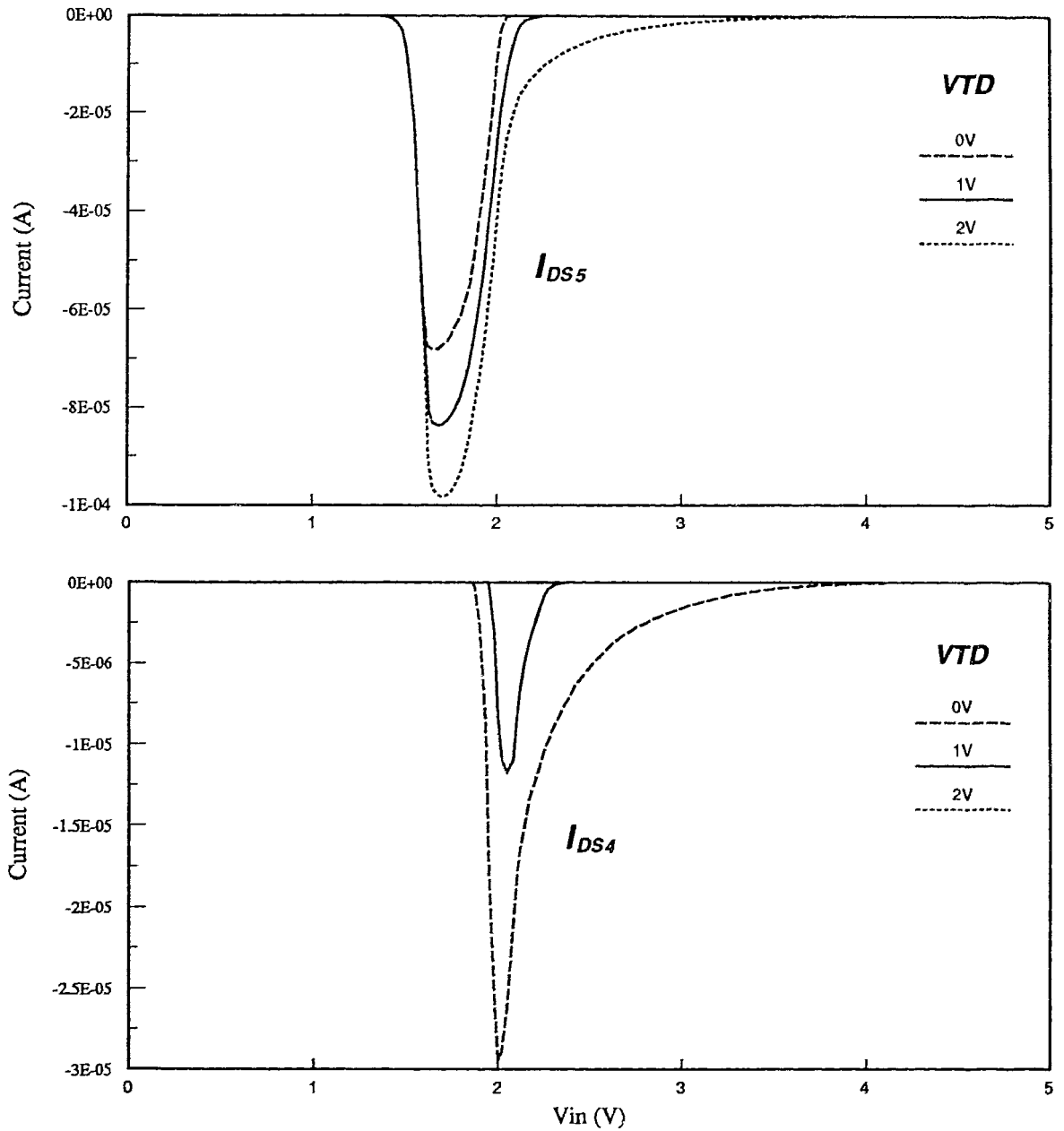
Figure 3.15   The drain current of I_{DS5} and I_{DS4} for V_{TD} varied from 0 to 2V

3.4.2 The effect of V_{TE} variation on voltage transfer curve

It is well know that this is the important parameter which would affect the DC characteristics for CMOS inverter, and also for the BiCMOS driver. Figs. 3.16 and 3.17 show the DC voltage transfer curves and the transient output response waveform with V_{TE} varying from 0.5 to 1.4V. It is clear to see that as the V_{TE} increases, the BiCMOS driver has better noise margin, especially for NM_L. On the other hand, the drain currents for both M1 and M2 become smaller with larger V_{TE} as described by the MOS device model equations (3.11) and (3.15). It would then cause slower transient output response for the BiCMOS circuit as shown in Fig. 3.17a for rising output response, and Fig. 3.17b for falling output response. So this V_{TE} value should be chosen with both DC and transient characteristics taken into consideration. The V_{TE} value of ~0.8V is the typical value widely used in the CMOS inverter and also the BiCMOS driver circuit.
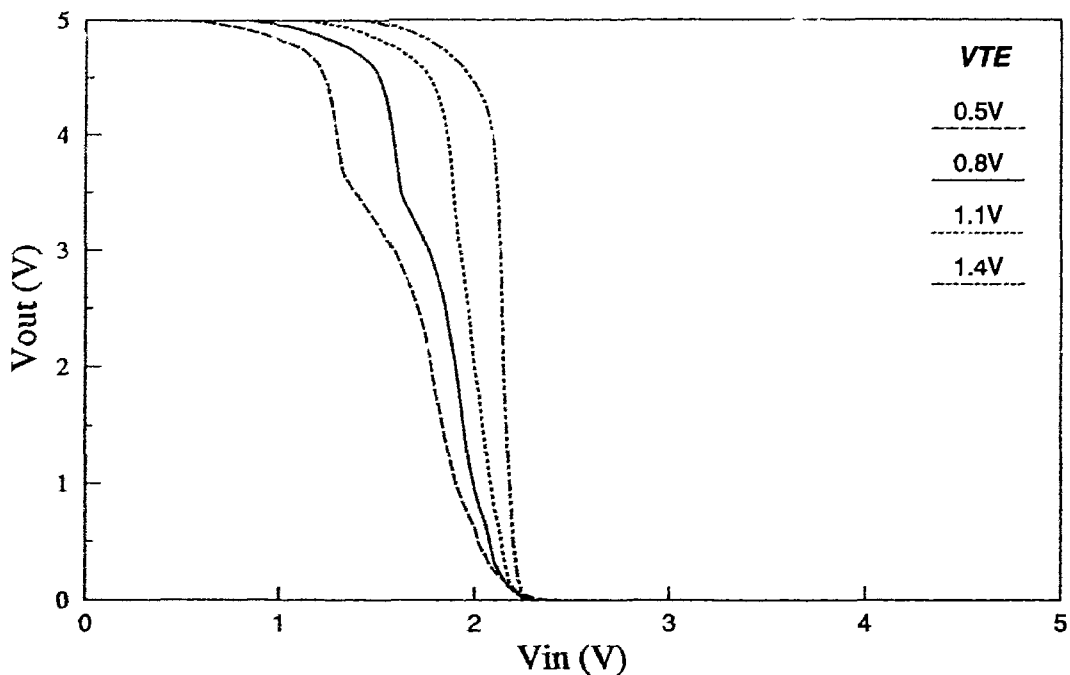
Figure 3.16 The DC voltage transfer curves for V_{TE} **varied from 0.5 to 1.4V**
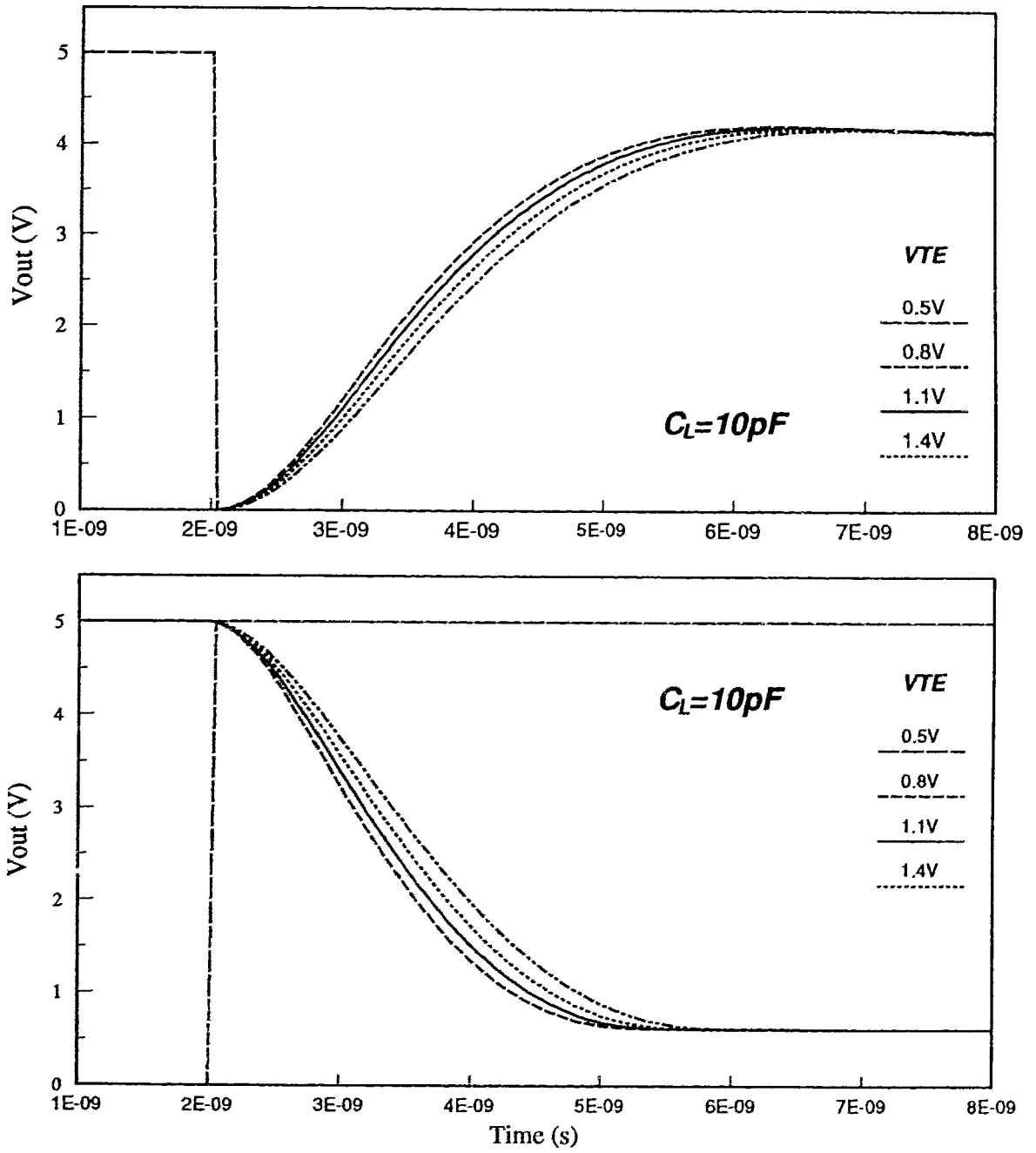
Figure 3.17 The transient output response for V_{TE} varied from 0.5 to 1.4V

CHAPTER 4:

IMPROVED ANALYTICAL DELAY MODEL

4.1 Introduction

Since the BiCMOS driver is mostly used in digital VLSI applications for its capability to drive large load capacitances at high speed, it is particularly important for the VLSI circuit designer to understand its transient characteristics quantitatively, and to derive its associated analytical delay model with an accurate and simple expression.

As discussed in chapter 2, the new BiCMOS driver (BCD-N) shown in Fig. 2.3 has approximately the same transient characteristics as that for the typical BiCMOS driver (BCD-4) shown in Fig. 2.1, so the transient response analysis carried out in this chapter was based on this typical BCD-4 circuit just for simplicity. All the model and results actually can be used directly for the case of BCD-N.

