

ASSESSMENT AND MODELING OF NON-QUASI-STATIC,
NON-LOCAL, AND MULTI-DIMENSIONAL EFFECTS IN
ADVANCED BIPOLAR JUNCTION TRANSISTORS

By

JOOHYUN JIN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1992

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Jerry G. Fossum, for giving me an opportunity to work as one of his privileged graduate students on interesting research topics. Without his devoted guidance, encouragement, concern, support and patience, this work could not have reached fruition. My interaction with him has been a most gratifying learning experience.

I also would like to thank the other members of my supervisory committee, Professors Dorothea E. Burk, Mark E. Law, Sheng S. Li, and Timothy J. Anderson, for their willingness to serve on my committee.

I am also indebted to numerous people I have interacted with during my stay in Gainesville. First I am grateful to Mr. D. FitzPatrick for his help in the MMSPICE software development. Thanks are also extended to many of my colleagues who helped me through technical discussions or by cheering me up in difficult times. I cannot mention all of them, but I should mention Drs. H. Jeong, Y. Kim, J. Choi, and Messrs. H. J. Cho, S. Lee, H. S. Cho, G. Hong, K. Green, D. Suh, P. Yeh, M. Liang, D. Apte, S. Krishnan. My deepest gratitude goes to my parents and sisters Hyesook and Minjung

for their endless love and encouragement throughout the years of my graduate study. Last but not least, I thank the Lord for His guidance in my life. I also acknowledge the financial support of the Semiconductor Research Corporation and Samsung Semiconductor & Telecommunication Co. Ltd.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	vi
CHAPTERS	
1 INTRODUCTION.....	1
2 MODELING OF MULTI-DIMENSIONAL CURRENTS.....	9
2.1 Introduction.....	9
2.2 Model Development.....	11
2.2.1 Experimental Characterization.....	11
2.2.2 Analytic Model.....	15
2.3 Simulations and Verification.....	21
2.4 Summary.....	27
3 NON-QUASI-STATIC MODELING OF BJT CURRENT CROWDING.....	29
3.1 Introduction.....	29
3.2 Model Development.....	31
3.2.1 Switch-on Case.....	37
3.2.2 Switch-off Case.....	39
3.3 NQS Model Implementation.....	42
3.4 Simulations.....	45
3.5 Summary.....	60
4 ANALYTIC ACCOUNTING FOR CARRIER VELOCITY OVERSHOOT....	64
4.1 Introduction.....	64
4.2 Model Development.....	67
4.2.1 Velocity Overshoot.....	67
4.2.2 Velocity Relaxation.....	73
4.2.3 Effective Saturated Drift Velocity.....	77
4.2.3.1 Junction SCR.....	79
4.2.3.2 Current-induced SCR.....	81
4.2.3.3 Special case.....	82
4.3 Comparisons with Energy Transport Model.....	84
4.4 Implementation.....	91
4.5 Simulations.....	93

4.6 Summary.....	109
5 MMSPICE-2 DEVELOPMENT.....	111
5.1 Introduction.....	111
5.2 New Features.....	112
5.2.1 Multi-dimensional Currents.....	112
5.2.2 Current Crowding.....	113
5.2.3 Velocity Overshoot.....	115
5.2.4 Extrinsic Collector-base Capacitance.....	115
5.2.5 Substrate Capacitance.....	119
5.3 Parameter Evaluation.....	119
5.4 Model Implementation.....	122
5.4.1 Subroutine Modifications.....	122
5.4.1.1 Subroutine MODCHK.....	122
5.4.1.2 Subroutine QBBJT.....	123
5.4.1.3 Subroutine QBCT.....	127
5.4.2 Subroutine Additions.....	127
5.4.2.1 Subroutine CROWD.....	127
5.4.2.2 Subroutine OVERSHOOT.....	129
5.5 Demonstration.....	131
5.6 Summary.....	144
6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK.....	145
APPENDICES	
A EVALUATION OF J_{SEO} , n_{EB} , J_{EOP} AND n_{EBP}	148
B DISCUSSION ON J_Q	150
C LIMITING $J_{EO(eff)}$ IN THE SWITCH-OFF SIMULATION.....	152
D VALIDITY OF THE DEPLETION APPROXIMATION.....	154
REFERENCES.....	156
BIOGRAPHICAL SKETCH.....	161

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ASSESSMENT AND MODELING OF NON-QUASI-STATIC,
NON-LOCAL, AND MULTI-DIMENSIONAL EFFECTS IN
ADVANCED BIPOLAR JUNCTION TRANSISTORS

By

JOOHYUN JIN

August 1992

Chairman: Dr. J. G. Fossum
Major Department: Electrical Engineering

This dissertation is concerned with assessment, modeling, and simulation of non-quasi-static (NQS), non-local, and multi-dimensional effects in advanced bipolar junction transistors. A simple analytic model for the sidewall injection of the base current, which is shown to be the most important multi-dimensional component in scaled devices, is developed based on the separation of the base current into internal and peripheral components. Simulation results for typical test BJTs with various emitter geometries are compared against corresponding measurements to support the model. A novel NQS model for transient current crowding in advanced BJTs is developed for circuit simulation. The new model, implemented based on a novel use of the previous time-step solution in the current time-step analysis,

characterizes a time-dependent effective bias on the emitter-base junction in a semi-numerical analysis, accounting for base conductivity modulation and the NQS nature of the crowding. The (dc) debiasing effect, which is important in analog circuits, is inherently accounted for as well. An analytic model for electron velocity overshoot resulting from non-local transport in advanced silicon-based BJTs is developed. The model, which characterizes an effective saturated drift velocity, larger than the classical value because of overshoot, is intended for circuit simulation. The model uses an augmented drift-velocity formalism that involves a length coefficient derived via Monte Carlo analysis. The associated velocity relaxation is characterized phenomenologically to be consistent with overshoot analysis. The developed charge-based models are implemented in MMSPICE-2, a semi-numerical mixed-mode device/circuit simulator, such that users may activate any combination of the new features by option. The resulting hierarchical tool, along with the parasitic charge (capacitance) models included to enhance the usefulness of the simulator, could indeed enable predictive yet computationally efficient mixed-mode simulations for bipolar (and BiCMOS) VLSI technology/manufacturing CAD. Utility of MMSPICE-2 is demonstrated by transient simulations of ECL circuits and devices.

CHAPTER 1 INTRODUCTION

In recent years, advances in process technology have led to the realization of high-performance bipolar junction transistors (BJTs). While continual improvement in the lithographic capability allows the lateral dimensions to be reduced, scaling down the BJT requires a coordinated change in both the lateral dimensions and vertical profile to achieve proper device operation and to improve the intrinsic device speed. Furthermore, in order to reduce the extrinsic portion of the bipolar device so that circuit performance can be more closely tied to the intrinsic device performance, various self-alignment schemes using polysilicon as base and emitter contacts have been developed. They all have a similar structure (see Fig. 1.1), and generally provide much improved performance over the conventional BJT structure via a reduction in base-collector junction area and base resistance.

Despite the impressive progress made in bipolar technology, computer simulation tools, which are essential to the optimization of device and circuit designs for the technology, have not kept pace with it. In integrated

circuit development and manufacturing today, a technology CAD (TCAD) system is essential for exploring alternative designs and evaluating various trade-offs without time-consuming and costly fabrications.

An effective TCAD system requires integrated, physics-based tools for predictive process, device, and (small-scale) circuit simulation. Computational efficiency is desirable and indeed essential if the TCAD system is to be used in manufacturing CAD involving statistical simulation. Conventional TCAD systems comprise robust, numerical process and device simulators which drive optimization of empirical device model parameters for circuit simulation. This optimization can miss parametric correlations, and hence the integrated system, although CPU-intensive, could yield nonunique (erroneous) predictions.

Numerical mixed-mode device/circuit simulation would obviate this deficiency, but with a high cost of computation time. Alternatively, improvement of the TCAD system can possibly be afforded by incorporation of semi-numerical device models into the circuit simulator which have physical parameters that relate directly to the device structure. The resulting tool is an application-specific, computationally efficient mixed-mode simulator that can easily be integrated with the process simulator by a program that evaluates the model parameters from the doping profile. The MMSPIICE

[Jeo90] is such a simulator, which is integrated with SUPREM [SUP88] by a parameter-extraction program, SUMM [Gre90].

The model development for MMSPICE has emphasized the advanced BJTs. A physical, one-dimensional charge-based model [Jeo89] has been developed and implemented. High-current effects, impact ionization, and non-reciprocal (trans)capacitances are physically accounted for in the semi-numerical model. This model is sufficient for many applications, but more work is needed to enhance the usefulness of MMSPICE.

In most advanced BJTs, the lateral dimension of the emitter has become the same order of magnitude as the emitter-base junction depth. Thus, multi-dimensional current effects in the peripheral region of the junction are expected to play a significant role in device performance. Especially, the variation in common-emitter forward current gain β with geometric shape and size is troublesome to IC designers [Hwa87]. Hence, some accounting of peripheral currents is needed for circuit simulation.

High-current effects (e.g., quasi-saturation and base widening, or pushout) are physically accounted for in the MMSPICE model, but emitter current crowding, caused by lateral voltage drops in the intrinsic base region, has not yet been considered. Today's advanced (scaled) BJTs commonly operate at high current density, and hence transient base

current can be much greater than the steady-state current; this clearly implies the non-quasi-static (NQS) nature of transient current crowding [Ham88]. Therefore, it can be significant even though dc crowding may be insignificant [Tan85].

In semiconductor devices where the electric field increases rapidly over distances comparable to the energy-relaxation mean free path, carrier velocity can overshoot the value corresponding to the local field because the carrier (kinetic) energy, which controls the collision time and hence limits the velocity, lags the field and remains relatively small [Ruc72]. This non-local effect on electron transport has been recognized as significant in MOSFETs and MESFETs for years, and now has become important in scaled BJT's [Lee89, Cra90]. Recent work [Fus92] has indicated that the velocity overshoot in scaled BJT's can be beneficial, and must be accounted for in the device and circuit design. However, the effect has not yet been accounted for in any existing circuit simulators, and indeed is missing in many device simulators because of the implied computational intensiveness.

For bipolar integrated circuits, reducing parasitic capacitances is one of the key issues for speed enhancements. The extrinsic collector-base junction capacitance (charge) has a predominant effect on the circuit performance because the extrinsic base region is not reduced in proportion as the

intrinsic device is scaled down. The collector-substrate capacitance (charge) is also important.

This dissertation addresses these problems; it is concerned with the development and implementation of new models to account for the aforementioned effects in the advanced BJTs. This work will enable not only truly predictive, scalable BJT simulations, but also computationally efficient (semi-numerical) mixed-mode device/circuit simulations for bipolar TCAD. The major contributions made in this work are as follows:

- (1) modeling of multi-dimensional current effects, based on the separation of the current into internal and peripheral components;
- (2) development of an NQS transient current-crowding model, based on a novel use of the previous time-step solution in the current time-step analysis;
- (3) development of an analytic model for electron velocity overshoot resulting from non-local transport in advanced silicon-based BJTs;
- (4) implementation of the new models, including both the extrinsic collector-base and collector-substrate capacitances (charges), in MMSPICE to create MMSPICE-2.

In Chapter 2, a simple analytic way of accounting for multi-dimensional current effects is described. The approach is based on the separation of the current into areal and

peripheral components. For high V_{BE} , an effective junction bias (described in Chapter 3) is necessarily defined to account for the emitter debiasing (a.k.a. crowding) effect. The model is supported by experimental results of test BJTs having varied emitter geometries.

In Chapter 3, a new NQS model for transient current crowding is presented. The model, which characterizes a time-dependent effective bias on the emitter-base junction in a semi-numerical analysis, follows the previous work by Hauser [Hau64], but physically accounts for base conductivity modulation and the NQS nature of the crowding. The novel modeling/implementation is based on the use of the previous time-step solution in the current time-step analysis, which in fact could enable general accounting of NQS effects in semi-numerical mixed-mode device/circuit simulation. The tool is supported by numerical simulations of advanced BJTs using PISCES [PIS84].

In Chapter 4, an analytic model for electron velocity overshoot in advanced BJTs is presented. The model, which characterizes an effective saturated drift velocity in the collector space-charge regions, is intended for circuit simulation. The model uses an augmented drift-velocity formalism that involves a length coefficient derived from Monte Carlo simulations. The associated relaxation of the carrier velocity is characterized phenomenologically to be

consistent with the overshoot analysis. Demonstrative simulation results are presented to assess the significance of the electron velocity overshoot in advanced bipolar and BiCMOS technologies, and to support model.

The developed charge-based models are implemented into MMSPICE-2 so that users may activate any combination of the new features by option. This hierarchical tool is discussed in Chapter 5. Representative simulations are presented, with descriptions of the new parameters.

In Chapter 6, the main accomplishments of this dissertation are summarized, and future research areas are suggested.