In its high speed applications with large load capacitance, the bipolar transistors (BJT) in the BiCMOS driver are usually operated under high-level injection condition, or in the saturation region [4.1-4.5]. Either mode of operation depends mainly on the parameters such as the Knee current I_{KF}, collector resistance R_C and load capacitance C_L. For the BiCMOS driver with the BJT operating in the saturation region, an analytical expression for the propagation delay τ_d can be found in [4.1,4.2]. However, when the BJT operates in the high-level injection region (so called Kirk effect), which is typical in the BiCMOS drivers, rather complicated analytical delay model were developed [4.4-4.6] based on Gummel-Poon large-signal BJT model [4.7]. In these models, the propagation delay time τ_d is usually expressed as the sum of τ_{d1} (referred to the turn-on time of the BJT), τ_{d2} (referred to the delay time in which the BJT is operating at low-level

injection condition) and τ_{d3} (referred to the delay time in which the BJT is operating at high-level injection condition). Since each delay time expression in these model are rather complicated (especially for τ_{d2} and τ_{d3}), they are obviously not easy to use. Meanwhile the empirical expression proposed in [4.2] seems quite simple and available for practical use. The propagation delay τ_d with the BJT under high-level injection condition was modelled in [4.2] as

$$\tau_d = \frac{V_{BE(on)}C_1}{I_{DS}} + \frac{AC_LV_S}{\sqrt{I_{KF}\beta_F I_{DS}}} \tag{4.1}$$

where A is an empirical constant. The first term in the above expression represents the turn-on time of the BJT, and is only important when driving small output capacitors. Although this empirical model is quite simple, we found that: (a) it has less accuracy than SPICE simulation results over certain ranges of parameter variation; (b) the empirical constant A in equation (4.1) is not easily determined; and (c) this τ_d expression does not depend on forward transit time τ_F, an important omission in equation (4.1).

In this chapter, an improved analytical propagation delay model, particularly suitable for high-level injection condition of the BJT in BiCMOS driver was developed [4.8] in section 2. This new model incorporates the Kirk effect into the Ebers-Moll BJT model so that the propagation delay τ_d in BiCMOS driver could be solved analytically. The following section shows detailed comparisons between this new model and the above empirical model based on the SPICE simulation results. Very good improvements compared to the model in [4.2] were achieved for this new model over the load capacitance ranges of 5 to 20pF with the typical parameter variations of I_{KF}, τ_F, β_F and I_{DS} respectively. In section 4, we discuss the effects of different V_{TH} or R_C value on this new model, and the optimal value for the bipolar device layout area design. Finally, the conclusions are presented in section 5.

4.2 Simple analytical delay model

Before deriving this new analytical delay model for BiCMOS driver, we first made the following assumptions for simplicity: (a) because the circuit of BCD-4 has similar characteristics when charging or discharging the load capacitance C_L due to emitter-follower configuration of bipolar transistor Q_1 or Q_2 respectively, only analysis of the rise time response was made in this paper, as is usually done [4.1,4.2]; (b) the propagation delay τ_d is defined as the rise time corresponding to 0-50% output voltage swing V_S, although this definition can also be extended to 10-90% rise time if necessary; (c) since the turn-on time of the bipolar transistor as expressed in the first term of the equation (4.1) is independent of either low-level or high-level injection mode of BJT, this term is temporarily neglected in the derivation, but will be finally added to our new model.

4.2.1 Equivalent circuit with Ebers-Moll BJT model
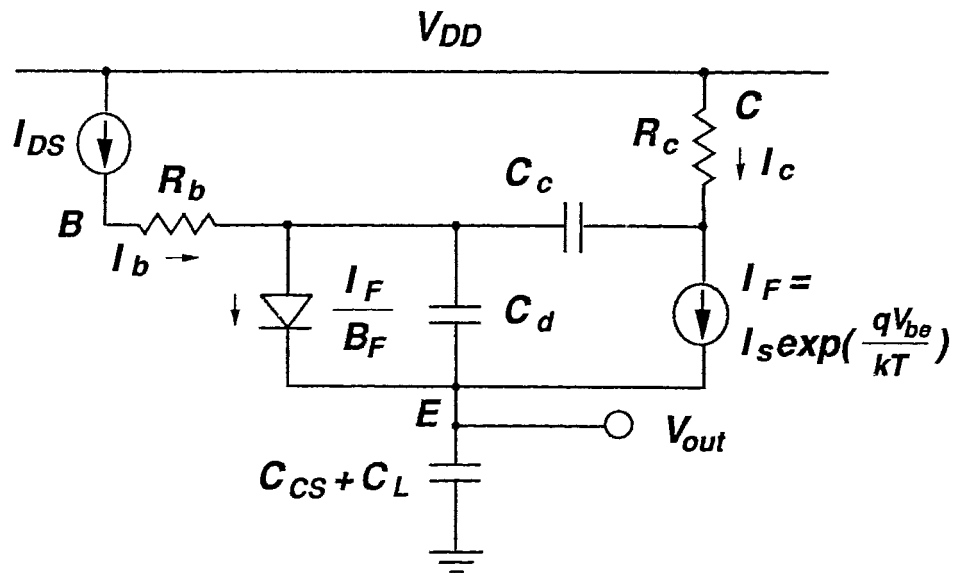
Figure 4.1 Equivalent circuit for rising response time analysis

Figure 4.1 shows the equivalent circuit of the BiCMOS driver associated with the rising response at the output. It is based on the Ebers-Moll (EM) large signal model for the BJT. All the parameters used in this equivalent circuit have the general physical meaning. Among these, two parameters I_F and C_D could be further expressed in the EM model [4.7] as

$$I_F = I_S \exp\left(\frac{qV_{BE}}{kT}\right)$$ (4.2)

$$C_D = \tau_F \frac{qI_F}{kT}.$$ (4.3)

Following the straightforward derivation in [4.1], the collector current transient response of $I_C(t)$ can be analytically expressed based on the equivalent circuit of figure 4.1 as

$$I_{C,l}(t) = \beta_F' I_{DS}\left\{1 - \exp\left(-\frac{t}{\beta_F' \tau_F'}\right)\right\}$$ (4.4)

with

$$\frac{1}{\beta_F'} = \frac{1}{\beta_F} + \frac{C_C}{C_{CS} + C_L}$$ (4.5)

$$\tau_F' = \tau_F + R_C C_C.$$ (4.6)

The subscript l in above $I_{C,l}(t)$ term indicates that this equation is only valid for low-level injection model of the BJT. For the case of $C_L \gg C_C$, which is the focus of our paper, β_F' is approximately equal to β_F. If the base-width modulation (Early effect) of the BJT is also included by using the associated parameter V_{AF} in modified EM model [4.7], then the forward current gain β_F should be further expressed as

$$\beta_F(V_{BC}) = \beta_F\left(1 - \frac{V_{BC}}{V_{AF}}\right).$$ (4.7)

The above equation (4.4) is also drawn in Fig. 4.2 as the solid line. It is clear to see the collector current $I_{C,I}(t)$ increases exponentially with time t. The increasing rate of $I_{C,I}(t)$ at t = 0 has the maximum value of

$$\frac{\delta I_C}{\delta t}\bigg|_{t=0} = \frac{I_{DS}}{\tau_F}. \tag{4.8}$$

As the time approaches infinity, the $I_{C,I}(\infty)$ will asymtotically reach the ideal static value (denoted as $I_{C,(EM)}$) and given by

$$I_{C,I}(\infty) = \beta_F' I_{DS} = I_{C,(EM)}. \tag{4.9}$$

Because no high-level injection effect is taken into account in above EM model, this maximum collector current $I_{C,(EM)}$ should be equal to current gain β_F' ($\approx \beta_F$) times base current I_{DS}.

In the above derivation, all EM model parameters C_E, C_C, C_{CS}, R_B, R_C, I_S, β_F, V_{AF} and τ_F are assumed to be constant, as required when using this model. The base current ($=I_{DS}$) could also be assumed constant (equal to I_{DSAT} of M1) during the time period of $0 < t < \tau_d$. This is because M_1 in Fig. 4.1 is mostly operated in the saturation region when its V_{DS} is reduced corresponding to this time period $0 < t < \tau_d$. The key parameters that determine whether the MOSFET is operated in saturation or linear region are the threshold voltage V_{TH}, and the maximum drift velocity v_{max} of carriers which is only important for short-channel length devices [4.4]. The effect on the delay time τ_d with different V_{TH} values is discussed in section 4.4.

4.2.2 Incorporation of high-level injection effect using Gummel-Poon BJT model

Since the BiCMOS circuit is usually used to drive large load capacitances with relatively short delay time, then the bipolar transistor Q_1 would be operating mostly under high-level injection conditions. Therefore, a more accurate Gummel-Poon (GP)

large signal BJT model with the associated parameter I_{KF} (called Knee current) is required [4.4,4.5]. This I_{KF} current has the specific physical meaning of representing the corner for the forward-beta high-current roll-off. It can be easily extracted from the Gummel plot of the BJT and its typical value is 0.1~10mA [4.4]. Due to this high-level injection effect, the current gain β_F at $I_C > I_{KF}$ is no longer a constant, so the static collector current I_C should be smaller than $I_{C,(EM)}$ in equation (4.9). The analytical expression for larger I_C in GP model can be obtained through the normalized majority base charge q_b as [4.7]:

$$
\begin{aligned}
I_C &= \frac{I_S}{q_b}\exp\left(\frac{qV_{BE}}{kT}\right) \\
&\approx \sqrt{I_{KF}I_S}\exp\left(\frac{qV_{BE}}{2kT}\right) \qquad for \ (I_C > I_{KF}).
\end{aligned}
\qquad (4.10)
$$

The base current I_B in the GP model is

$$
I_B = \frac{I_S}{\beta_F}\exp\left(\frac{qV_{BE}}{kT}\right) \qquad (4.11)
$$

so the static collector current of the BJT in BiCMOS circuit (with $I_B=I_{DS}$) under high-level injection condition (denoted as $I_{C,(GP)}$) could be expressed as

$$
I_{C,(GP)} = \sqrt{I_{KF}\beta_F I_{DS}} \qquad for \ (I_C > I_{KF}). \qquad (4.12)
$$

It is clear to see that the high-level injection effect $(I_C > I_{KF})$ will lead to reducing the static collector current significantly from $\beta_F I_{DS}$ in EM model to $\sqrt{I_{KF}\beta_F I_{DS}}$ in GP model. On the other hand, for low-level injection mode $(I_C < I_{KF})$, the collector current expression $I_C=\beta_F I_B$ is still valid. Also, close examination of equation (4.1) reveals that the delay time τ_d used in this empirical model is actually obtained by assuming a constant collector current (equal to the static value of $I_{C,(GP)}$) charging the load capacitance C_L. It obviously overestimates the collector charging current $I_C(t)$ during the time period of $0 < t < \tau_d$, which actually increases exponentially from 0 to the final static value of $I_{C,(GP)}$.

Since it is complicated to use the complete GP BJT model, which covers both low-level and high-level injection condition [4.7], to solve for the collector current transient response $I_C(t)$ in the circuit of Fig. 4.1, we modified the EM model-related equation (4.4) by incorporating the Kirk effect in the following way

$$
\begin{aligned}
I_{C,h}(t) &= \beta_F^* I_{DS} \left\{ 1 - \exp\left(-\frac{t}{\beta_F^* \tau_F'} \right) \right\} \\
&\approx \beta_F^* I_{DS} \left\{ 1 - \exp\left(-\frac{t}{\tau^*} \right) \right\}
\end{aligned}
\tag{4.13}
$$

with

$$
\beta_F^* = \sqrt{\frac{I_{KF} \beta_F'}{I_{DS}}}
\tag{4.14}
$$

$$
\tau^* = \beta_F^* \tau_F'.
\tag{4.15}
$$

Here, the subscript h in above $I_{C,h}(t)$ stands for high-level injection condition. This equation has the following two features

i). For the time period of $I_{C,h}(t) > I_{KF}$, the collector current will reduce significantly and reach the final static value of $\beta_F^* I_{DS} = \sqrt{I_{KF} \beta_F' I_{DS}}$ at $t = \infty$, which is approximately the same as $I_{C,(GP)}$ given by equation (4.12). It also means the current gain decreases correspondingly to the equivalent value of β_F^* given by equation (4.14);

ii). For the time period when $I_{C,h}(t) < I_{KF}$, the collector current should have almost same response behavior as described by the EM model equation (4.4), especially the slope of $\delta I_C / \delta t$ at $t = 0$ which is exactly the same as I_{DS}/τ_F' in equation (4.8).
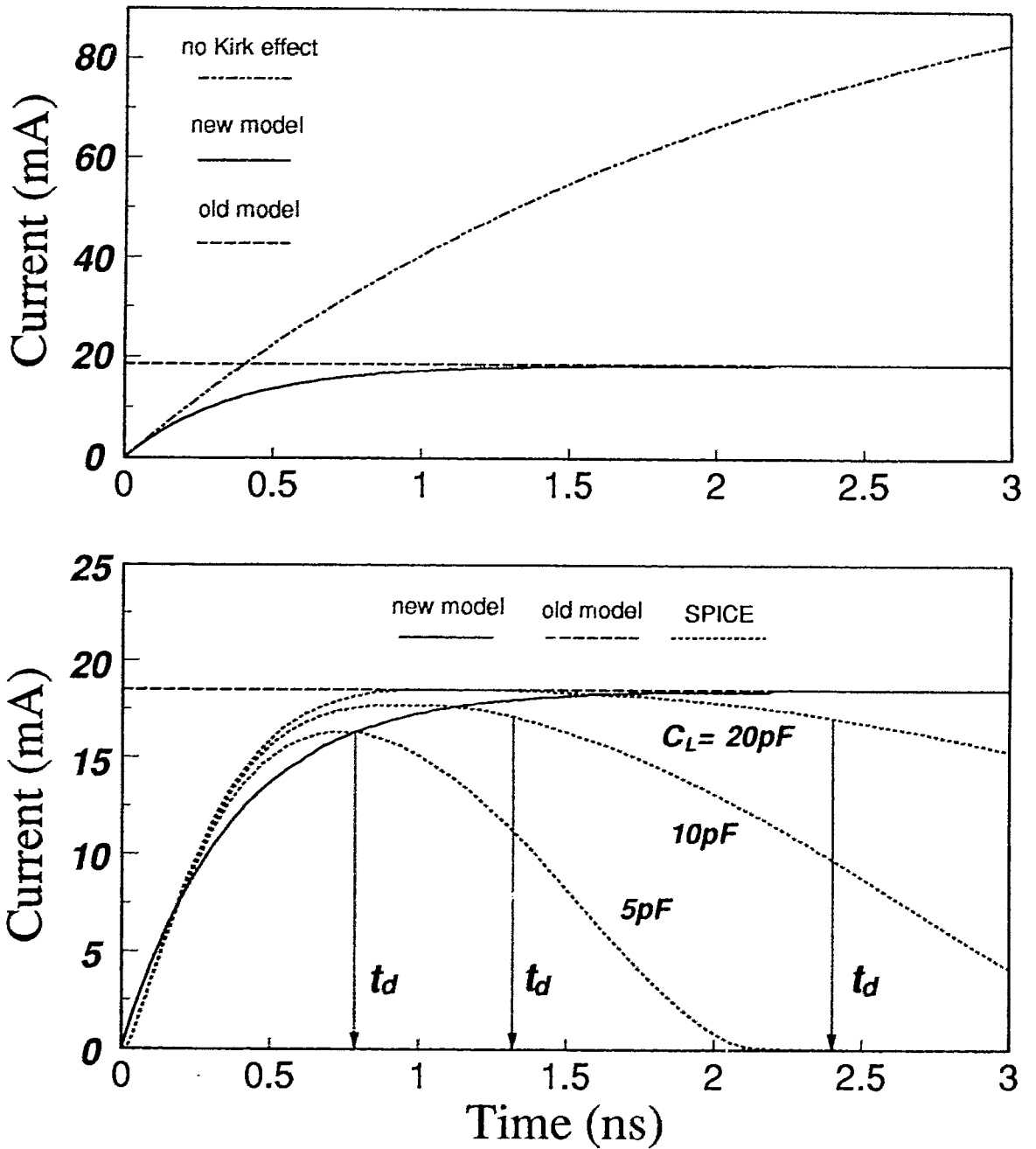
Figure 4.2 Collector current transient response of NPN BJT
with C_L variations of 5-20pF (τ_d from SPICE results)

Figure 4.2 also shows this $I_{C,h}(t)$ curve, together with the SPICE simulation result of

$I_C(t)$ for BiCMOS circuit shown in Fig. 2.1. The device parameters used in SPICE

simulation are listed in Table 4.1.

NPN BJT:	IS=1E-16A, BF=100, IKF=3E-3A, VAF=20V, RB=100 Ω, RC=20Ω, RE=2Ω, CJE=50fF, CJC=50fF, CJS=100fF, TF=20pF, TR=500pF
NMOS/ PMOS:	Level=3, L=1.0μm, W=10μm(M_1,M_3)/5μm(M_2,M_4), VTO=1.0/-1.0V, TOX=20μm, GAMMA=0.6/-0.4 $\sqrt{V}$, UO=500/200cm^2/(Vs), THETA=0.05V^{-1}, VMAX=5E4m/sec

Table 4.1. The NPN BJT and NMOS/PMOS devices parameters used in SPICE simulations of transient response.

From Fig. 4.2 it is apparent that, for $0 < t < \tau_d$, the curve $I_{C,h}(t)$ is much closer to

the simulated $I_C(t)$ from SPICE for both low-level and high-level injection conditions.

By comparison, the curve $I_{C,l}(t)$ derived from EM model is only correct for the smaller

collector current of $I_C < I_{KF}$, and the constant current curve of $I_{C,(GP)}$ used in the empirical

model [4.2] is only valid for larger collector current of $I_C > I_{KF}$. In this Fig. 4.2, the

simulated $I_C(t)$ decreases on further increase of time $t > \tau_d$. This corresponds to the

decrease of the base current I_{DS}, due to the PMOS device M_1 entering the linear region

from its saturated mode as the output voltage increases.

4.2.3 The output voltage response $V_O(t)$ and delay model τ_d

Using the above modified $I_{C,h}(t)$ equation (4.13), the rising response of the output

voltage V_O can be easily derived as

$$\frac{dV_O(t)}{dt} = \frac{\beta_F^* I_{DS}}{C_{CS} + C_L}\left\{1 - \exp\left(-\frac{t}{\tau^*}\right)\right\}.$$

(4.16)

By integrating the above equation with the output voltage swing of $0.5V_s$ and the corresponding time period of 0 to τ_d, the delay time τ_d is found to be

$$\tau_0 = \tau_d - \tau^* \left\{ 1 - \exp\left(-\frac{\tau_d}{\tau^*}\right) \right\}. \tag{4.17}$$

with τ_0 given by

$$\tau_0 = \frac{0.5V_S(C_{CS} + C_L)}{\sqrt{I_{KF}\beta_F' I_{DS}}}. \tag{4.18}$$

Because the equation (4.17) is a hypergeometric function, it is very difficult to solve τ_d directly. Analytical expressions for τ_d can be only approximated for the following two cases

i) $\tau_d > \tau^*$.

In this case, since the delay time τ_d is larger than τ^* (where $\tau^* = \beta_F^* \tau_F'$), then the exponential term in equation (4.17) $\tau^* \cdot \exp(-\tau_d/\tau^*) \approx 0$ so that $\tau_d \approx \tau_0 + \tau^* \approx \tau_0$ since $\tau_d > \tau^*$. Therefore, to the first order approximation, by substituting τ_0 for τ_d of the exponential term in the equation (4.17), the delay time τ_d could be expressed as

$$\tau_{d,h} = \tau_0 + \tau^* \left\{ 1 - \exp\left(-\frac{\tau_0}{\tau^*}\right) \right\}. \tag{4.19}$$

The physical meaning of this equation (4.19) is now described. The first term (τ_0) is just the delay time used in the empirical model equation (4.1) with A = 0.5, and represents the time constant for charging load capacitance C_L with constant collector current $I_{C,(GP)}$. The second term represents the compensation due to the overestimation of $I_C(t)$ in τ_0, and also incorporates the time constant $\tau^* = \beta^*(\tau_F + R_C C_C)$ required for charging C_D and C_C. Therefore, the above new delay time model is suitable for the BiCMOS driver with larger load capacitance, where the BJT is mostly operated under high-level

injection condition. It is also expected that the accuracy of this model would be improved by the additional second term in equation (4.19), this point will be demonstrated in the following section.

ii) $\tau_d < \tau^*$.

This is usually the case for very load capacitance C_L. By expanding the exponential term in equation (4.17) using a Taylor series and keeping the first three terms, τ_d can be solved analytically as

$$\tau_{d,l} = \sqrt{2\tau^*\tau_0}$$

$$= \sqrt{\frac{(\tau_F + R_C C_C)(C_{CS} + C_L)V_S}{I_{DS}}}. \tag{4.20}$$

This expression is similar to that developed in [4.1], where only the EM BJT model was used and low-level injection condition was assumed. Therefore the modified collector current equation (4.13) proposed here can also be used for low-level injection condition of the BJT.

The criterion to distinguish the above two operating conditions of the BJT in BiCMOS has already been discussed in [4.2]. It can also be roughly estimated using this new model by comparing τ_0 and τ^* in equations (4.18) and (4.15) respectively. Based on the discussion above and $\tau_0 > \tau^*$, then $\tau_{d,h}$ given by equation (4.19) should be used, otherwise $\tau_{d,l}$ given by equation (4.20) is desirable. Furthermore, this criterion could be evaluated by the load capacitance and related device parameters as

$$\frac{\tau_0}{\tau^*} = \frac{0.5V_S(C_{CS} + C_L)}{\sqrt{I_{KF}\beta_F' I_{DS}}} / \sqrt{\frac{I_{KF}\beta_F'}{I_{DS}}(\tau_F + R_C C_C)}$$

$$\approx \frac{0.5V_S(C_{CS} + C_L)}{I_{KF}\beta_F'(\tau_F + R_C C_C)}. \tag{4.21}$$

78

Therefore if the calculated value of τ_0/τ^* is great than one, it indicates the BJTs in the

BiCMOS drivers is operating in the high-level injection mode, otherwise this BJTs is

under the low-level injection condition.

4.2.4 The final analytical propagation delay expression of τ_d

Since the new delay model equations (4.19-4.20) derived above are based on the

assumption that the BJT is already turned-on at $t = 0$, then the final propagation delay

expression τ_d for BiCMOS driver should include the additional turn-on time of the BJT

as already used in equation (4.1):

i) for high-level injection condition of BJT, equation (4.19) is modified to

$$\tau_{d,h} = \frac{V_{BE(on)}}{I_{DS}} + \tau_0 + \tau^* \left\{ 1 - \exp\left(-\frac{\tau_0}{\tau^*} \right) \right\}. \tag{4.22}$$

ii) for low-level injection condition of BJT, equation (4.20) is modified to

$$\tau_{d,l} = \frac{V_{BE(on)}C_1}{I_{DS}} + \sqrt{\frac{(\tau_F + R_C C_C)(C_{CS} + C_L)V_S}{I_{DS}}}. \tag{4.23}$$

Usually this turn-on time expression of the BJT is only important for the $\tau_{d,l}$ given

by equation (4.23). For the larger load capacitance under high-level injection condition,

first term on the right hand side of equation (4.22) is so small that it could be neglected

without introducing significant errors in the results.