CHAPTER 2
MODELING OF MULTI-DIMENSIONAL CURRENTS

2.1 Introduction

For bipolar integrated circuits, reducing parasitic effects and achieving shallow profiles are two of the key issues in improving performance. Many self-aligned bipolar technologies have been developed to achieve low parasitic capacitance and low base resistivity. They all have a similar device structure using polysilicon as base and emitter contacts. In the scaled structure, the distance between base and emitter contacts is greatly reduced as determined by the boot-shaped sidewall spacer (see Fig. 1.1). The lateral dimensions of the device have also been scaled down; for example, the emitter width of today's most advanced transistors has become the same order of magnitude as the emitter-base junction depth. Thus, multi-dimensional effects in the peripheral region of the junction can play a significant role in device performance [Hur87].

For digital applications, a most predominant multi-dimensional effect is the lateral injection of significant base current along the emitter sidewall, which is controlled by the morphology of the link region [Li88]. One simple way

to reduce this sidewall current component is to increase the width of the spacer [Dej88, Saw88]. However, many desirable features of the device depend on the limitation of the spacer width. For example, as the spacer width increases, the base resistance and parasitic capacitances increase. Also, the emitter-collector punchthrough current increases due to insufficient extrinsic-intrinsic base overlap in the emitter periphery [Chu87, Saw88], while an increase in the extrinsic-intrinsic base overlap results in excessive perimeter tunneling current [Sto83] and hence reduced emitter-base breakdown voltage. Thus, the control of spacer thickness is vital to the performance of the device.

The peripheral component of the base current does not modulate the collector current, and is therefore a parasitic that degrades the dc current gain β in proportion to the ratio of its magnitude relative to that of the areal component. Hence, β is degraded more as the perimeter-to-area ratio (P_E/A_E) increases. This implies that the sidewall effect can be an obstacle for down-scaling the emitter size [Hwa87, Dej88]. Therefore, some accounting of peripheral currents for a given process is needed for a circuit simulator, e.g., MMSPICE, which actually gives an extra degree of freedom to the IC designer [Ver87].

In Section 2.2, a simple model based on measurements is presented to account for the peripheral currents in the

advanced BJT structure. This model, combined with the current-crowding analysis described in Chapter 3, will be the basis for a more predictive and scalable BJT model for MMSPICE. In Section 2.3, experimental results of test BJTs having varied emitter geometries are presented to support our formalism. In fact, interpretation of these results requires the crowding model of Chapter 3, which was hence developed in conjunction with the work described in this chapter.

2.2 Model Development

2.2.1 Experimental Characterization

For digital applications, the most important peripheral current is the sidewall component of the base current. However, the peripheral component of the collector current is not significant compared with the areal component, provided the extrinsic base is well-linked with the intrinsic base [Li88].

This fact is also supported by our own measurements of representative (advanced) BJTs provided by Dr. D. Verret of Texas Instruments. The lateral geometries of the test devices are described in Table 2.1; L_E and W_E are the effective (or actual) length and width of the emitter, and P_E ($=2L_E+2W_E$) and A_E ($=L_EW_E$) are the perimeter and area

TABLE 2.1
LATERAL EMITTER GEOMETRIES OF TEST DEVICES

L_E [μm]	W_E [μm]	P_E/A_E [μm^{-1}]
9.2	5.2	0.60
9.2	4.2	0.69
9.2	3.2	0.84
9.2	1.7	1.39
9.2	1.2	1.88
9.2	0.7	3.08
9.2	0.45	4.66

respectively. The spacer width of these devices is estimated to be $0.4\mu\text{m}$. Fig. 2.1 shows the base (J_B) and collector (J_C) current densities versus P_E/A_E for the devices with L_E fixed at $L_E=9.2\mu\text{m}$ when $V_{BE}=0.4$ or 0.7V . Since J_C is almost constant regardless of P_E/A_E as well as V_{BE} and V_{BC} , we infer that the peripheral collector current can be neglected at least for relatively low V_{BE} . On the contrary, J_B clearly increases with P_E/A_E , obviously implying a significant lateral-injection component. We note that this parasitic effect becomes more significant as V_{BE} is reduced, which we believe reveals that the peripheral base current is due to the recombination of excess carriers in the peripheral junction space-charge-region (SCR) near or at the oxide-silicon interface.

The lateral injection can be understood better if the peripheral component of base current is quantified. Empirically, the total base current I_B can be separated into areal and peripheral parts as follows [Rei84]:

$$\begin{aligned}
 I_B &= I_{BA} + I_{BP} \\
 &= C_A A_E \left[\exp\left(\frac{V_{BE}}{n_A V_T}\right) - 1 \right] + C_P P_E \left[\exp\left(\frac{V_{BE}}{n_P V_T}\right) - 1 \right]
 \end{aligned} \tag{2.1}$$

where C_A , n_A , C_P , and n_P are (process-dependent) empirical constants, which can easily be evaluated using the basic experimental method discussed in Appendix A. In (2.1), the

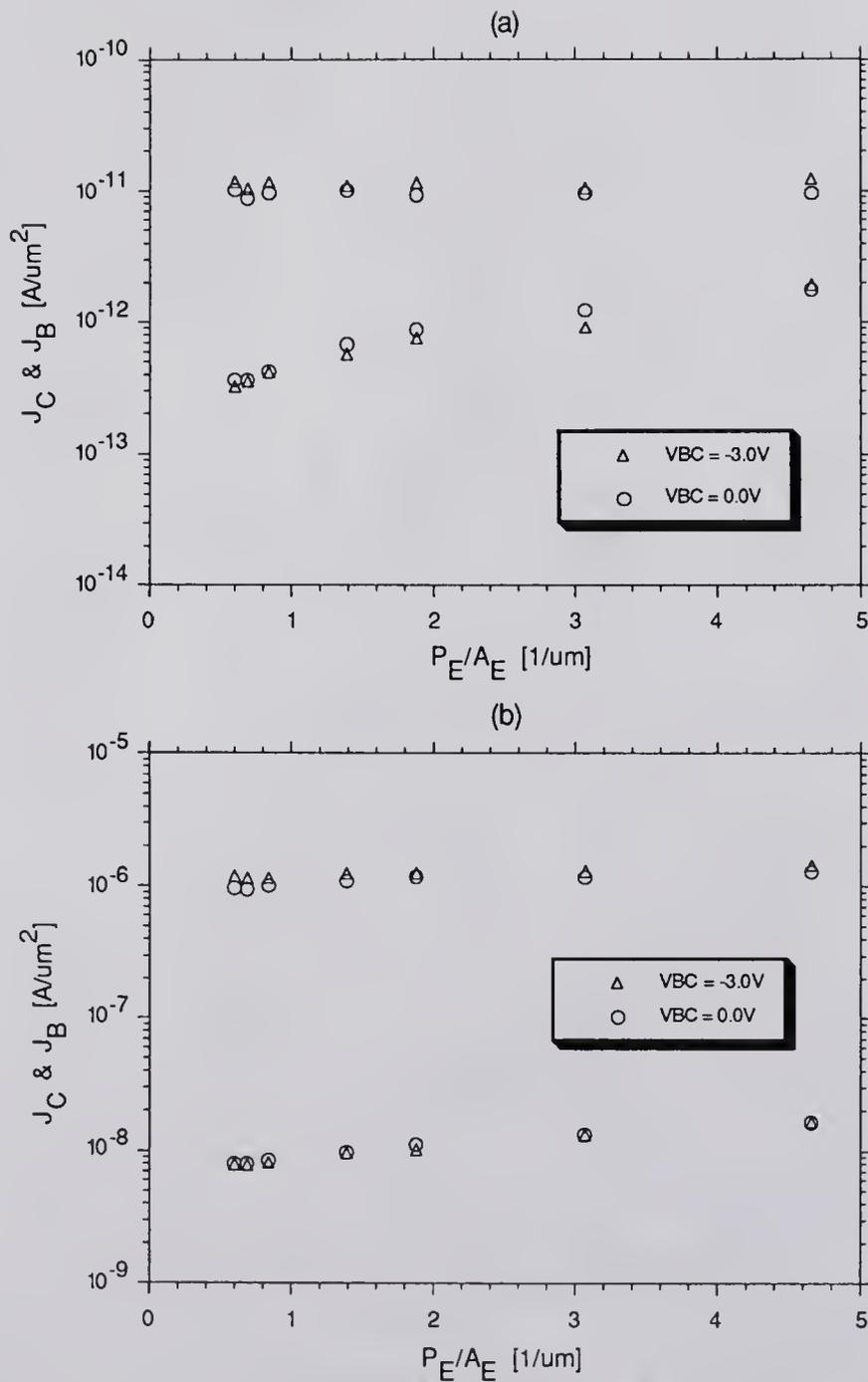


Fig. 2.1 Base and collector current densities versus P_E/A_E for devices with $L_E = 9.2\mu\text{m}$: (a) $V_{BE} = 0.4\text{V}$; (b) $V_{BE} = 0.7\text{V}$.

voltage drop across the extrinsic base resistance is neglected for low-current conditions.

Based on this formalism, it is possible to calculate the contribution of the peripheral current to the total base current. Doing this for the devices previously characterized yields in Fig. 2.2 I_{BP}/I_B versus P_E/A_E for $V_{BC}=0.0V$. As discussed before, the peripheral base component increases with P_E/A_E . For example, when $V_{BE}=0.7V$ and $P_E/A_E=0.60/\mu m$ (actually, this is equivalent to the device with $W_E=5.2\mu m$), I_{BP} is only 16% of the total base current, but it increases to 50% when $P_E/A_E=3.1/\mu m$ (i.e., $W_E=0.7\mu m$). For reduced V_{BE} , the effect of lateral injection becomes more significant in accord with our previously stated recognition; when $V_{BE}=0.4V$, the mentioned ratios are changed to 47% and 82% respectively. Our other simulations and measurements show that the peripheral collector current evaluated via this methodology is about 10% of the total collector current on the average.

2.2.2 Analytic Model

With this insight, we can extend the MMSPICE BJT model to account for the peripheral base region, at least to first order. The extended model is restricted to include only the lateral injection of the base current, which has been shown to be the most important multi-dimensional effect in modeling

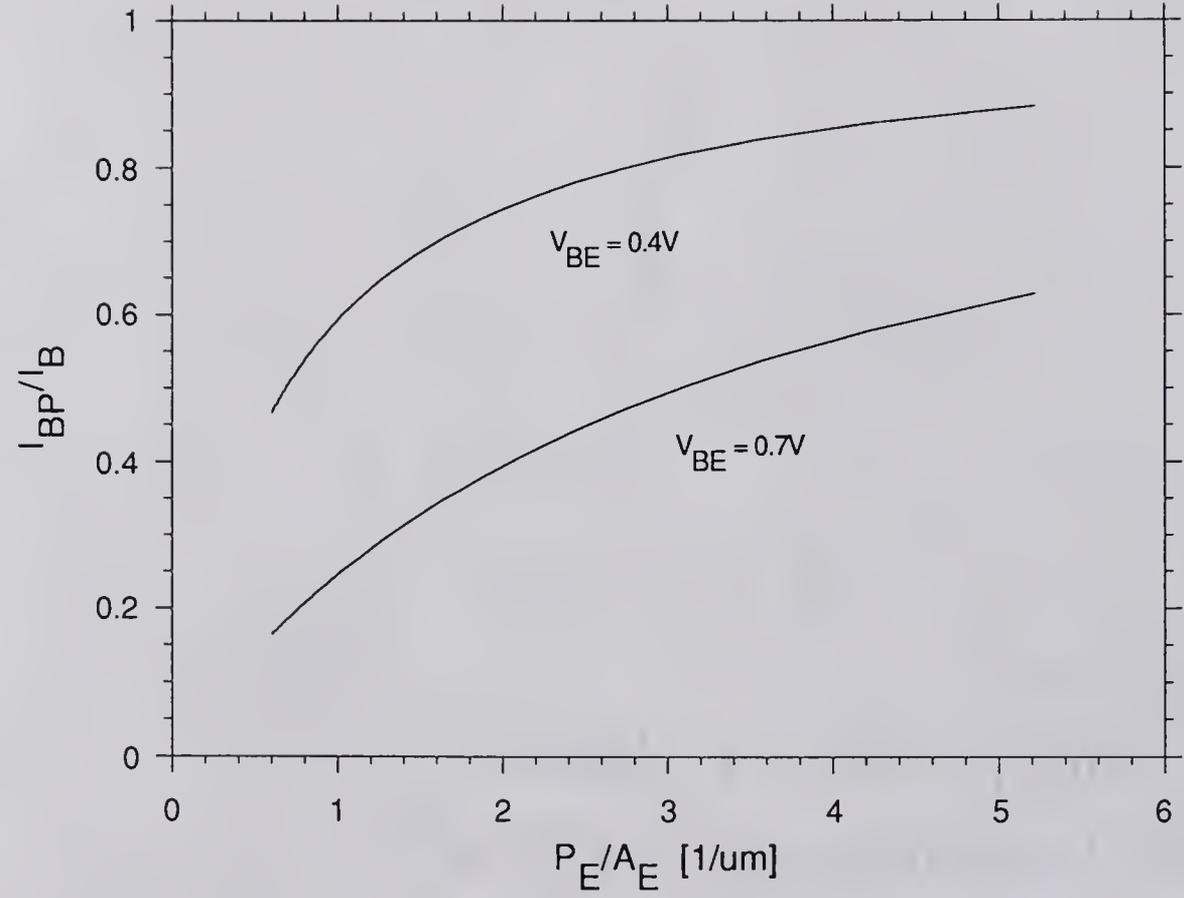


Fig. 2.2 Simulated I_{BP}/I_B versus P_E/A_E for the devices used in Fig. 2.1.

advanced BJTs. Based on the insight derived from the measurements, we add only a peripheral component of base current to the existing BJT routine in MMSPICE. This additional component is proportional to the emitter perimeter P_E , and represents peripheral SCR recombination near the surface. The peripheral base current I_{BP} can be expressed as

$$I_{BP} = J_{EOP} P_E \left[\exp\left(\frac{V_{BE}}{n_{EBP} V_T}\right) - 1 \right] \quad (2.2)$$

where J_{EOP} and n_{EBP} represent the peripheral saturation current density (per unit length) and the peripheral emission coefficient respectively. The sidewall injection effect could also be dependent on the emitter junction depth, but we assume that this dependence is implicitly included in the above formalism.