4.3 Comparison with SPICE simulation results

In order to verify this new analytical propagation delay model given by equations

(4.22) and (4.23), especially for the $\tau_{d,h}$ model given by equation (4.22) which is more

important in BiCMOS drivers, SPICE [4.9] is used to compare its simulated results with

those calculated with the new model. The SPICE program includes not only a completely built-in GP BJT model, but it also has the capability to simulate the integrated circuit performance. The device parameters used for the SPICE simulations are listed in Table 1. The load capacitance is chosen from 5pF to 20pF to meet the high-level injection condition of the BJT. The delay time τ_d is then calculated with the definition of 50% output voltage response. This simulation procedure was also repeated for varying parameters like I_{KF}, τ_F, β_F and I_{DS} respectively. Meanwhile, the new model equation (4.22) and the empirical model equation (4.1) with A = 0.5 are also used for the calculated results. The turn-on time term in both equations (4.1) and (4.22) are neglected, as explained before. The detailed comparison between SPICE and analytical results are now briefly described.

4.3.1 I_{KF} variations from 1mA to 9mA

Figure 4.3 shows the propagation delay τ_d versus Knee current I_{KF}. It is clear to see that for the entire range of I_{KF} and C_L, the new model results are in good agreement to the simulation results, with only small difference (about 2~10%) for smaller C_L and larger I_{KF}, or larger C_L and smaller I_{KF}. On the other hand, the τ_d values from the empirical model equation (4.1) predicts too low τ_d (about 20-50%), especially for larger I_{KF}. This is because the empirical model equation (4.1) misses the third term used in our new model equation (4.22) which has the meaning of compensating the overestimation of the charging current of $I_C(t) \equiv I_{C,(GP)}$ for τ_d calculation, and this term is roughly proportional to $\sqrt{I_{KF}}$.
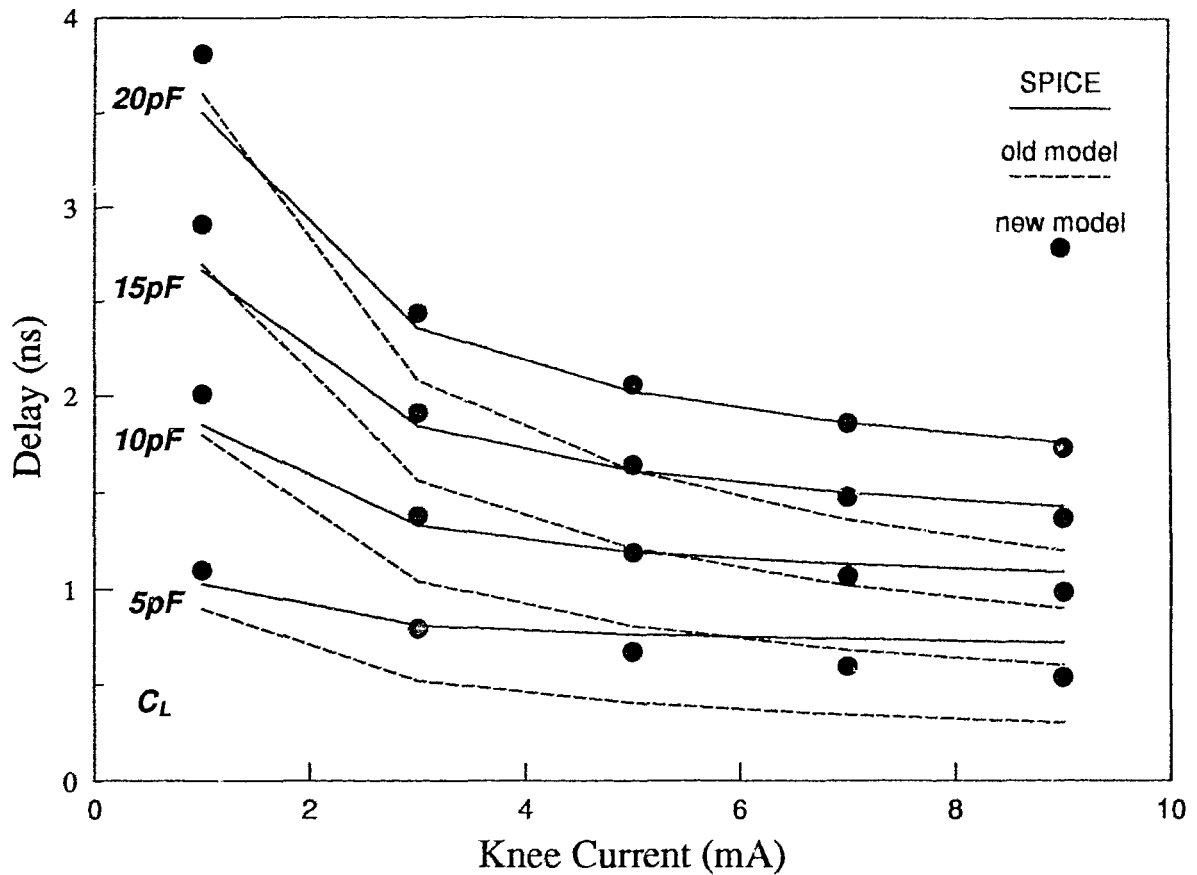
Figure 4.3 Propagation delay τ_d **vs. Knee current** I_{KF} **of NPN BJT**

with C_L **variations of 5-20pF**

4.3.2 β_F variations from 50 to 250

Figure 4.4 shows the propagation delay τ_d versus current gain β_F. The agreements between the SPICE results and the new model calculation are also very good for the β_F and C_L variations. Some small error (about 5-8%) is expected with the smallest C_L of 5pF and larger β_F. The empirical model shows a similar large error behavior as the above parameter I_{KF} does. It is simply due to the same reason for I_{KF} as described above, since the term $I_{KF}\beta_F$ is present in our new model equation (4.22) and the empirical model

equation (4.1).
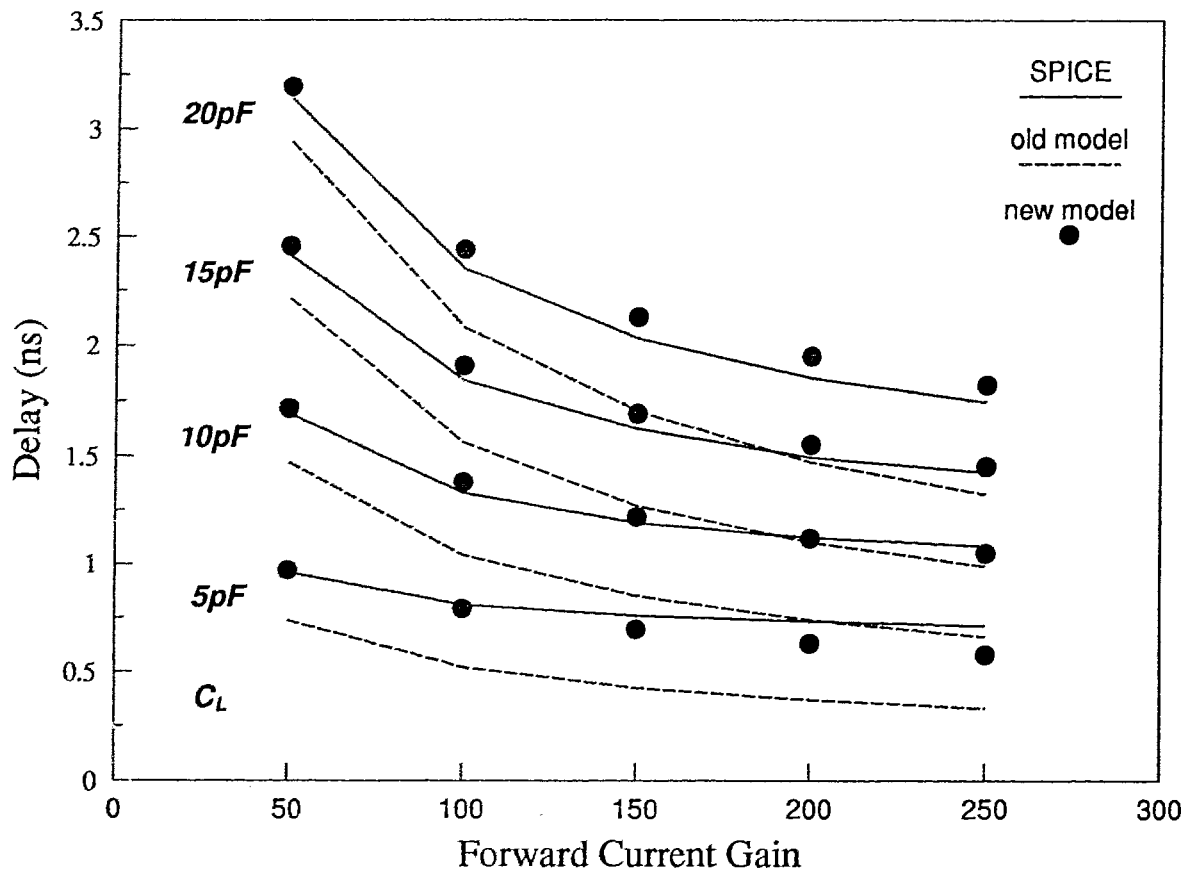
Figure 4.4 Propagation delay τ_d **vs. forward current gain** β_F **of NPN BJT**

with C_L **variations of 5-20pF**

4.3.3 τ_F variations from 10ps to 50ps

The propagation delay τ_d versus forward transit time τ_F is shown in Fig. 4.5. Due to the fact that τ_F is not present in the empirical model, the calculated τ_d values is independent of τ_F. The SPICE simulation indicates that the delay time τ_d is almost linearly proportional to the τ_F. The new model equation (4.22) also implies this relationship, and shows quite good agreements with the simulation results. Only some error ($< 10\%$)

appears for larger τ_F values. Since the forward transit time τ_F is linearly proportional to the emitter-base diffusion capacitance C_D, it could be expected that this τ_F value should have some kind of linear relationship with the turn-on switching time of BJT, and also correspondingly the propagation delay τ_d of the BiCMOS drivers.
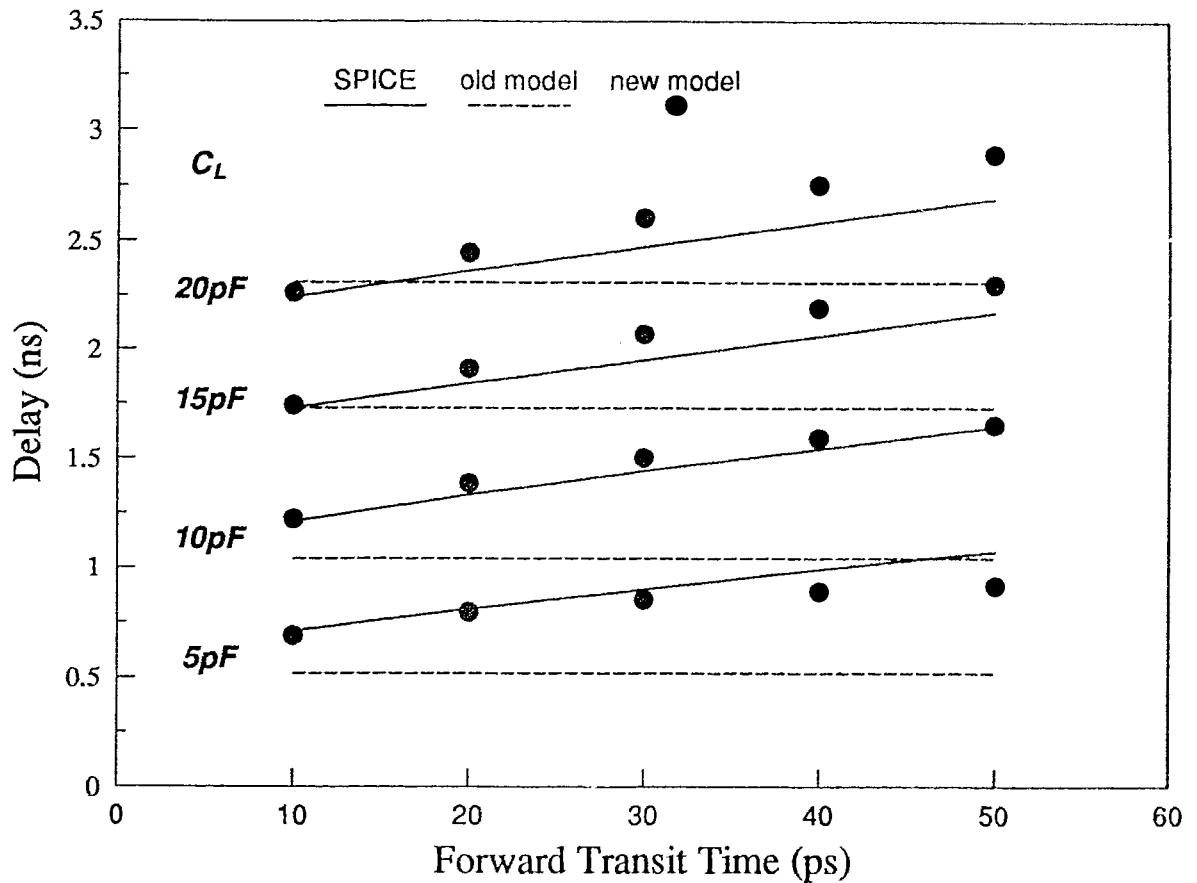
Figure 4.5 Propagation delay τ_d **vs. forward transit time** τ_F **of NPN BJT with** C_L **variations of 5-20pF**

4.3.4 I_{DS} variations from 0.54mA to 2.70mA

Figure 4.6 shows the propagation delay τ_d versus current gain I_{DS}. The I_{DS} variation is implemented by altering the channel length over channel width ratio (W/L) from 5/1

to 25/1 for the MOS devices in SPICE simulations. The new model results are also in very good agreement with the SPICE results. The empirical model results are about 5-20% lower than the simulation results for all C_L values used. For smaller I_{DS}, this error becomes more severe. This is because the second term in new model equation (4.22) is inversely proportional to $\sqrt{I_{DS}}$.
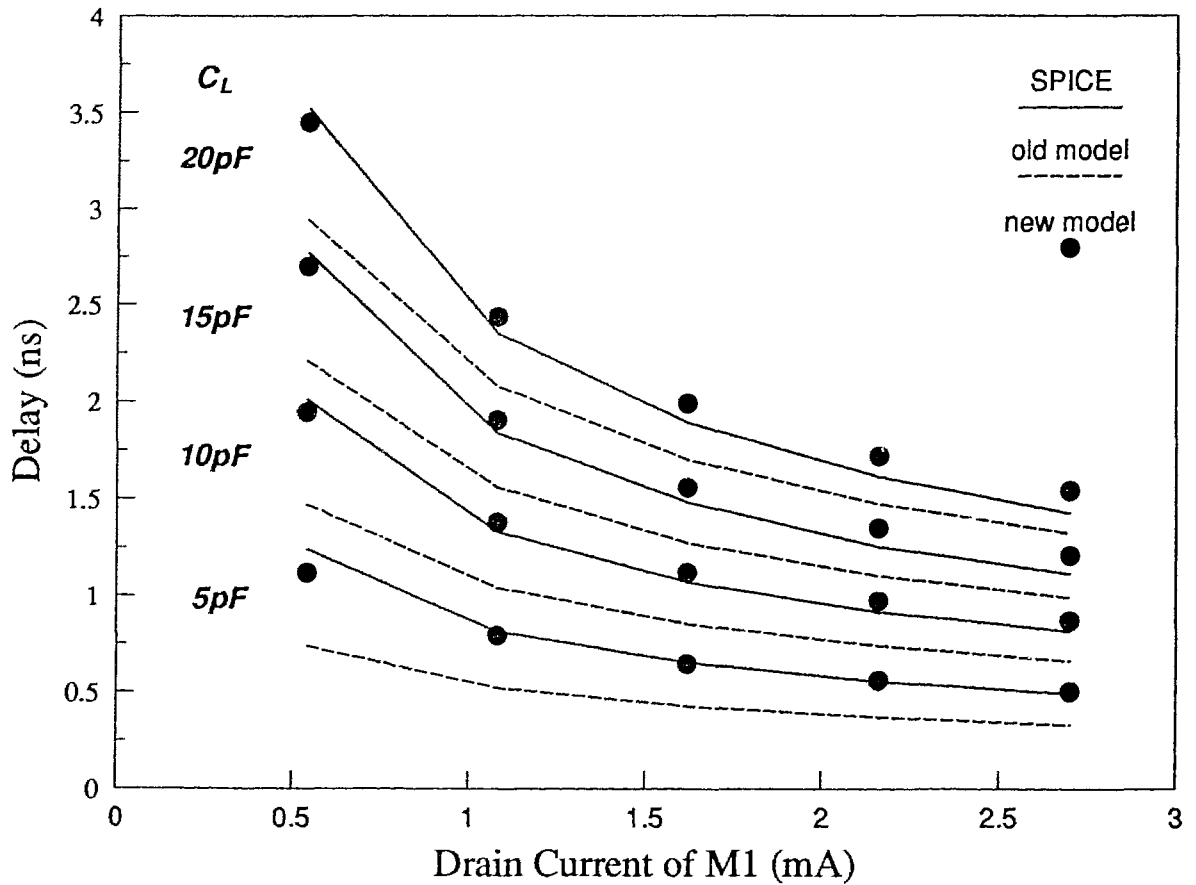
Figure 4.6 Propagation delay τ_d vs. drain current I_{DS1} of M1 with C_L variations of 5-20pF

Based on the comparisons above, the new model equation (4.22) proposed in this paper shows significant improvement with respect to the empirical model equation (4.1). Although the accuracy of the empirical model could be improved by adjusting the

empirical constant A carefully, it would be very difficult for it to be optimized over the entire range of device parameter variations and load capacitance variations as demonstrated above.

4.4 Effect of some important parameter variations

4.4.1 The effect of V_{TH} variation on propagation delay τ_d

In section 4.2, the new analytical delay model τ_d was developed by the assumption of the constant I_{DS} within the time period of $0 < t < \tau_d$. This is usually correct because the PMOS device M_1, which offers the drain current I_{DS} as shown in Fig. 4.1, is mostly operated in the saturation region at this time period. As the output voltage rises from $V_{OL} \approx V_{CE,sat}$ to $0.5(V_{OH} + V_{OL}) \approx 0.5V_{DD}$ [4.3], the drain voltage V_{DS} (in absolute value) of M_1 would correspondingly reduce from $\sim(V_{DD} - V_{BE} - V_{CE,sat})$ to $\sim(0.5V_{DD} - V_{BE})$. The criterion for MOS devices to operate in the saturation region is $V_{DS} \geq V_{D,sat}$, here the saturation voltage $V_{D,sat}$ could be expressed in MOS model Level 3 of SPICE for modern short channel length MOS device as

$$V_{D,sat} = V_a + V_b - \sqrt{V_a^2 + V_b^2} \tag{4.24}$$

with

$$V_a = \frac{V_{GS} - V_{TH}}{1 + F_B} \tag{4.25}$$

$$V_b = \frac{v_{max} L_{eff}}{\mu_s}. \tag{4.26}$$
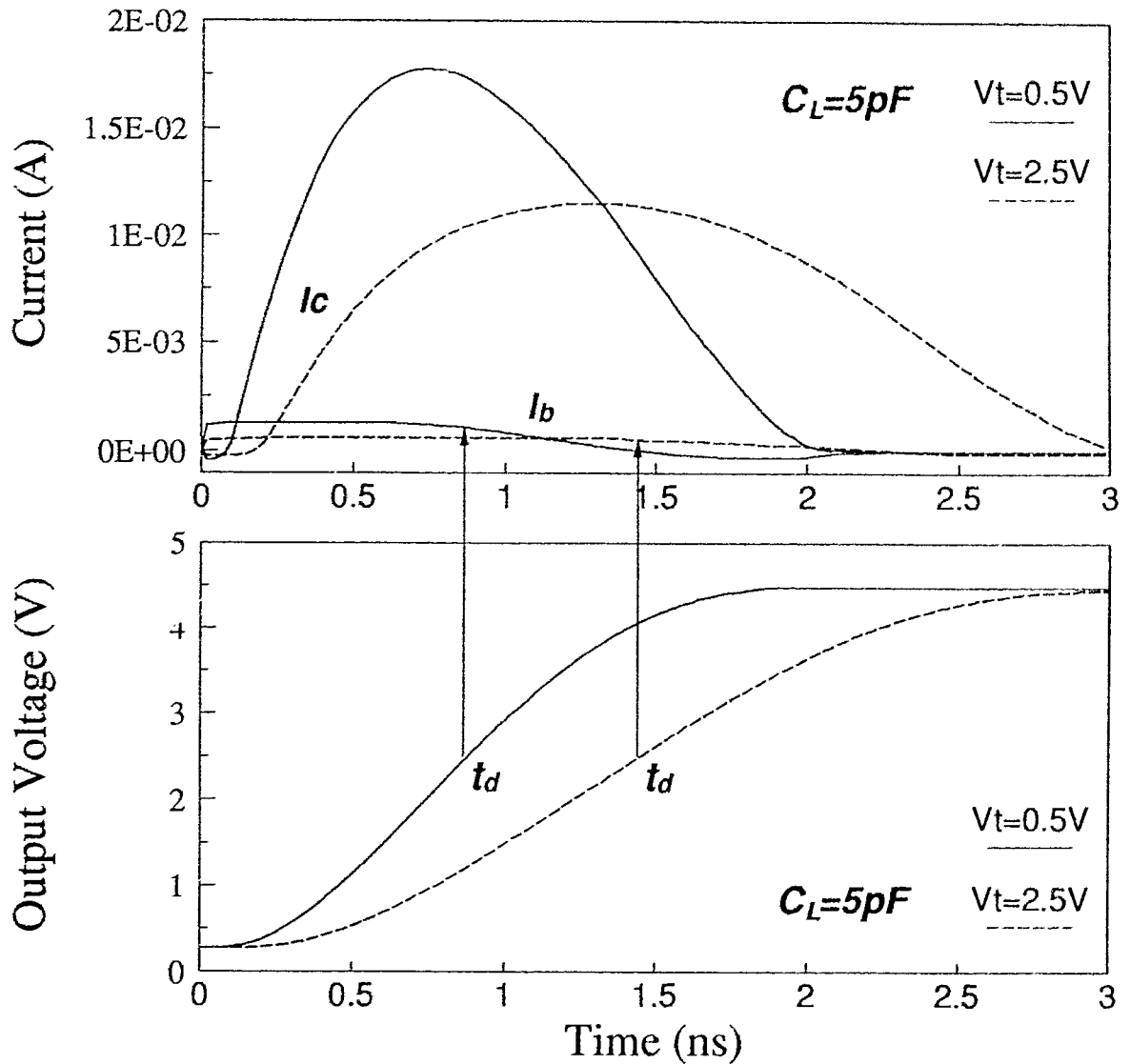
Figure 4.7 Transient response of I_C, I_B **and** V_{out} **with** V_{TH}

of 0.5V and 2.5V respectively

Therefore it is clear to see that the $V_{D,sat}$ is not only determined by the threshold voltage

V_{TH}, but also reduced considerably by the maximum drift velocity v_{max} of carriers for

short-channel length device. Larger V_{TH} or smaller L_{eff} would lower the $V_{.,sat}$ value in

favor of meeting the constant I_{DS} condition. Figure 4.7 shows the simulated transient

response of the drain current $I_{DS}(t)$, collector current $I_C(t)$ and output voltage $V_{OUT}(t)$ for

BiCMOS driver from SPICE. The parameters used here are the same as given previously in Table 4.1, with V_{TH} = 2.5V and 0.5V respectively, which corresponds the possibly-chosen maximum and minimum value for suitable inverter operation. Also the load capacitance C_L is fixed at 5pF for the worst case simulation. The results confirmed very well that for typical 1μm MOS technology with typical v_{max} value of 5E4m/sec and the lowest V_{TH} value of 0.5V, the $I_{DS}(t)$ value in the time period of $t \leq \tau_d$ could still be kept almost constant. As the threshold voltage V_{TH} increases further to 2.5V, this constant current characteristic would become more obvious due to a higher saturation voltage. In case of some extremely low V_{TH} and C_L values happened, we suggest that in stead of using the maximum I_{DS} value above, choosing the average value of I_{DS} in the time period of $0 < t < \tau_d$ is still good enough for the new delay model equations (4.22) and (4.23) presented in this paper as also used in reference [4.2].