In a dc case, the predominant components of the areal base current are typically back-injection current from the base to the emitter and the recombination current at the (emitter-base) junction SCR. (Recombination in the quasi-neutral base and the epi collector is neglected here since it is typically insignificant in advanced BJTs.) Hence, the total base current I_B can be expressed as

$$I_B = I_{BA} + I_{BP}$$

$$\begin{aligned}
&= J_{EO}A_E \left[\exp\left(\frac{V_{BE}}{V_T}\right) - 1 \right] + J_{SEO}A_E \left[\exp\left(\frac{V_{BE}}{n_{EB}V_T}\right) - 1 \right] \\
&+ J_{EOP}P_E \left[\exp\left(\frac{V_{BE}}{n_{EBP}V_T}\right) - 1 \right] \tag{2.3}
\end{aligned}$$

where J_{EO} is the (areal) emitter saturation current density, and J_{SEO} and n_{EB} are (areal) SCR saturation current density and SCR emission coefficient respectively.

Although (2.3) is sufficient for many operating ranges, it is necessary to examine whether it is valid for high-current operation where additional effects are significant. In this case the actual (peripheral) junction bias V'_{BE} cannot be approximated as the terminal voltage V_{BE} ; V'_{BE} is considerably less than V_{BE} since the voltage drops across the extrinsic base and emitter resistances are no longer negligible. Furthermore, the areal component is degraded by the lateral voltage drops in the intrinsic base region. In fact, interpretation of data necessitated the current-crowding modeling described in Chapter 3. Hence we modify (2.3):

$$\begin{aligned}
I_B &= J_{EO}A_E \left[\exp\left(\frac{V_{BE(eff)}}{V_T}\right) - 1 \right] + J_{SEO}A_E \left[\exp\left(\frac{V_{BE(eff)}}{n_{EB}V_T}\right) - 1 \right] \\
&+ J_{EOP}P_E \left[\exp\left(\frac{V'_{BE}}{n_{EBP}V_T}\right) - 1 \right] \tag{2.4}
\end{aligned}$$

where $V_{BE(\text{eff})}$ is defined (in Chapter 3) as the effective bias on the emitter-base junction to account for the debiasing (a.k.a. current crowding) in terms of the actual (peripheral) bias V'_{BE} . Note that in (2.4), the peripheral current term is not threatened by the current crowding because the peripheral junction voltage is always fixed at V'_{BE} . Although the debiasing effect was classically characterized by Hauser [Hau64], his treatment is inadequate for advanced BJTs because it neglects conductivity modulation of the base. On the contrary, the concept of the effective bias can account for the high-current effects via the charge-based BJT model [Jeo89]. When the debiasing effect is significant, the effective bias is of course less than the actual junction bias V'_{BE} . (In this case, V'_{BE} is also significantly less than V_{BE} .) Otherwise, $V_{BE(\text{eff})}$ would be almost the same as V'_{BE} . This effective bias is derived from the quasi-three-dimensional crowding analysis, which involves a coupling of the vertical and lateral carrier-transport analyses in the base region. Details are described in Chapter 3.

Fig. 2.3 illustrates $(V'_{BE} - V_{BE(\text{eff})})/V_T$ versus W_E predicted by the debiasing analysis for typical advanced devices with $L_E = 9.2 \mu\text{m}$. When $V_{BE} = 0.7\text{V}$, the debiasing effect is, as expected, negligible resulting in $V_{BE(\text{eff})} \approx V'_{BE} \approx V_{BE}$ regardless of W_E and V_{BC} . However it becomes noticeable for higher V_{BE} and especially for greater W_E , due to the increased

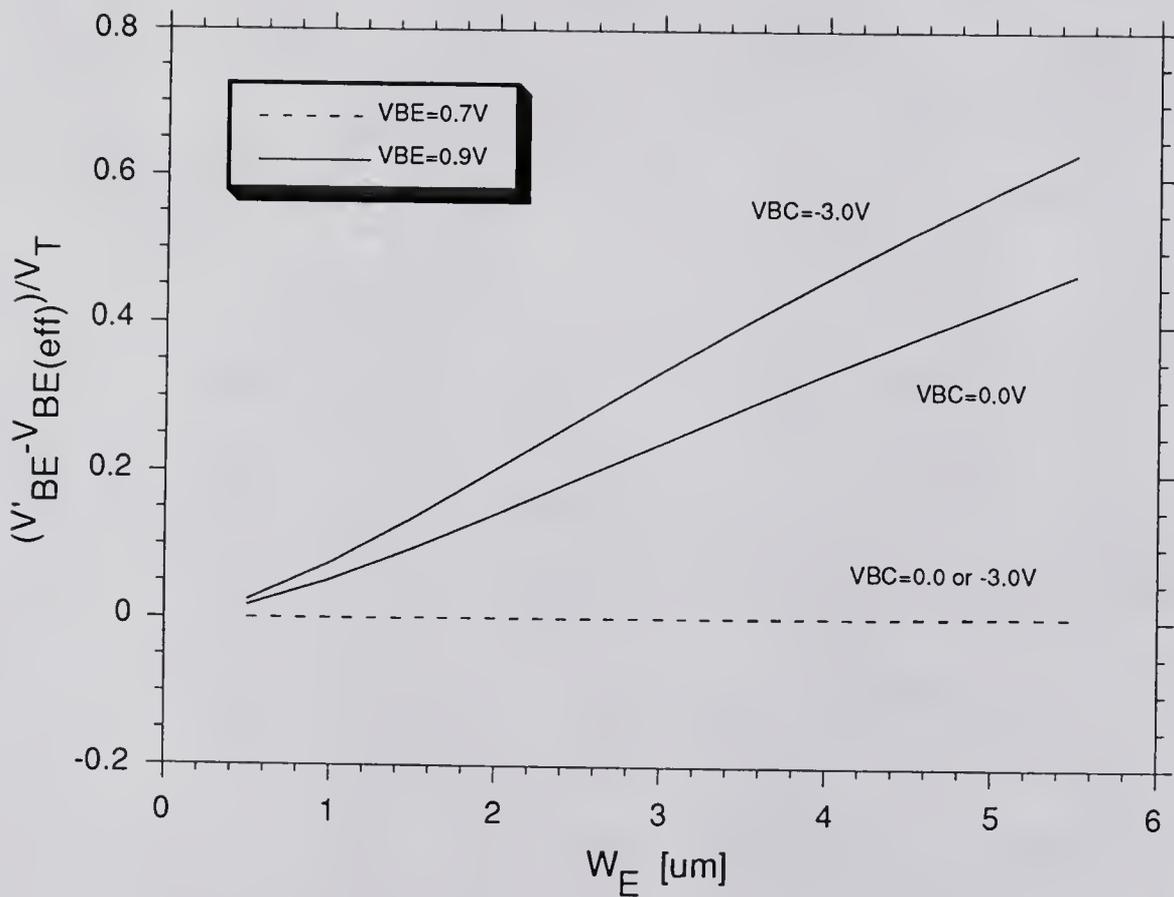


Fig. 2.3 Simulated $(V'_{BE} - V_{BE(eff)})/V_T$ versus W_E for typical advanced BJTs with $L_E = 9.2\mu\text{m}$.

voltage drops in the intrinsic base region. The debiasing effect also becomes more important with increasing reverse bias on the base-collector junction because the base resistivity increases correspondingly. For contemporary scaled BJTs however, it is not significant [Tan85]; for $W_E=2\mu\text{m}$ at $V_{BE}=0.9\text{V}$ and $V_{BC}=-3.0\text{V}$, the voltage difference between the actual and effective bias is about 20% of the thermal voltage.

2.3 Simulations and Verification

The test devices, representative of the advanced bipolar technology, were used to verify the model. The devices, from Texas Instruments, were fabricated using a double-polysilicon process in conjunction with a sidewall spacer technique, which enables a self-aligned submicrometer emitter structure. In order to identify significant multi-dimensional effects, transistors with different P_E/A_E (see Table 2.1) were measured.

Simulations were done with MMSPICE-2, which includes the peripheral base current [eq. (2.2)] and the current-crowding model as described in Chapter 3. At first, the model parameters associated with the lateral injection were extracted as described in Appendix A. Then, with no additional parameter extraction, all BJTs were simulated with

reasonably good accuracy simply by scaling A_E .

Simulated I_C/W_E and I_B/W_E compare quite well with the corresponding measurements in Fig. 2.4(a) when $V_{BE}=0.4V$ and $V_{BC}=0.0V$. Note that the lateral injection effect on the base current becomes significant as W_E is scaled down; I_B/W_E increases because the ratio of the peripheral to the areal component increases. However, the contribution of the peripheral collector current is negligible for each device. Note that if I_{BP} had not been accounted for, I_B/W_E would have been predicted to be a constant, since the voltage drops across the extrinsic resistances are negligible for each device at this bias point. For the corresponding β shown in Fig. 2.4(b), the simulations are excellent. As expected, β is reduced with decreasing W_E . Although β -degradation is an obstacle for down-scaling W_E , we expect that our first-order accounting of the lateral injection could give an extra degree of freedom to the circuit designer.

The peripheral collector current is still negligible when V_{BE} is increased to 0.7V, as shown in Fig. 2.5(a). Still, the sidewall injection of the base current, although not as significant as in the low-current region, is important especially for devices with small W_E . The simulations are good, although there is a small discrepancy between the measured and predicted values of I_B/W_E for submicron devices. Indeed this discrepancy seems to be inevitable because the

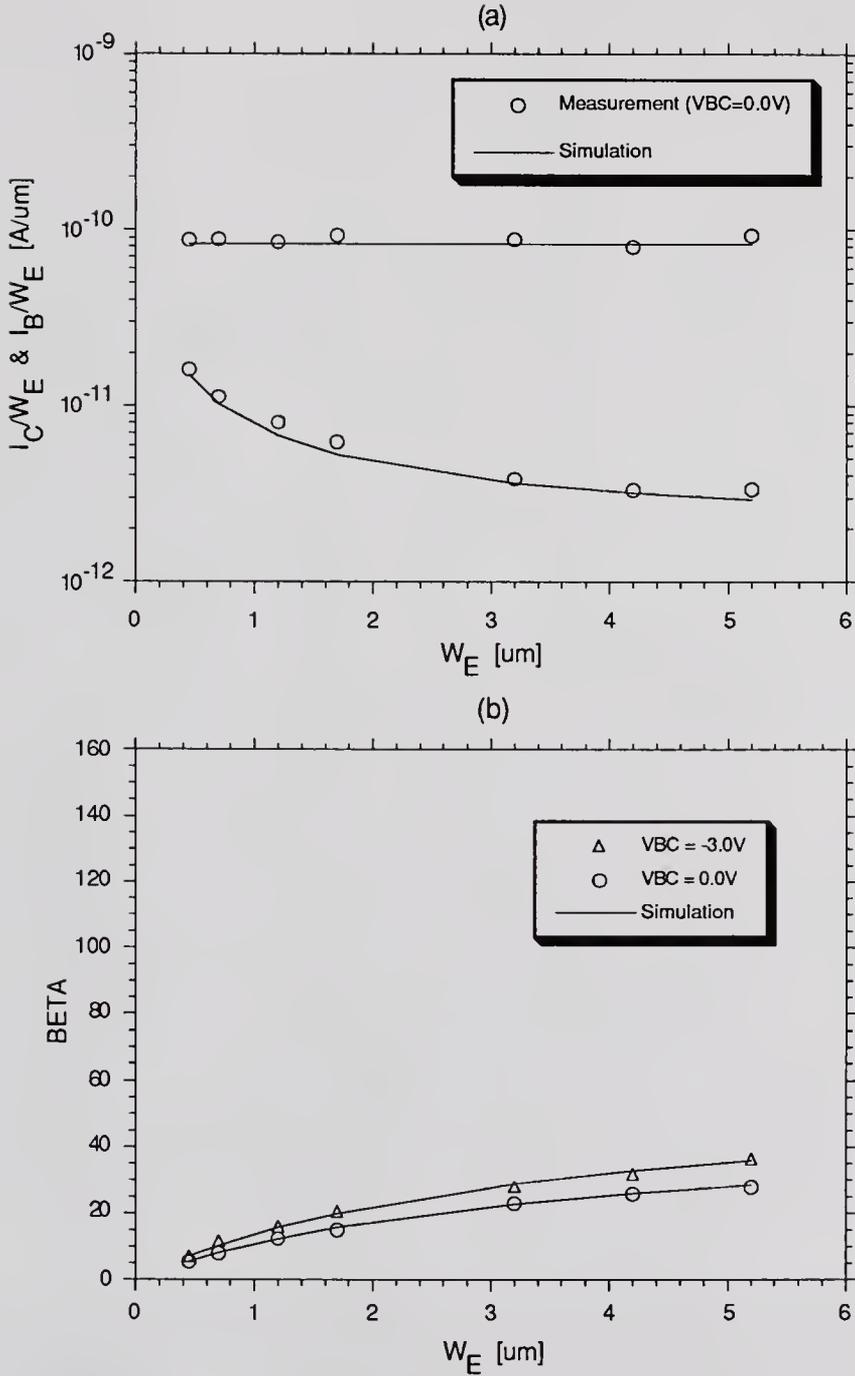


Fig. 2.4 Measured and simulated I_C/W_E , I_B/W_E in (a) and β in (b) for the test BJTs with $L_E = 9.2 \mu\text{m}$ for $V_{BE} = 0.4\text{V}$.

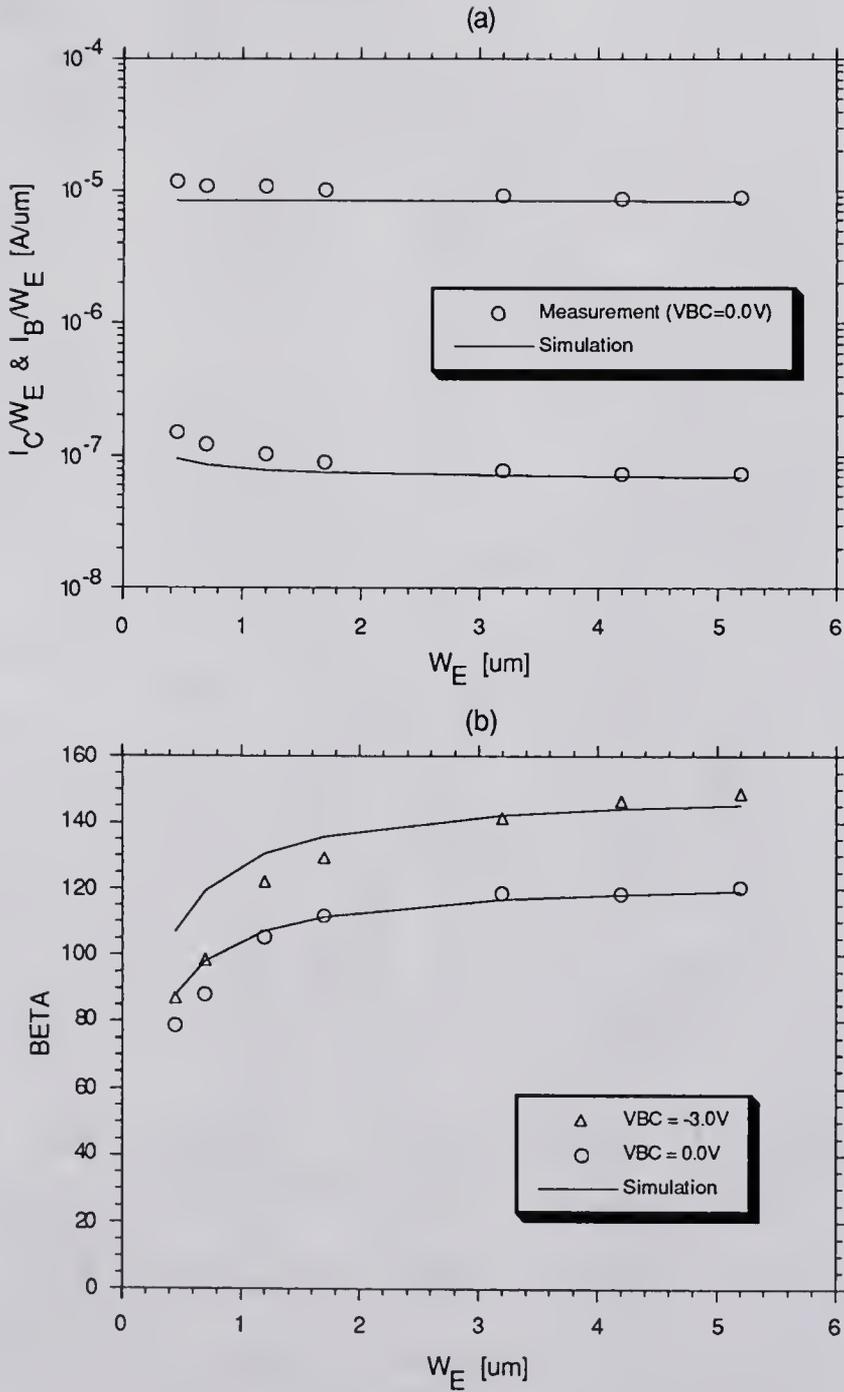


Fig. 2.5 Measured and simulated I_C/W_E , I_B/W_E in (a) and β in (b) for the test BJTs with $L_E = 9.2 \mu\text{m}$ for $V_{BE} = 0.7\text{V}$.

I_{BP} -related model parameters were evaluated from the devices operating in the low current region; according to (2.3), the P_E -dependent term would become negligible with increasing V_{BE} . However, our model seems adequate, as implied by the corresponding β results in Fig. 2.5(b).

For $V_{BE}=0.9V$ in Fig. 2.6(a), the simulations are also reasonably good. We note that I_C/W_E and I_B/W_E decrease with increasing W_E , not because the lateral injection becomes less significant as in Figs. 2.4 and 2.5, but because both the debiasing of the internal junction and high-current-induced voltage drops across the extrinsic resistances, including base resistance, increase with W_E . From the figure however, we can infer that the voltage drops, which become greater for large devices due to the increased terminal currents, are most dominant. The effect of current crowding on β is well illustrated in Fig. 2.6(b); of course, the better simulations obtain with debiasing accounted for. However the debiasing seems to be insignificant for contemporary scaled devices, as discussed before. Our other simulations show that for devices with $W_E > L_E$, the debiasing effect is almost the same for each device, since the predominant base current flow under the rectangular emitter is laterally along the shorter emitter dimension (L_E in this case).

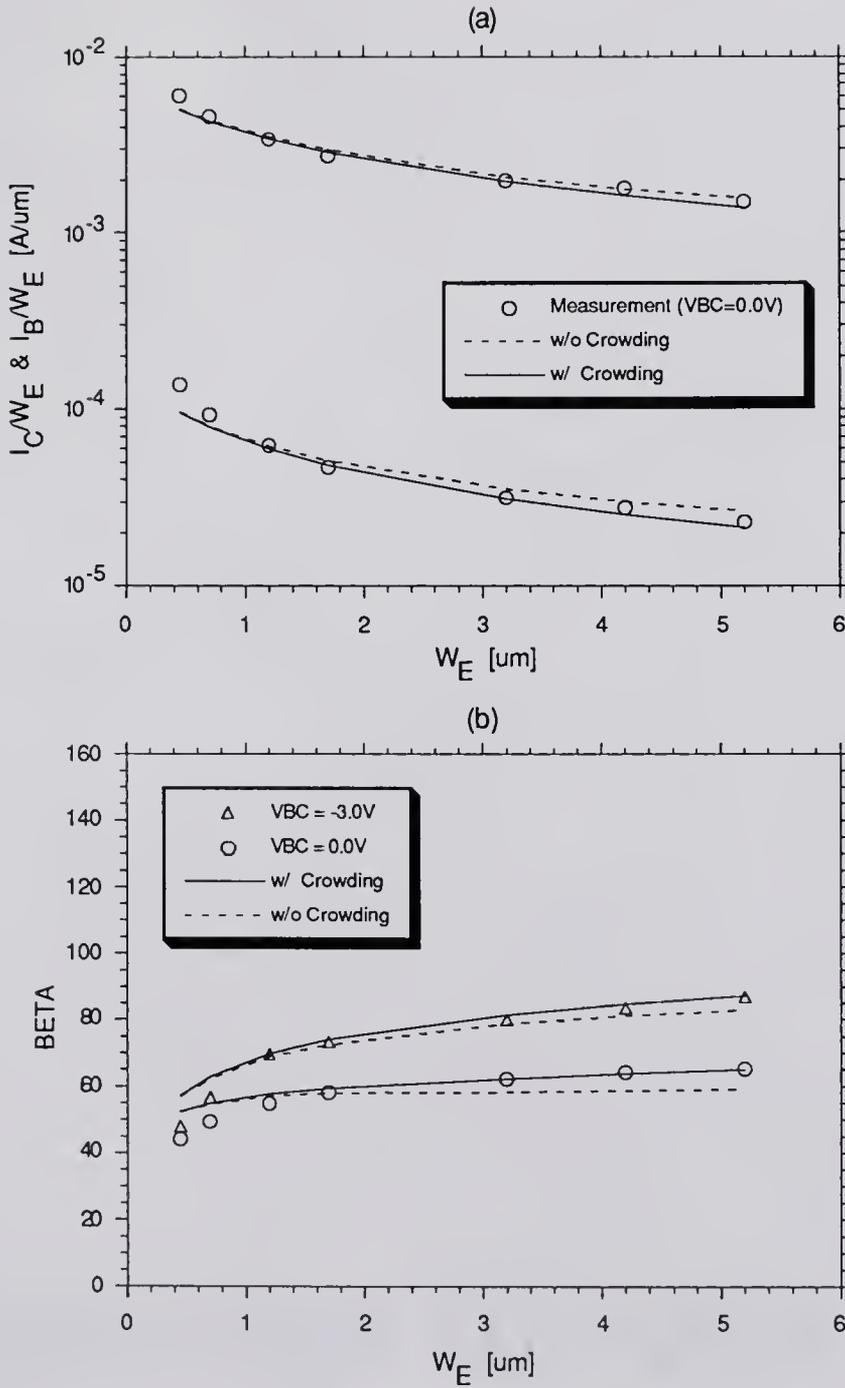


Fig. 2.6 Measured and simulated I_C/W_E , I_B/W_E in (a) and β in (b) for the test BJTs with $L_E=9.2\mu\text{m}$ for $V_{BE}=0.9\text{V}$.

2.4 Summary

A simple analytic model for the lateral injection of base current, which is shown to be the most predominant multi-dimensional current effect in advanced BJTs, has been developed by separating the base current into internal and peripheral components. The model is intended for (digital) circuit simulation and has been implemented in MMSPICE-2. For high V_{BE} , the effective bias (see Chapter 3) on the emitter-base junction is defined to account for the debiasing effect. The tool is well supported by experimental results of test BJTs having varied emitter geometries, despite the fact that the simulation for each device was done by scaling only A_E for a given parameter set. Therefore, this lateral injection model, combined with the current-crowding analysis, can be the basis for more predictive and scalable BJT simulation for TCAD.

For analog circuit simulations, more precision is usually required. In this case, it is possible to analyze more physically the multi-dimensional effects by cascading a second (peripheral) BJT to the intrinsic one, each represented by the one-dimensional BJT model in MMSPICE; the composite transistor is also useful to account for the parasitics associated with the extrinsic base region as well as the lateral injection effect, for example in RF IC design

applications [Jaf92].

From the measurement and simulation results for contemporary BJTs, the following conclusions were reached:

- (1) The lateral injection of the base current becomes more significant with decreasing V_{BE} , which reveals that the nature of this perimeter effect is recombination at the peripheral junction SCR near the oxide-silicon interface.
- (2) The peripheral component of collector current is typically negligible.
- (3) In high-current regions, the voltage drops across the extrinsic resistances are most predominant, and the dc debiasing effect seems to be negligible for contemporary BJTs.

CHAPTER 3
NON-QUASI-STATIC MODELING OF BJT CURRENT CROWDING

3.1 Introduction

In contemporary digital circuits containing advanced (scaled) BJTs, high transient base current can be much greater than the steady-state current; this clearly implies the non-quasi-static (NQS) nature of transient current crowding. (We generally define an NQS effect in the time [or ac-frequency] domain as one that cannot be inferred nor characterized from steady-state [dc] conditions.) Hence it can be significant even though dc crowding may be insignificant [Tan85]. The classical treatment of emitter current crowding by Hauser [Hau64], although useful, is inadequate for advanced BJTs because it neglects conductivity modulation of the base, which can occur because of high injection and/or base widening, and because it assumes steady-state or quasi-static conditions. In fact, transient current crowding is NQS, as well as being dependent on the base conductivity modulation [Ham88].