4.4.2 The optimal AREA value of the BJT based on minimum propagation delay τ_d

It is clear that the optimal design of the BJT is quite important for minimizing the delay time of BiCMOS driver. This could be a very complicated problem to solve analytically, with too many parameters to consider. For example, when the BJT is operated under high-level injection conditions, we saw from Fig. 4.3 that it results in delay time τ_d would monotonically decrease as the Knee current I_{KF} increases. Due to the linearly proportional relationship of this parameter versus the device area [4.6], it seems possible to increase the I_{KF} by designing a larger BJT device area. This design would also reduce the parasitic resistance like RC, RB and RE, but would inevitably increase those parasitic junction capacitance like CJE, CJC and CJS. Therefore the optimal design for BJT layout dimension should be examined. Although it is hard to express this feature in a simple form based on our new delay model, we feel it is worth

while to see how it works by simulation since there is a parameter called the AREA factor used in SPICE for the BJT model, which represents the number of equivalent parallel devices of a specified model. Figure 4.8 shows the related simulation results of τ_d vs. AREA with C_L variations of 5pF to 20pF. It is quite interesting to see that as the load capacitance C_L increases from 5pF to 20pF, the optimal value of parameter AREA would correspondingly increase from ~1 to ~3. Of course, this result is strongly related to the specified model used in SPICE simulation, but the existence of the optimal value for the BJT device area design is apparent and must be carefully considered in design.
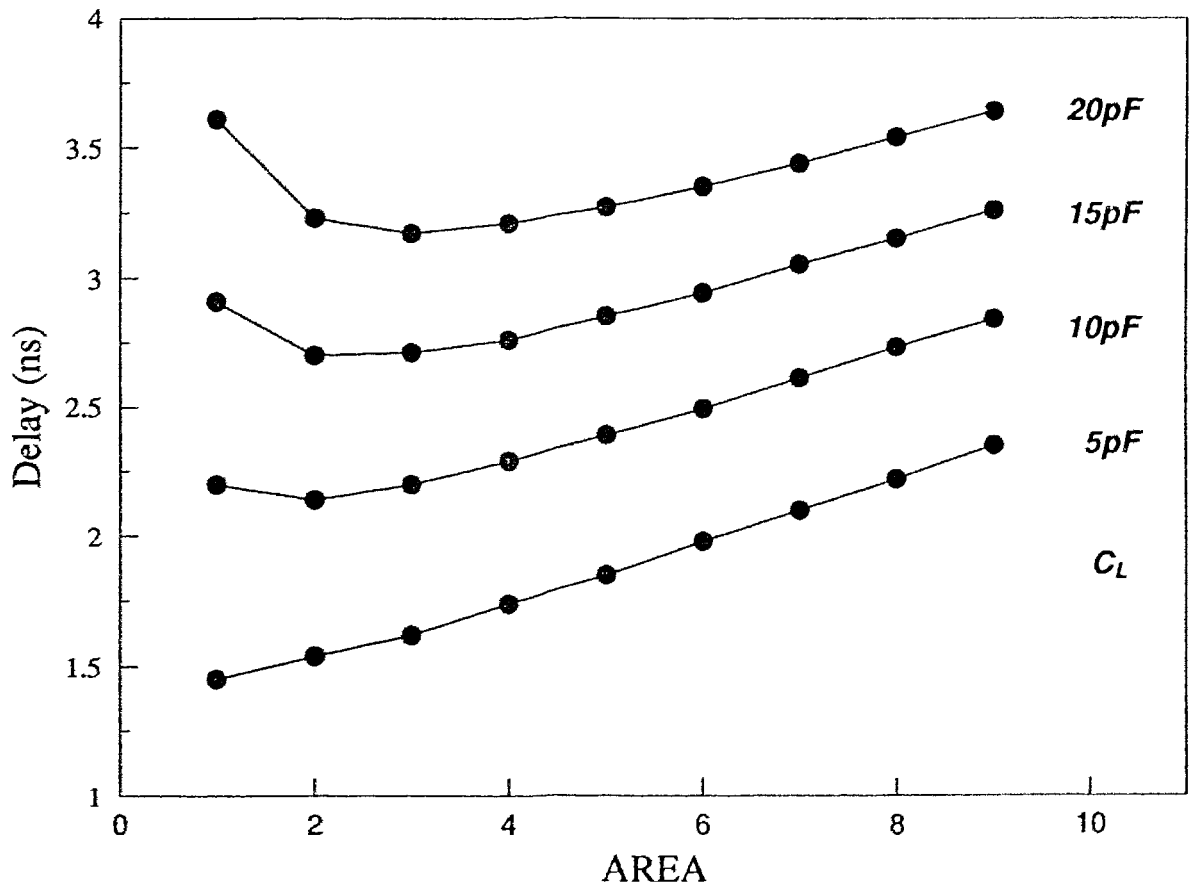
Figure 4.8 Propagation delay τ_d vs. the AREA factor of NPN BJT with C_L variations of 5-20pF

CHAPTER 5:

SUGGESTED FUTURE WORK

5.1 Further improvement for BiCMOS driver circuit design

Based on the above detailed description of the new BiCMOS driver circuit shown previously in Fig. 2.3, it is clear from the circuit design point of view, that devices such as M1, M2, M3 (form CMOS inverter), and Q1, Q2 (offer large current) play a fundamental role for the BiCMOS driver function, and that two other devices M4 and M5 also have an important impact on both DC and transient characteristics for this BiCMOS driver. The ideal design for these two devices are

i) when the BiCMOS driver is operating in the stable output state of either low "0" level or high "1" level, these two devices (M4 and M5) should be on in order to reach the lowest output of $V_{OL} = V_{SS}$, or the highest output of $V_{OH} = V_{DD}$ for better DC characteristics of this circuit;

ii) when the BiCMOS driver is operating in the switching state of either rising response or falling response, then these two devices (M4 and M5) should be off in order to offer the maximum charge or discharge base current to Q1 or Q2 for better transient characteristics of the driver circuit.

Using the above conclusions, the new BiCMOS driver circuit proposed in this thesis could be further improved by the following modifications.

i). For M5, its "on-resistor" behavior does not affect the rising output response significantly, as already explained in chapter 2. However, if this negative effect from "on-resistor" M5 has to be diminished, one could either try to find a similar "dynamic-resistor-behavior" device for M5 to meet the above ideal requirement, or just replace this depletion mode M5 in BCD-N back with the enhancement mode M5 in

BCD-5 for a minor improvement. As mentioned earlier in section 2.2.3, it is not very easy to substitute this M5 by a dynamic-resistor-like device, but it is still worth investigating for possible circuit performance improvement.

ii). For M4, it has been designed to function like a "dynamic-resistor-behavior" device, so it really demonstrates the improvement of the falling output response while maintaining the lowest static logic "0" output of $V_{OL} = V_{SS}$. But this M4 device also turns off when the output voltage approaches $V_{OH} = V_{DD}$. This off-state of M4 has no problem with static output of V_{OH} as already shown in Fig. 2.4, but it would cause the transient logic "1" voltage to be a little lower, as mentioned previously in section 2.4.3. This undesirable transient response waveform with higher output voltage becomes more serious for smaller load capacitance as shown in Fig. 5.1. The reason is because during the rise-time period, the base terminal of Q2 is floating due to the off-state of both M4 and M3, then Q2 could be turned-on temporarily through the parasitic capacitors between nodes (4) and (3), while the output voltage increases. The faster output response with smaller C_L would results in more current flowing through Q2. Fig. 5.2 clearly shows the required charge current $I_{C1}(t)$ and the undesirable current $I_{C2}(t)$ with $C_L = 2pF$ and 10pF respectively. The ratio of the maximum $I_{C2}(t)$ over the maximum $I_{C1}(t)$ is ~8.8% ($C_L = 2pF$) or ~3.6% ($C_L = 10pF$). Also this $I_{C2}(t)$ lasts less than 10ns. By comparison, in BCD-5 this undesirable current $I_{C2}(t)$ also occurs, but with much lower current ratio of ~3.6% ($C_L = 2pF$) and ~0.8% ($C_L = 10pF$) respectively as seen in the same figure 5.2. This is because during this rise-time period, M4 in BCD-5 is on and the base terminal of Q2 is connected to the ground through a relatively small equivalent resistor. But the simulation results in Figs. 5.1 and 5.2 show clearly that (a) this problem does not affect most of the rising transition response, or increases the propagation delay time τ_d; and (b) the lowest output voltage for logic "1" is still larger than 4V, high enough for the

reliable logic function. Since it is apparent that the BiCMOS driver is usually used for the application with large load capacitance, then this should not be a serious problem. However, if this off state of M4 during the rising time period is not desirable for some specific applications like very small load capacitance, then this BiCMOS driver circuit might have to be redesigned for minimizing this potential problem.