There has been some modeling done addressing the NQS nature of current crowding, but generally involving distributed lumped-model representations of the base region.

Indeed NQS effects can be physically accounted for by cascading a sufficient number of elemental quasi-static models, but computational efficiency must be sacrificed. Rey [Rey69] used a more novel approach to model ac crowding and derived a frequency-dependent base impedance for an equivalent-circuit model.

In this chapter we extend the one-dimensional BJT model in MMSPICE-1 to account for three-dimensional transient current crowding in advanced, self-aligned devices which have peripheral base contacts. The formalism includes a novel methodology for semi-numerically modeling general NQS effects in transient device/circuit simulation. The new model characterizes a time-dependent effective bias on the emitter-base junction for each Newton-Raphson iteration of the circuit nodal analysis at each time-step. The semi-numerical analysis follows Hauser, but physically accounts for base conductivity modulation and the NQS nature of the crowding. The latter extension is effected by the novel modeling/implementation that involves the use of the previous time-step solution in the current time-step analysis. The model naturally accounts for dc crowding as well, which is important in analog circuits, and which was needed in Chapter 2 to interpret the multi-dimensional current measurements in the BJT. It does not require a lumped intrinsic base resistance [Jo90], which is commonly used in BJT circuit

models.

The NQS model, implemented in MMSPICE-2, enables a semi-numerical mixed-mode device/circuit simulation capability for application-specific TCAD. The tool is supported by numerical simulations of advanced BJT structures using PISCES [PIS84]. It is used to clarify the nature of the added (NQS) delay due to current crowding in switch-on and switch-off transients in representative BJT inverting circuits, and it reveals the significance of transient crowding even in submicron devices.

3.2 Model Development

The intrinsic base of the advanced (self-aligned) BJT is surrounded by a high-conductivity extrinsic base. Hence the predominant base current flow under a rectangular emitter is along the shorter emitter dimension (W_E); this is assumed in our (quasi-three-dimensional) crowding analysis. Consider a section of the base of an npn BJT as shown in Fig. 3.1, where W_E is shorter than the emitter length L_E . For transient conditions at a point in time, let $i_B(y)$ be the lateral base current which causes the crowding in the emitter-base junction. Then, the emitter-base junction voltage $v(y)$ can be expressed as

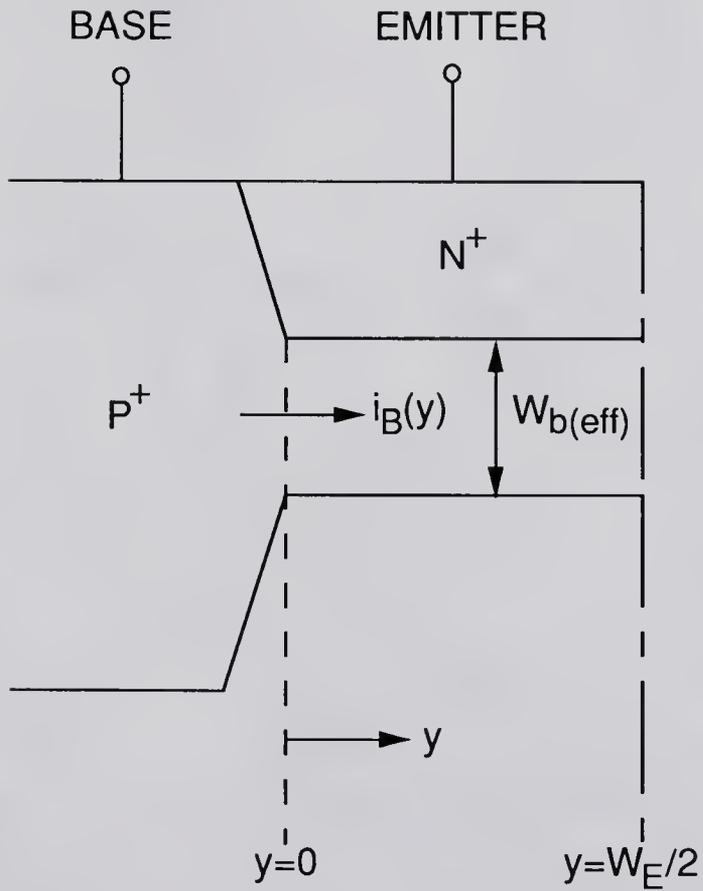


Fig. 3.1 Cross section of the advanced (symmetrical) bipolar junction transistor. $W_{b(eff)}$ is the widened (due to possible quasi-saturation) base width.

$$\begin{aligned}
v(y) &= v(0) - \int_0^y dv \\
&= v_{BE} - \int_0^y i_B(y) dR_{Bi} \\
&= v_{BE} - \int_0^y i_B(y) \rho dy
\end{aligned} \tag{3.1}$$

where v_{BE} is the peripheral junction voltage and ρ is the specific base resistivity,

$$\rho = \frac{dR_{Bi}}{dy} = \frac{1}{\bar{p}q\mu_p L_E W_{b(\text{eff})}} \approx \frac{W_E}{2\mu_p(Q_{BB} + Q_{QNR})} . \tag{3.2}$$

In (3.2), \bar{p} represents an average hole density at y , which we assume can be represented in terms of the total hole charge ($Q_{BB}+Q_{QNR}$) in the quasi-neutral base (possibly widened to $W_{b(\text{eff})}$ due to quasi-saturation); Q_{BB} , the hole charge in the metallurgical base region, and Q_{QNR} , the hole charge in the widened base region, both integrated over the emitter area A_E as well as over the base width, are characterized in the one-dimensional model [Jeo89]. This assumption in (3.2) is consistent with a quasi-two-dimensional analysis (to be described) which links the one-dimensional ambipolar

transport to the lateral hole flow. Implicit in the assumption is a neglect of lateral hole diffusion, which indeed is typically small compared to the lateral drift current when crowding is significant. The model deficiency resulting from this neglect will be shown to be inconsequential later. The hole mobility at y is also approximated by an average value $\overline{\mu_p}$, which is reasonably estimated from common sources. Note that the factor of 2 in the denominator of (3.2) accounts for the symmetry of the transistor obvious in Fig. 3.1.

For transient excitation, the main components of the intrinsic base current i_B are typically hole current back-injected from the base to the emitter (I_{BE}) and majority-hole charging/discharging current (dQ_{BE}/dt). Note that Q_{BE} includes components of ($Q_{BB}+Q_{QNR}$) communicating with the emitter [Jeo89]. It comprises space charge (e.g., junction depletion charge) as well as quasi-neutral-region charge in the intrinsic device structure. Generally, $I_{BE}(y)$ can be expressed as

$$I_{BE}(y) = I_{BE}(0) - \int_0^y 2J_{E0}L_E \left[\exp\left(\frac{v(y)}{V_T}\right) - 1 \right] dy \quad (3.3)$$

where J_{E0} is the (constant) emitter saturation current density. We assume that the y -dependence of $dQ_{BE}(y)/dt$, at a

particular point in time, may be similarly expressed as

$$\frac{dQ_{BE}^{m+1}(y)}{dt} = \frac{dQ_{BE}^{m+1}(0)}{dt} - \int_0^y 2J_Q^m L_E \left[\exp\left(\frac{v(y)}{V_T}\right) - 1 \right] dy \quad (3.4)$$

where $J_Q(t)$ is a transient (time-dependent) counterpart to J_{E0} . Implicit in (3.4) is an idea that J_Q can be estimated from the previous time-step ($t=t^m$) solution for dQ_{BE}/dt for use in the current time-step ($t=t^{m+1}$) analysis as follows:

$$J_Q^{m+1} \approx J_Q^m \equiv \frac{\frac{dQ_{BE}^m(0)}{dt}}{L_E W_E \left[\exp\left(\frac{v_{BE}^m(\text{eff})}{V_T}\right) - 1 \right]} \quad (3.5)$$

where $v_{BE(\text{eff})}$ is an NQS effective bias on the emitter-base junction defined (see (3.7)) to account for the current crowding (see the discussion in the Appendix B). So our model, when implemented based on the previous time-step solution, accounts for transient crowding non-quasi-statically. The approximation in (3.5) is viable even for fast transients because of the automatically controlled time-step reduction in the simulator, which is needed to ensure acceptable truncation error and convergence of the time-point solution.

With (3.3) and (3.4), the intrinsic base current $i_B(y)$ is written as

$$\begin{aligned}
i_B(y) &= I_{BE}(y) + \frac{dQ_{BE}(y)}{dt} \\
&= \left[I_{BE}(0) + \frac{dQ_{BE}(0)}{dt} \right] - \int_0^y 2(J_{EO} + J_Q)L_E \left[\exp\left(\frac{v(y)}{V_T}\right) - 1 \right] dy \\
&= i_B(0) - \int_0^y 2J_{EO(\text{eff})}L_E \left[\exp\left(\frac{v(y)}{V_T}\right) - 1 \right] dy \tag{3.6}
\end{aligned}$$

where the time-dependent $J_{EO(\text{eff})}$ is defined as the sum of J_{EO} and $J_Q(t)$. To facilitate an analytic accounting for the crowding (reflected by the integral in (3.6)), we define $v_{BE(\text{eff})}$ based on the total intrinsic base current:

$$i_B(0) = L_E W_E J_{EO(\text{eff})} \left[\exp\left(\frac{v_{BE(\text{eff})}}{V_T}\right) - 1 \right] . \tag{3.7}$$

Note that (3.7) is consistent with (3.5).

Now, following Hauser's classical analysis [Hau64], we differentiate (3.6) combined with (3.1) to get

$$\frac{\partial i_B}{\partial y} = -2J_{EO(\text{eff})}L_E \left[\exp\left(\frac{1}{V_T}(v_{BE} - \int_0^y i_B(y)\rho dy)\right) - 1 \right] . \tag{3.8}$$

This integral-differential equation for $i_B(y)$ may be transformed into a closed-form second-order differential

equation by differentiating it. This differentiation, with

$$\exp\left[\frac{v(y)}{V_T}\right] \gg 1 \quad (3.9)$$

for all values of y , which is generally valid for problems of interest, yields

$$\frac{\partial^2 i_B}{\partial y^2} + \frac{\rho}{V_T} i_B \frac{\partial i_B}{\partial y} = 0 \quad . \quad (3.10)$$

For transient crowding, (3.10) has two different types of solution depending on the sign of $\partial i_B / \partial y$. We consider the two cases separately.

3.2.1 Switch-on Case

When the BJT is switched-on, $i_B > 0$ tends to cause peripheral-emitter current crowding, as in dc crowding [Tan85]. In this case, $\partial i_B / \partial y$ is negative, and the solution of (3.10) is

$$i_B(y) = A \tan\left[\frac{A\rho}{2V_T} B \left(1 - \frac{y}{B}\right)\right] \quad (3.11)$$

where A and B are arbitrary constants of integration. The constants can be evaluated from the boundary conditions of

the problem. For the structure shown in Fig. 3.1, we have due to the symmetry

$$i_B\left(\frac{W_E}{2}\right) = 0 \quad , \quad (3.12)$$

which gives $B=W_E/2$. Then from (3.11),

$$i_B(y) = A \tan\left[z\left(1 - \frac{2y}{W_E}\right)\right] \quad (3.13)$$

where $z \equiv A\rho W_E/(4V_T)$. Hence, the total base current is

$$i_B(0) = A \tan(z) \quad , \quad (3.14)$$

which is equated to (3.7) to characterize $v_{BE(\text{eff})}$.

Using (3.13) in (3.1) and doing the integration yields

$$v(y) = v_{BE} - 2V_T \ln\left[\frac{\cos\{z(1 - 2y/W_E)\}}{\cos z}\right] \quad . \quad (3.15)$$

Note for this case that

$$v\left(\frac{W_E}{2}\right) = v_{BE} + 2V_T \ln(\cos z) < v_{BE} = v(0) \quad .$$

Now using (3.15) in the integration in (3.6), with the boundary condition (3.12), yields another expression for the

total base current:

$$i_B(0) = L_E W_E J_{EO(\text{eff})} \exp\left(\frac{v_{BE}}{V_T}\right) \frac{\text{sinz} \text{cosz}}{z} \quad . \quad (3.16)$$

With (3.7), (3.14), and (3.16) we now have a set of three nonlinear equations in three unknowns ($v_{BE(\text{eff})}$, A , and $i_B(0)$), which can be numerically solved by the iterative Newton-Raphson method. An interesting relationship is an expression relating $v_{BE(\text{eff})}$ to v_{BE} . This is obtained by equating (3.7) to (3.16):

$$\exp\left(\frac{v_{BE(\text{eff})}}{V_T}\right) = \exp\left(\frac{v_{BE}}{V_T}\right) \frac{\text{sinz} \text{cosz}}{z} \quad (3.17)$$

for $\exp[v_{BE(\text{eff})}/V_T] \gg 1$. Note that $v_{BE(\text{eff})}$ is always less than v_{BE} for the switch-on case since $(\text{sinz} \text{cosz}/z)$ is less than unity.

The accounting for dc crowding in the model is inherent in the switch-on analysis described above. For the dc case, $J_Q=0$ and $J_{EO(\text{eff})}=J_{EO}$ in (3.16).

3.2.2 Switch-off Case

For the switch-off case, $i_B < 0$ tends to cause central-emitter current crowding. The analysis is very similar to

that for switch-on, except that now $\partial i_B / \partial y$ is positive. Actually this condition does not obtain instantaneously when the BJT is abruptly turned off from an on-state. A very fast transient occurs during which holes diffuse out of the intrinsic base periphery to support the central-emitter crowding that ultimately controls the predominant switch-off transient. Our model presented below is invalid during this fast transient since it neglects lateral diffusion flow. However this brief invalidity is typically inconsequential with regard to simulating the predominant transient. Note that the fast (diffusion) transient is governed by a lateral quasi-neutral base transit time for minority electrons; it is proportional to $(W_E/2)^2 / \overline{D_n}$ where $\overline{D_n}$ is an average diffusion constant for electrons.

With the same boundary condition (3.12), the solution of (3.10) with $\partial i_B / \partial y > 0$ is

$$i_B(y) = -A \tanh \left[z \left(1 - \frac{2y}{W_E} \right) \right] . \quad (3.18)$$

So, the total base current is now

$$i_B(0) = -A \tanh(z) . \quad (3.19)$$

Once again we define the NQS effective bias $v_{BE(\text{eff})}$ by (3.7), in which $J_{EO(\text{eff})}$ is now negative because predominant

discharging current flows in this case. Following the steps in the switch-on analysis, we get

$$v(y) = v_{BE} + 2V_T \ln \left[\frac{\cosh z}{\cosh\{z(1 - 2y/W_E)\}} \right] . \quad (3.20)$$

Note here that

$$v\left(\frac{W_E}{2}\right) = v_{BE} + 2V_T \ln(\cosh z) > v_{BE} = v(0) .$$

The total base current can now be derived, analogously to (3.16), as

$$i_B(0) = L_E W_E J_{EO(\text{eff})} \exp\left(\frac{v_{BE}}{V_T}\right) \frac{\cosh z \sinh z}{z} . \quad (3.21)$$

Once again we have a system of three nonlinear equations, (3.7), (3.19), and (3.21), that define $v_{BE(\text{eff})}$ semi-numerically via iterative solution. Another interesting relationship between v_{BE} and $v_{BE(\text{eff})}$ is obtained from (3.7) and (3.21):

$$\exp\left(\frac{v_{BE(\text{eff})}}{V_T}\right) = \exp\left(\frac{v_{BE}}{V_T}\right) \frac{\cosh z \sinh z}{z} \quad (3.22)$$

for $\exp[v_{BE(\text{eff})}/V_T] \gg 1$. Note that $v_{BE(\text{eff})}$ is always greater than v_{BE} in the switch-off case since $(\cosh z \sinh z/z)$ is greater than unity.

We note that the switch-off analysis described above has no solution for extremely large negative $J_{EO(eff)}$, which tends to obtain when the discharging current dQ_{BE}/dt (viz., J_Q in (3.5)) becomes too large compared with the dc current I_{BE} . This condition is non-physical, and reflects the deficiency of our model during the initial fast (diffusion) transient discussed previously.

The no-solution problem can be avoided by limiting $J_{EO(eff)}$. Such limitation results in a solution, albeit invalid, that most importantly carries the simulation through the fast transient to the most significant lateral-drift-controlled switch-off transient. So, for each iteration at each time-step, we calculate a hypothetical maximum absolute value of $J_{EO(eff)}$ for which the system of equations is solvable, and then compare it with the actual $J_{EO(eff)}$; the smaller value is used for the analysis. Details are given in the Appendix C. This hypothetical limit for $J_{EO(eff)}$ is, as expected, used only at the very beginning of the switch-off transient, where the model is non-physical anyway, and indeed is insignificant with regard to the predominant transient.

3.3 NOS Model Implementation

Our novel NOS modeling/implementation in MMSPICE-2 of the BJT current crowding involves a coupling of the vertical

and lateral carrier-transport analyses in the base region. For the npn device, the analysis of the two-dimensional hole flow semi-numerically defines $v_{BE(eff)}$ for each Newton-Raphson iteration of the circuit nodal analysis at each time step. The implemented transient-crowding model algorithm is flowcharted in Fig. 3.2. The calculation of J_Q from the previous time-step solution for use in the current time-step is done only in the first iteration at each time step, and the value is used for all subsequent iterations. With the terminal biases v_{BE} and v_{BC} passed in from the nodal analysis, the one-dimensional model routine in MMSPICE solves the ambipolar transport, accounting for constant extrinsic terminal resistances, and characterizes the base charge in both the metallurgical (Q_{BB}) and widened (Q_{QNR}) base regions. These charges define the specific base resistivity (ρ) for the current time-step analysis, which is needed in the solution of the hole transport to derive a new $v_{BE(eff)}$. As discussed in Section 3.2, this derivation requires a Newton-Raphson iterative solution because of nonlinearities due to the conductivity modulation.

Note in Fig. 3.2 that $v_{BE(eff)}$ is not iteratively coupled to the one-dimensional model solution; that is, ρ is not updated to correspond with $v_{BE(eff)} \neq v_{BE}$. Although this one-pass derivation of $v_{BE(eff)}$ using $\rho(v_{BE})$ might seem incomplete, it is proper. A complete iterative solution, which would

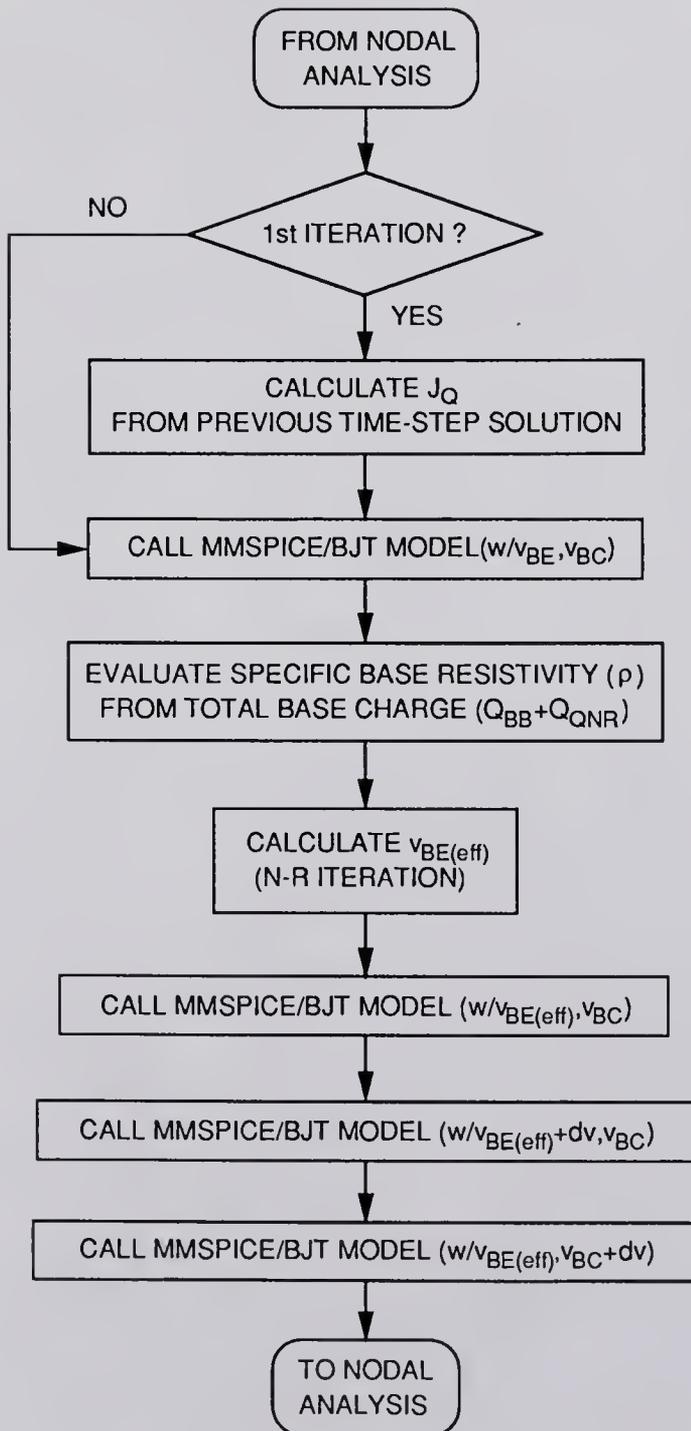


Fig. 3.2 Flowchart of the MMSPICE-implemented transient current crowding analysis, for every iteration at each time step.

require an outer Newton-like loop in the algorithm, would be non-physical. The reason is that in the switch-on case where $v_{BE(eff)}$ is less than v_{BE} , the smaller $v_{BE(eff)}$ in the one-dimensional model would not adequately account for possible high-current effects at the periphery, and that in the switch-off case where $v_{BE(eff)}$ is greater than v_{BE} , the larger $v_{BE(eff)}$ in the one-dimensional model would tend to diminish the central crowding effects by implying a smaller ρ .

With $v_{BE(eff)}$, the one-dimensional MMSPICE model routine is called again to obtain the nominal bias-point solution. Since the model is semi-numerical, analytic derivatives of the currents and charges cannot be given explicitly. Thus, numerical (divided-difference) approximations are used to evaluate (trans-)conductances and (trans-)capacitances for use in the subsequent nodal analysis. In order to do that, the model routine is called twice more with perturbed values of $v_{BE(eff)}$ and v_{BC} as indicated in Fig. 3.2. The admittance matrix is then loaded, and ordinary circuit nodal analysis follows.

3.4 Simulations

Examples of transient simulations using MMSPICE-2 are presented in this section. One circuit chosen for simulation is a single-transistor inverter shown in Fig. 3.3, with no

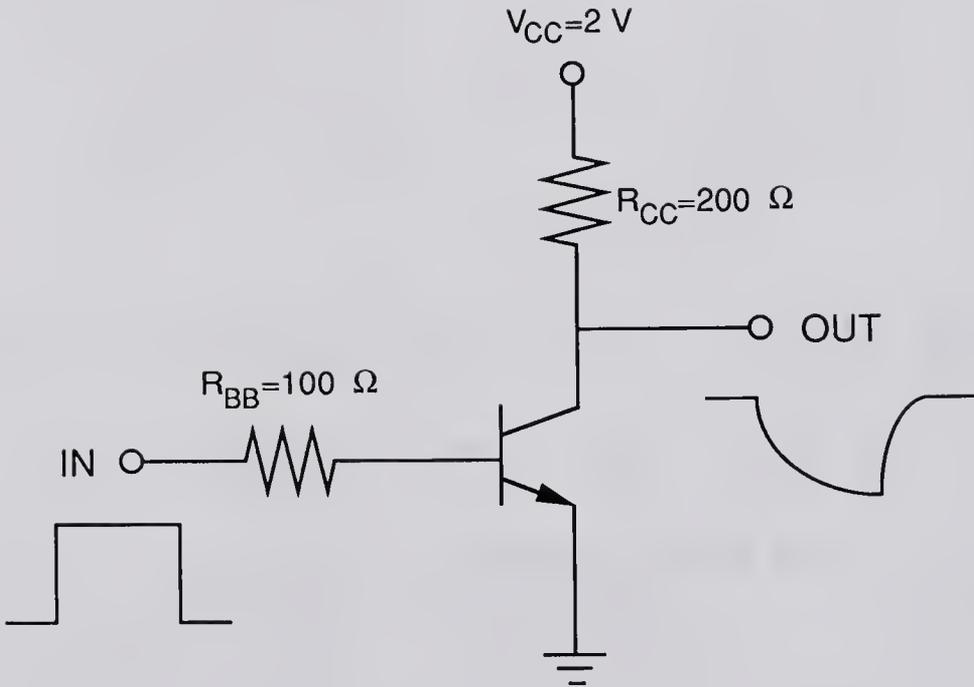


Fig. 3.3 A single transistor inverter circuit. The base terminal is driven with a voltage pulse that is delayed by 200ps and then ramped up (down) from 0.4V (0.9V) to 0.9V (0.4V) at a rate of 0.1V/ps.

load on the output. The assumed BJT model parameters characterize a typical advanced device structure with $W_E=1.2\mu\text{m}$. The peak base doping density is $1.5\times 10^{18}\text{cm}^{-3}$ and the metallurgical base width is $0.15\mu\text{m}$. For the switch-on transient, the NQS nature of the transient current crowding is well illustrated in Fig. 3.4 where the simulated time-dependent J_Q , defined in (3.5), is compared with J_{E0} . Note that J_Q is several orders of magnitude greater than J_{E0} at the moment the device is switched-on. It decreases monotonically with time and finally becomes less than J_{E0} only when the device nears steady state. In the switch-off case, J_Q is negative, and its magnitude is not so large as for the switch-on case. This is due to the $\exp[v_{BE(\text{eff})}/V_T]$ term in the denominator of (3.5), which is large when the device is switched off.