iii) Symmetric DC characteristics can be achieved by increasing V_{TE} as shown previously in Fig. 3.16. However this is undesirable as it results in increased propagation delay. Symmetric transient characteristics can be obtained by optimizing the MOS device size of M1 and M3. This should be investigated in future before the practical realization of BCD-N occurs.
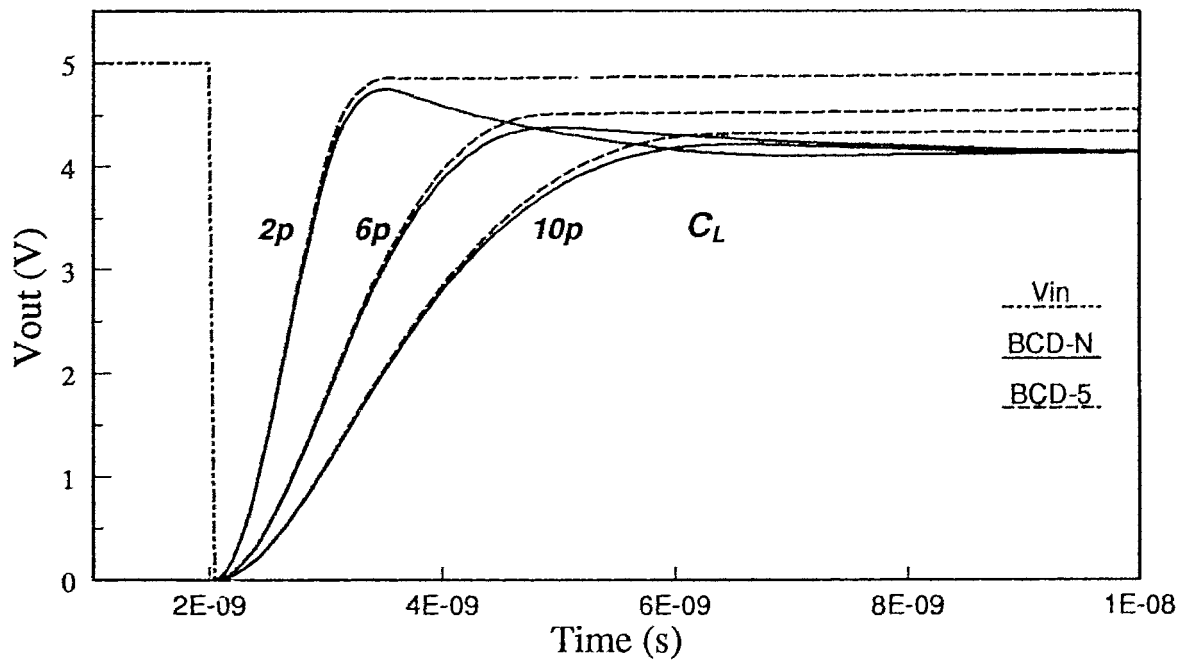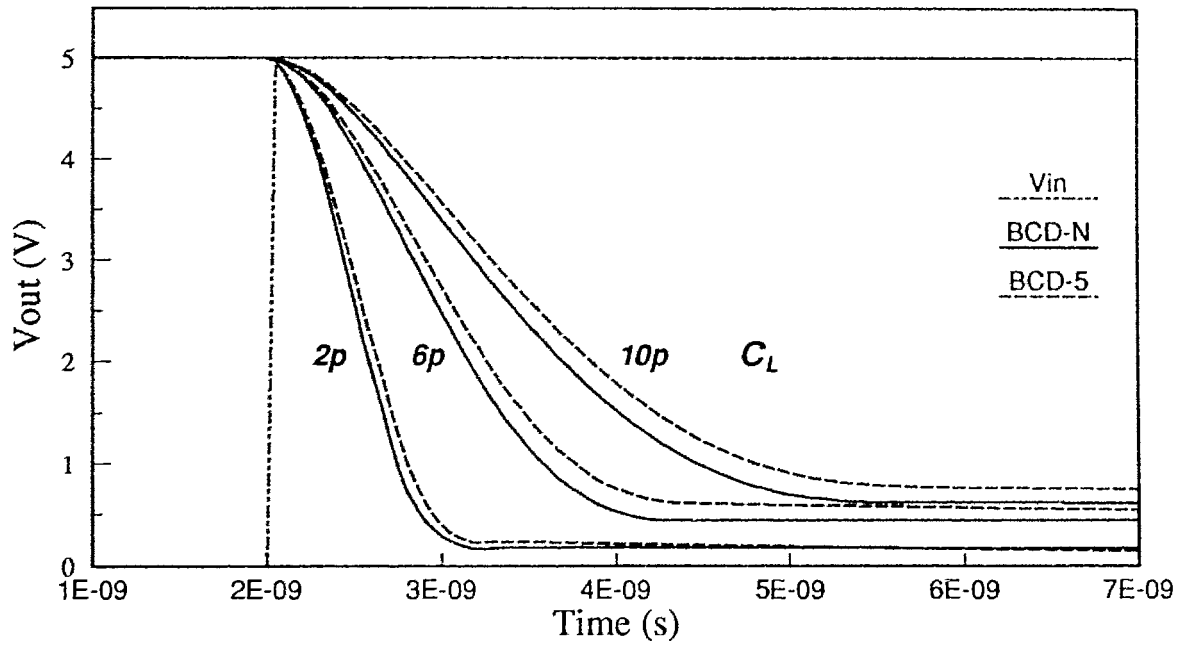
Figure 5.1 The transient output response waveform for C_L varied from 2 to 10pF
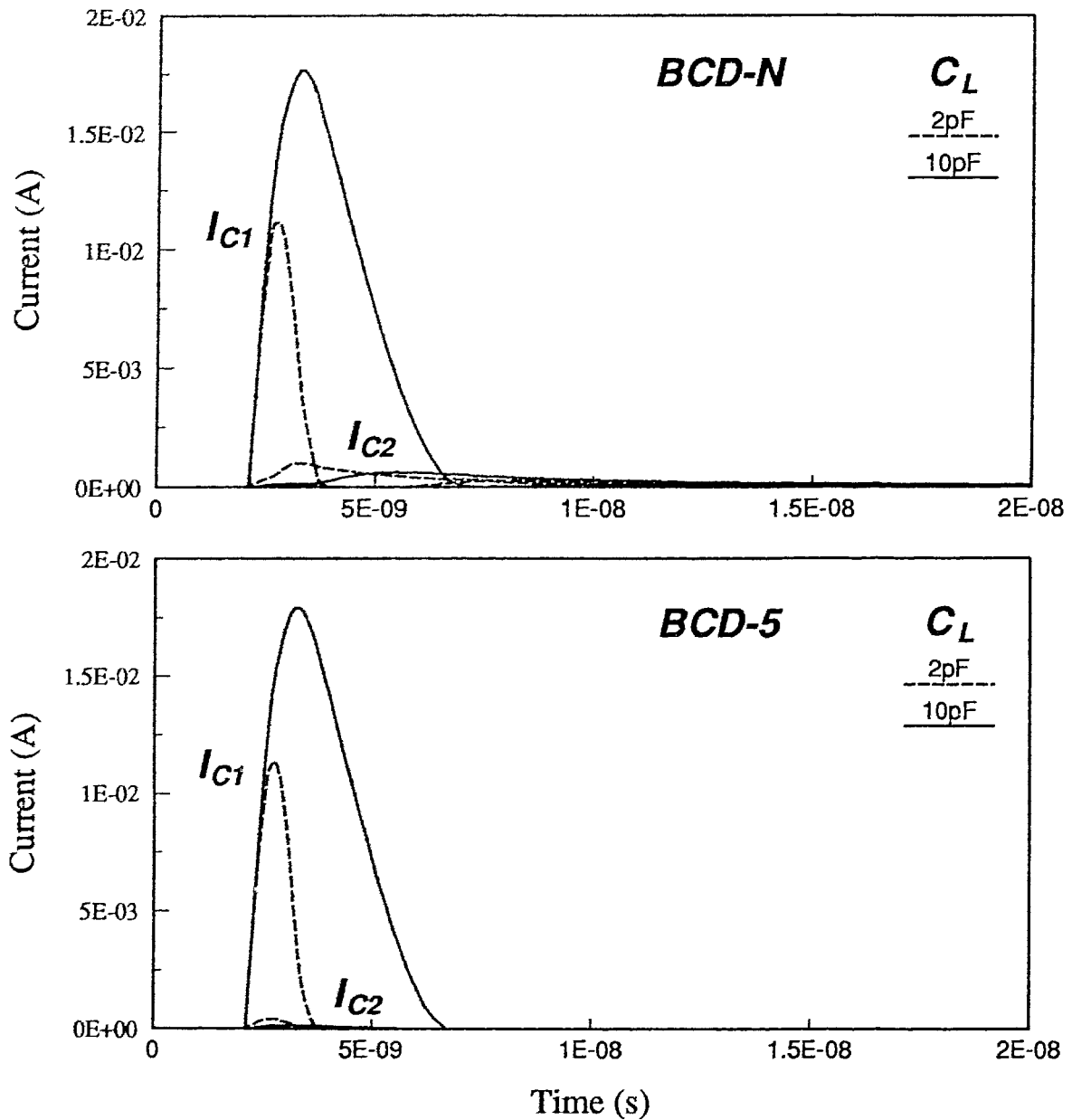
Figure 5.2 The collector currents response of $I_{C1}(t)$ and $I_{C2}(t)$ during the rise time for C_L varied from 2 to 10pF

5.2 Further improvement for BiCMOS analytical delay model

As mentioned in the beginning of the chapter 4, the delay model τ_d determined by

the R_C-limited saturation region of the BJT in BiCMOS driver could be found in [4.1,4.2].

In the model calculations and SPICE simulations in chapter 4, the R_C value was chosen

as 20Ω. This collector resistance is low enough to avoid the BJT entering the saturation

region, because the maximum collector current allowed for this bipolar transistor with

$V_{DD} = 5V$ is about 240mA, which is much larger than the collector current required for

charging the load capacitance of 5-20pF. Of course, too large R_C values would reduce

the maximum allowed collector current significantly, even much lower than the static

collector current $I_{C,(GP)}$ forced by the Knee current I_{KF}. It obviously results in large errors

if the above new delay model equation (4.22) is still used. Figure 5.3 shows the delay

time τ_d versus collector resistance R_C from SPICE simulation results and the new model

calculations. This figure shows that the new model would not be accurate for $R_C >$

~150-200Ω. Therefore an unified analytical delay model which could cover both the

high-level injection effect of the BJT and also the R_C-limited saturation region of the

BJT will be a very interesting topic for further improvement of this simple BiCMOS

delay model.
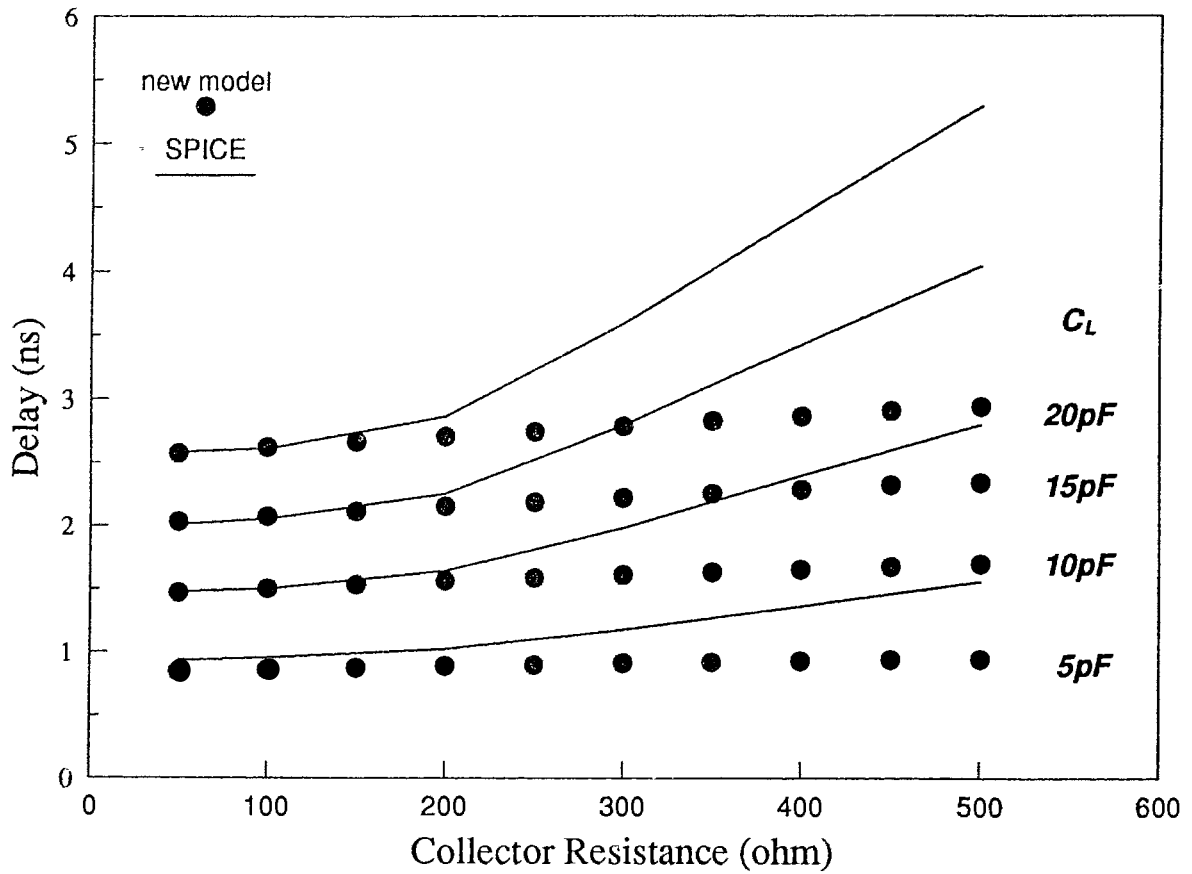
Figure 5.3 Propagation delay τ_d vs. collector resistance R_C of NPN BJT with C_L variations of 5-20pF

CHAPTER 6:

CONCLUSIONS

A new BiCMOS driver circuit (BCD-N) is proposed in this thesis. In order to improve both DC and transient characteristics for BiCMOS driver, this new circuit uses a "dynamic-resistor-behavior" device (depletion mode PMOSFET) to replace the original "on-resistor-behavior" device (enhancement mode NMOSFET). The simulation results show this new BCD-N circuit indeed has the advantages of both rail-to-rail static logic swing from BCD-5 and the faster transient output response from BCD-4. In addition, this new BiCMOS driver shows more compact structure with merged PMOS-NPN devices in BCD-N and larger tolerance for additional process step control. The Table 6.1 below compares the major characteristics for the above three BiCMOS driver circuits.