For the complete switch-on/switch-off cycle, Fig. 3.5 contrasts the simulated $v_{BE(\text{eff})}$ with v_{BE} in time, accounting for constant extrinsic/external base resistance, which is reflected by the discrepancies between v_{BE} and the input voltage v_{in} . The moment the device is switched-on, $v_{BE(\text{eff})}$ becomes, as mentioned earlier, less than v_{BE} due to the high transient base current-induced crowding, but then increases steadily with time to a value that corresponds to dc crowding, which is relatively insignificant. For the switch-off transient, $v_{BE(\text{eff})}$ is greater than v_{BE} , but the difference

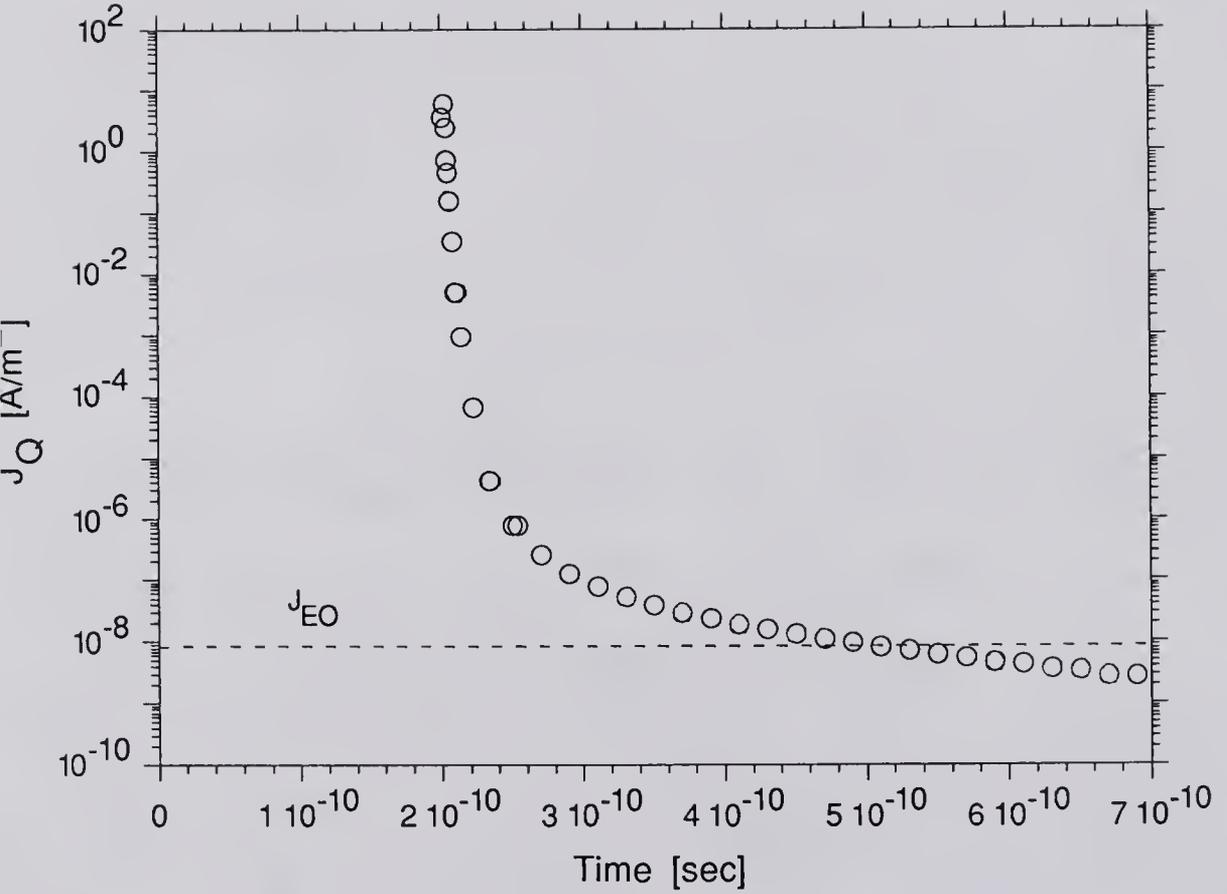


Fig. 3.4 Simulated J_Q versus time in the switch-on case. J_{EO} is the emitter saturation current density.

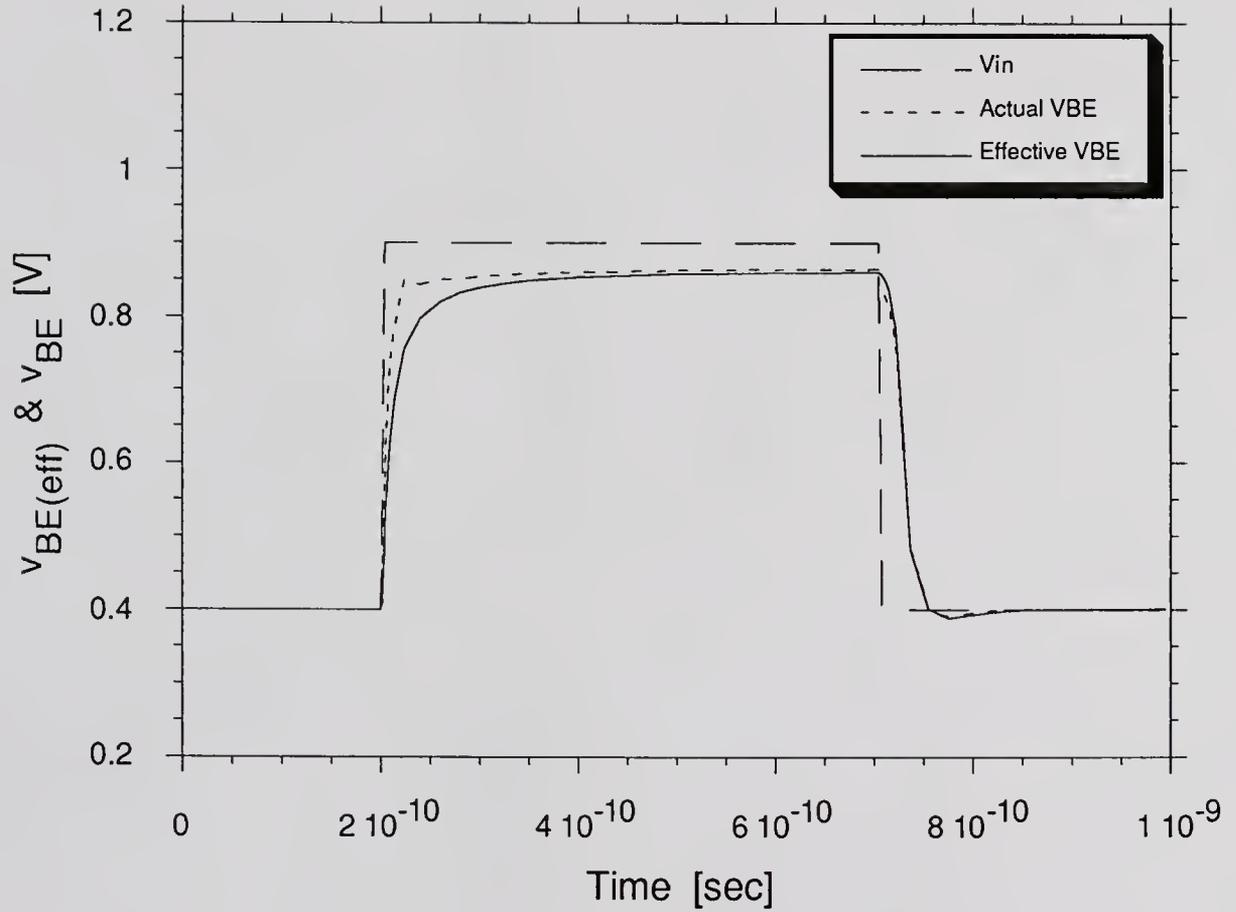


Fig. 3.5 Simulated $v_{BE(eff)}$ versus time for the complete switch-on/switch-off cycle. The input pulse and the actual (peripheral) base-emitter junction voltage are shown for comparison.

is not so noticeable as for the switch-on case. These results suggest that the central-emitter current crowding during a switch-off transient is much less significant than the peripheral-emitter crowding during a switch-on transient. This can be attributed to the level of base conductivity modulation (reflected by ρ) at the initial stages of the respective transients.

Fig. 3.6 shows the output voltage characteristics of the inverter simulated with (MMSPICE-2) and without (MMSPICE-1) the current crowding accounted for. In accord with conclusions drawn from Fig. 3.5, the result of the switch-on transient crowding is a substantively slower response, while the added delay is insignificant for the switch-off transient. Other simulations show that accounting for only quasi-static crowding (due to J_{E0} in (3.6)) yields an output voltage characteristic which is virtually identical to that predicted by the simulation in Fig. 3.6 for which crowding was completely neglected.

Predicted switch-on delays of the single transistor inverter versus W_E , with the emitter area fixed ($A_E = L_E \times W_E = 9.2 \times 2.0 \mu\text{m}^2$), and with the emitter area scaled with W_E , are plotted in Fig. 3.7. The emitter width W_E was varied using the values 0.1, 0.4, 1.2 and 2.0 μm . The delay was defined as the time for the output current to reach 50% of its final (high) value. The effect of the crowding is made

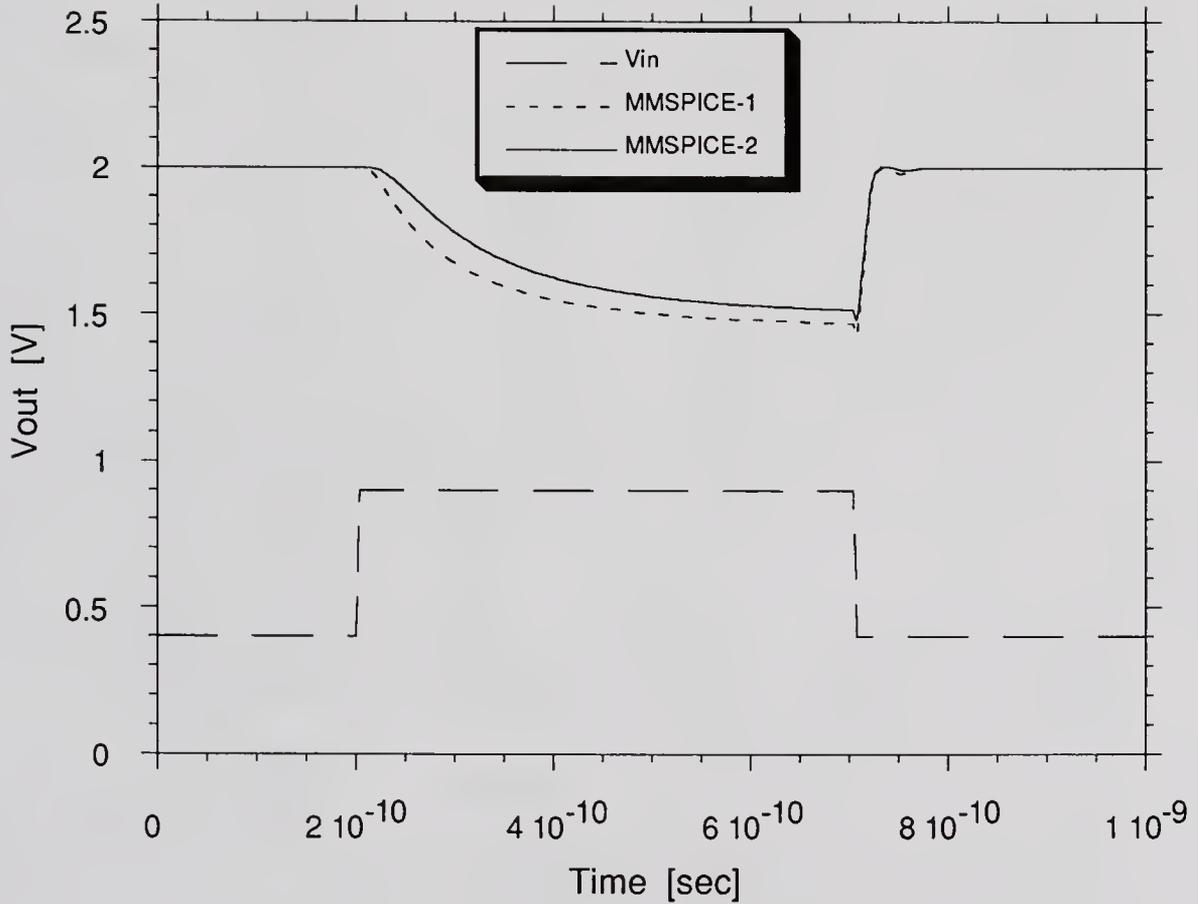


Fig. 3.6 Output voltage characteristics of the single transistor inverter simulated with (MMSPICE-2) and without (MMSPICE-1) the transient current crowding accounted for.