Detailed DC analysis of the voltage transfer curve (VTC) for this new BiCMOS driver circuit was presented for quantitative evaluation. Due to the highly nonlinear function of the large signal I-V characteristics for both NPN BJT and MOSFET, the piecewise linear function approximation model is used for deriving the slope of $\delta V_{out}/\delta V_{in}$ individually in each segment of VTC. The calculated values from this model are in very good agreement with the simulated results by SPICE.

An improved analytical delay model for the BiCMOS drivers was also developed. By incorporating the Kirk effect directly into collector current transient response derived from the EM BJT model, a simple but accurate propagation delay τ_d model was derived. This new model is particularly suitable for large load capacitors C_L. The effects of those important device parameters like Knee current I_{KF}, forward current gain β_F, the forward transit time τ_F and the charge drain current I_{DS} on the propagation delay τ_d are inherent

in this model, and shown in good agreement with simulated results from SPICE.

Further improvement on the BiCMOS driver circuit design with faster and reliable rising output response, and on the analytical BiCMOS delay model covering also the R_C-limited saturation region of the BJT were suggested for the future research work.

Characteristics:	BCD-4	BCD-5	BCD-N	Comparison (BCD-N/-5)
Static logic swing	lower (~4V)	good (5V)	good (5V)	same
Noise margin NM_L	lower (~0.95V)	low (~1.18V)	good (~1.52V)	~28% improved
Noise margin NM_H	lower (~2.5V)	good (~2.9V)	good (~2.9V)	same
Static power dissipation	no DC current	no DC current	no DC current	same
Rising output response $\tau_{d,r}$ (with $C_L = 10pF$)	good (~1.65ns)	slower (~1.76ns)	slower (~1.79ns)	~1.7% longer
Falling output response $\tau_{d,f}$ (with $C_L = 10pF$)	good (~1.31ns)	slower (~1.55ns)	better (~1.41ns)	~9.0% improved
Transient logic swing (with $C_L = 10pF$)	about the same as in DC case (~3.8V)	about the same as BCD-4 (~3.8V)	a little smaller (~3.4V)	~10% smaller
Process complexity	typical	typical	one more step may be required	one more mask step

Table 6.1: The major characteristics comparisons between BCD-4, BCD-5 and BCD-N

REFERENCES

1.1   A.R. Alvarez, BiCMOS Technology and Applications, Kluwer Academic Publishers, 1989.

1.2   ⁴ "BiCMOS, Is it the Next Technology Driver?", Electronics, February 4, 1988.

1.3   J.E. Buchanan, BiCMOS/CMOS Systems Design, McGraw-Hill, Inc., 1991.

1.4   W. Heimsch, B. Hoffmann, R. Krebs, E. Muellner, B. Pfaeffel, K. Ziemann, "Merged CMOS/Bipolar Current Switch Logic", IEEE ISSCC, pp. 112-113, 1989.

1.5   T.Y. Chiu, G.M. Chin, M.Y. Lau, R.C. Hanson, M.D. Morris, K.F. Lee, "Non-overlapping Super Self-Aligned BiCMOS with 87ps Low Power ECL", IEEE IEDM pp.752-755, 1988.

1.6   H. Klose, T. Meister, B. Hoffmann, B. Pfaffel, P. Weger, I. Maier, "Low Cost and High Performance BiCMOS Processes: A Comparison", IEEE pp.178-181, 1989.

1.7   H. Fukuma, T. Yamauchi, Y. Okajima, "High-Speed BiCMOS Technology with Poly-silicon Emitter Structure", Fujitsu Sci. Tech. J., 24, 4, pp. 384-390, December 1988.

1.8   T. Hotta, T. Bandof, A. Hotta, T. Nakano, S. Iwamoto, S. Adachi, "High-Speed Digital BiCMOS ICs", IEEE ISSCC pp. 124-125, 1989.

1.9   Y. Kobayashi, C. Yamaguchi, Y. Amemiya, and T. Sakai, "High Performance LSI Process Technology: SST CBI-CMOS", IEEE IEDM pp. 760-761, 1988.

1.10   T. Hanibuchi, M. Ueda, K. Higashitani, M. Hatanaka, and K. Mashiko, "A Bipolar-PMOS Merged Base Cell for 0.8μm BiCMOS Sea-of-Gates", IEEE ISSCC pp.4.2.1-4.2.4, 1990.

1.11   M. P. Brassington, M. H. El-diwany, R. R. Razouk, M. E. Thomas, and P. T. Tuntasood, "An Advanced Single-Level Polysilicon Submicrometer BiCMOS Technology", IEEE Transactions on Electron Devices, Vol. 36, No. 4, pp. 712-719, 1988.

1.12   J.L. de Jong, R. Lane, B. Schravendijk, and G. Conner "Single Polysilicon Layer Advanced Super High-speed BiCMOS Technology", IEEE pp. 182-184, 1989.

1.13   V. Ilderem, A. Iranmanesh, A. Solheim, L. Lam, C. Blair, R. Lahri, S. Leibiger, L. Bouknight, M. Biswal, and B. Bastani, "A 0.8μm Advanced Single Poly BiCMOS Technology for High Density and High Performance Applications", IEEE CICC pp. 18.6.1-18.6.4, 1990.

1.14   T. Yuzuriha, T. Yamaguchi, and J. Lee, "Submicron Bipolar-CMOS Technology using $16GH_zf_r$ Double Poly-Si Bipolar Devices", IEEE IEDM pp.748-751, 1988.

1.15   H. Momose, T. Maeda, K. Inoue, I. Kamohara, T. Kobayashi, Y. Urakawa, and K. Maeguchi, "Characterization of Speed and Stability of BiCMOS Gates with A Bipolar and PMOSFET Merged Structure", IEEE IEDM pp. 231-234, 1990.

1.16   T.C. Mele, J. Hayden, F. Walczyk, M. Lien, Y. C. See, D. Denning, S. Cosentino, and A. H. Perera, "A High Performance $0.5\mu m$ BiCMOS Triple Polysilicon Technology for 4Mb FAST SRAMs", IEEE IEDM pp. 18.4.1-18.4.4, 1990.

1.17   N. Rovedo, S. Ogura, J. Acocella, K. Barnes, A. Dally, T. Yanagisawa, C. Ng, J. Burkhardt, E. Valsamakis, J. Hamers, T. Buti, C. Richwine, "Process Design for Merged Complementary BiCMOS", IEEE IEDM pp. 18.5.1-18.5.4, 1990.

1.18   M. El-Diwany, M. Brassington, R. Razouk, P. Wijnen, and V. Akylas, "Low Voltage Performance of An Advanced CMOS/BiCMOS Technology Featuring $18GH_z$ Bipolar fT and Sub-70ps CMOS Gate Delays", IEEE IEDM pp. 18.6.1-18.6.4, 1990.

1.19   S. Konaka, T. Kobayashi, T. Matsuda, M. Ugajin, K. Imai, and T. Sakai, "HSST/BiCMOS Technology with 26ps ECL and 45ps 2V CMOS Inverter", IEEE IEDM pp. 18.7.1-18.7.4, 1990.

1.20   H.V. Tran, D.B. Scott, K. Fung, R. Havemann, RR.E. Eklund, T.E. Ham, R.A. Haken, A.H. Shah, "An 8ns Battery Backup Submicron BiCMOS 256K ECL SRAM," IEEE ISSCC pp. 188-189, 1988.

1.21   M. Matsui, H. Monose, Y. Urakawa, T. Maeda, A. Suzuki, N. Urakawa, K. Sato, K. Makita, J. Matsunaga, K. Ochii, "An 8ns 1Mb ECL BiCMOS SRAM," IEEE ISSCC pp. 38-39, 1989.

1.22   T. Hotta, I. Masuda, H. Maejima, A. Hotta, "CMOS/Bipolar Circuits for 60MHz Digital Processing," IEEE ISSCC pp. 190-191, 1986.

2.1   A.R. Alvarez, BiCMOS Technology and Applications, Kluwer Academic Publishers, 1989.

2.2   G. Zimmer, W. Esser, J. Fichtel, B. Hosticka, A. Rothermel, and W. Schardein, "BiCMOS: technology and circuit design", Microelectronics Journal Vol. 20 Nos. 1-2 c, pp. 59-75, 1989.

2.3   G.P. Rosseel and R.W. Dutton, "Influence of Device Parameters on the Switching Speed of BiCMOS Buffers", IEEE J. Solid-State Circuits, vol 24, no.1, pp. 90-99, February 1989.

2.4   L.W. Nagel, SPICE2: A computer Program to Simulate Semiconductor Circuits, Electronics Research Laboratory Rep. No. ERL-M520, University of California, Berkeley, 1975.

3.1   N.E. Weste and K. Eshraghian, Principles of CMOS VLSI Design, Addison-Wesley Publishing Company, Massachusetts, 1985.

3.2   P. Antognetti and G. Massobrio, Semiconductor Device Modeling with SPICE, McGraw-Hill, Inc., New York, 1988.

3.3   I.E. Getreu, Modeling the Bipolar Transistor, Tektronix, Inc., Beaverton, Ore., 1976.

4.1   E.W. Greeneich and K.L. McLaughlin, "Analysis and Characterization of BiCMOS for High-Speed Digital Logic", IEEE J. Solid-State Circuits, vol 23, no.2, pp. 566-572, April 1988.

4.2   G.P. Rosseel, R.W. Dutton, K. Mayaram and D.O. Pederson, "Delay Analysis for BiCMOS Drivers", IEEE Bipolar Circuits and Technology Meeting, pp. 220-222, 1988.

4.3   A.R. Alvarez, BiCMOS Technology and Applications, Kluwer Academic Publishers, Boston, 1989.

4.4   G.P. Rosseel and R.W. Dutton, "Influence of Device Parameters on the Switching Speed of BiCMOS Buffers", IEEE J. Solid-State Circuits, vol 24, no.1, pp. 90-99, February 1989.

4.5   C.H. Diaz, S.M. Kang and Y. Leblebici, "An Accurate Analytical Delay Model for BiCMOS Driver Circuits", IEEE Trans. Computer-Aided Design, vol 10, no.5, pp. 577-588, May 1991.

4.6   P. Antognetti and G. Massobrio, Semiconductor Device Modeling with SPICE, McGraw-Hill, Inc., New York, 1988.

4.7   I.E. Getreu, Modeling the Bipolar Transistor, Tektronix, Inc., Beaverton, Ore., 1976.

4.8   M.J. Deen and Z.X. Yan "A Simple Propagation Delay Model for BiCMOS Driver Circuits", Solid State Electronics, in press (1991)

4.9   L.W. Nagel, SPICE2: A computer Program to Simulate Semiconductor Circuits, Electronics Research Laboratory Rep. No. ERL-M520, University of California, Berkeley, 1975.