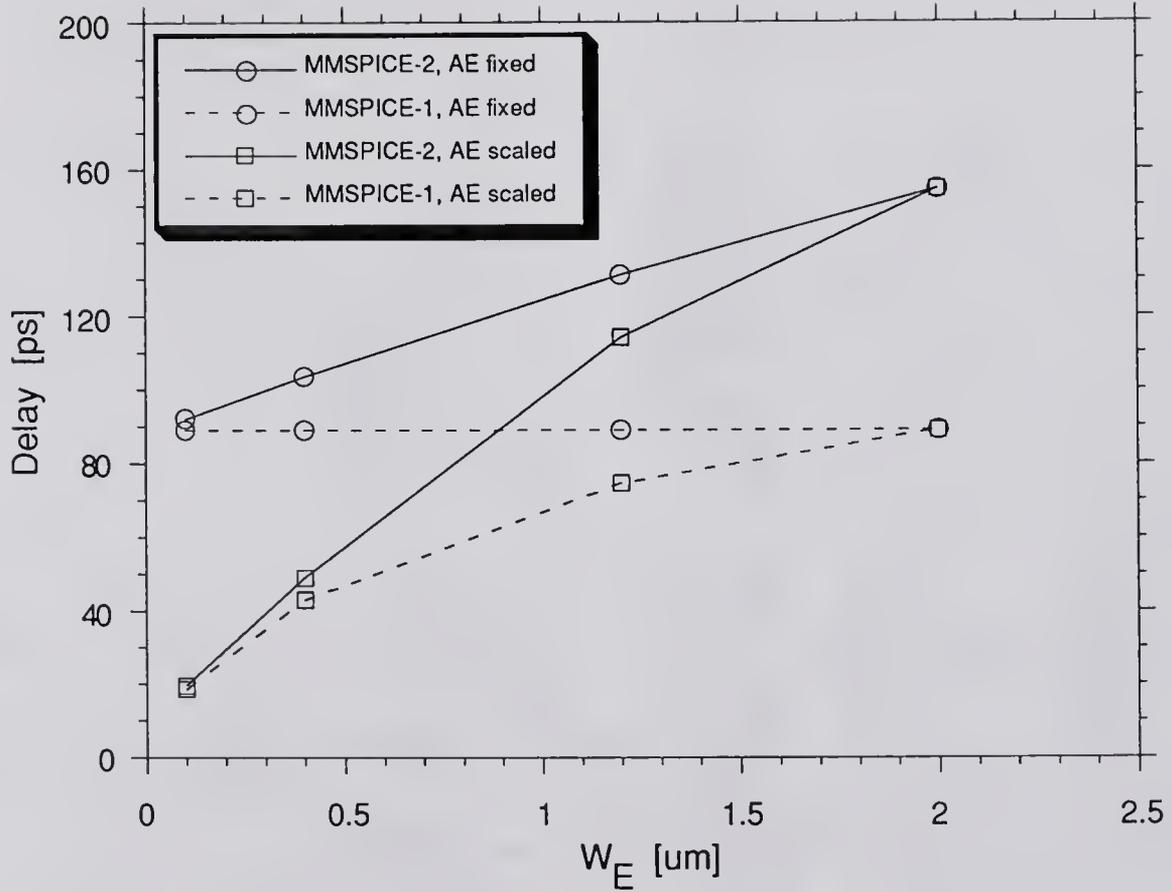


Fig. 3.7 Predicted switch-on delays of the single transistor inverter versus W_E , with the emitter area fixed ($A_E = 9.2 \times 2.0 \mu\text{m}^2$), and with the emitter area scaled with W_E .

apparent by including in the figure delays predicted by one-dimensional (MMSPICE-1) simulations. For the switch-on transient, the results, consistent with previous work [Tan85], show that peripheral-emitter crowding causes an added delay, one that tends to become insignificant only when W_E is reduced to deep-submicron values [Ham88]. Note in Fig. 3.7 that when the emitter area is scaled with W_E , the delay is more sensitive to W_E . The reason of course is that, in addition to the crowding effect, the amount of charge that must be stored in the BJT varies with W_E . Other simulations show that the relative importance of the crowding varies inversely with the extrinsic (plus external) base resistance.

Results of switch-off simulations with varying W_E show that the added delay due to central-emitter crowding is negligible, at least for $W_E < 2\mu\text{m}$. Indeed the simulations predict that the reduced delay of a scaled (W_E and A_E) device is due predominantly to the reduced charge storage in the BJT.

The effect of the emitter length L_E on the current crowding is reflected in Fig. 3.8, which shows normalized predicted switch-on delays versus W_E for devices with A_E fixed at $9.2 \times 2.0\mu\text{m}^2$ or $3.2 \times 2.0\mu\text{m}^2$. Note that for a fixed W_E , the crowding effect on the delay diminishes with increasing L_E . This is due to the decreasing specific resistivity ρ in (3.2).

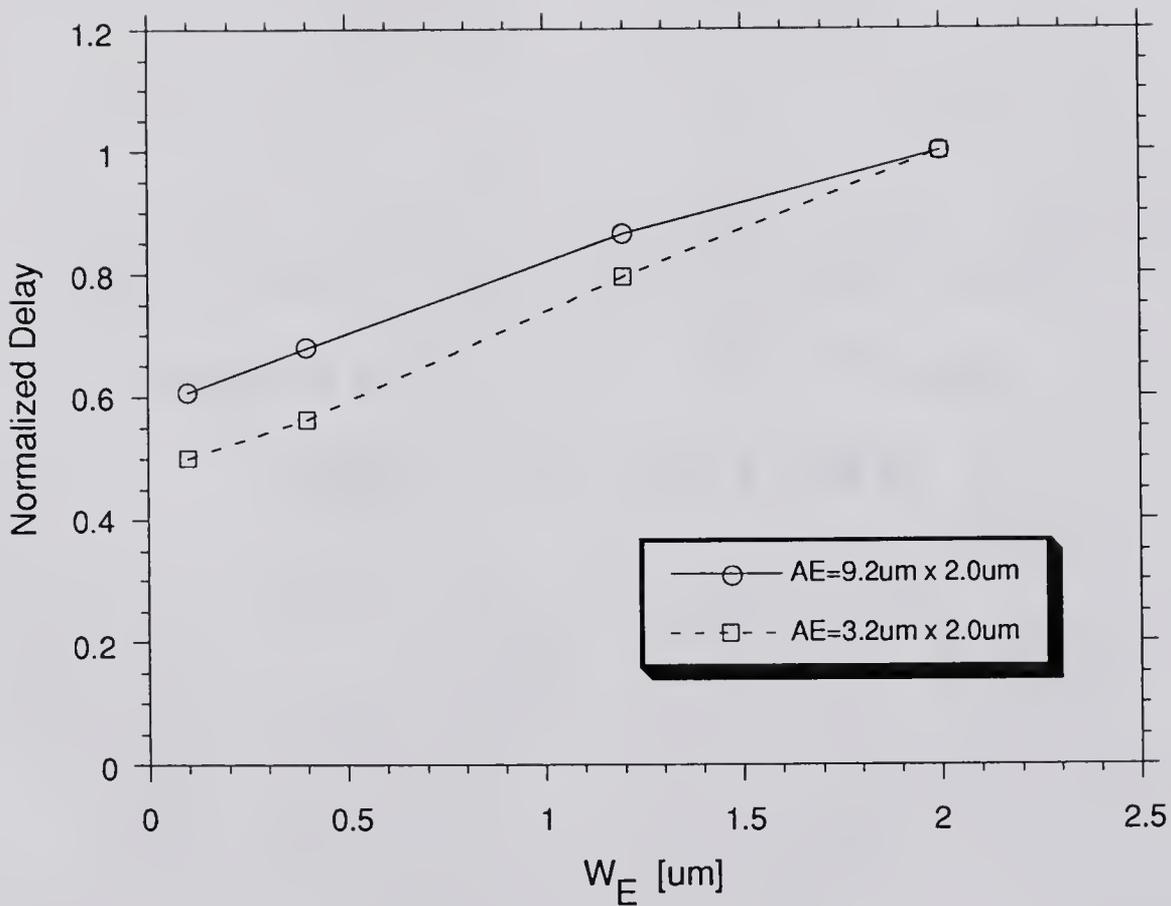


Fig. 3.8 Predicted normalized switch-on delays versus W_E , with fixed A_E , for devices with different L_E .

The influence of the nominal base resistivity, viz., the Gummel number, on the added switch-on delay due to crowding is revealed in Fig. 3.9 where predicted normalized delays are plotted versus W_E (with fixed $A_E=9.2 \times 2.0 \mu\text{m}^2$) for three different metallurgical base widths W_{BM} . The peak base doping density was fixed at $1.5 \times 10^{18} \text{cm}^{-3}$. The plots show how the transient crowding becomes more significant as W_{BM} is scaled down, independent of the increasing current gain of the BJT since there is no load on the inverter (Fig. 3.3).

In order to verify our model, two-dimensional numerical simulations of the nominal BJT inverter were performed using PISCES [PIS84], the results of which for varying W_E are shown in Fig. 3.10. In these switch-on and switch-off simulations, the actual emitter length was fixed at $1 \mu\text{m}$ because the output currents of PISCES are always normalized by the length perpendicular to the simulated structure. Also, the values of W_E used for the plots are the effective emitter widths, which are about $0.2 \mu\text{m}$ wider than the poly-emitter windows because of lateral diffusion. The contact resistances at the collector and base terminals were specified to include the external resistances in the inverter circuit. Included in Fig. 3.10 are corresponding MMSPICE device/circuit simulations, with $L_E=1 \mu\text{m}$. In the switch-on case, the transient current crowding is significant and is faithfully predicted by MMSPICE-2, as contrasted by the inaccurate

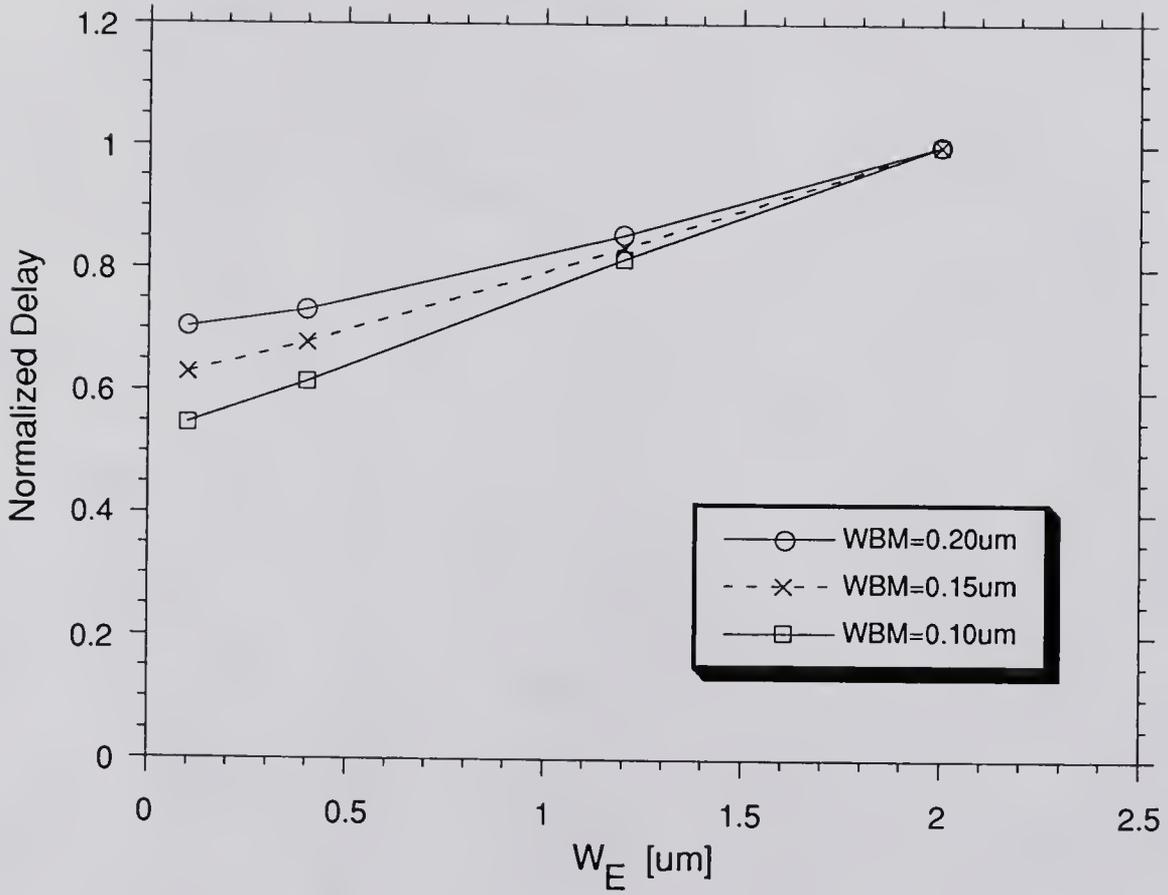


Fig. 3.9 Predicted normalized switch-on delays versus W_E , with fixed A_E , for devices with different W_{BM} .

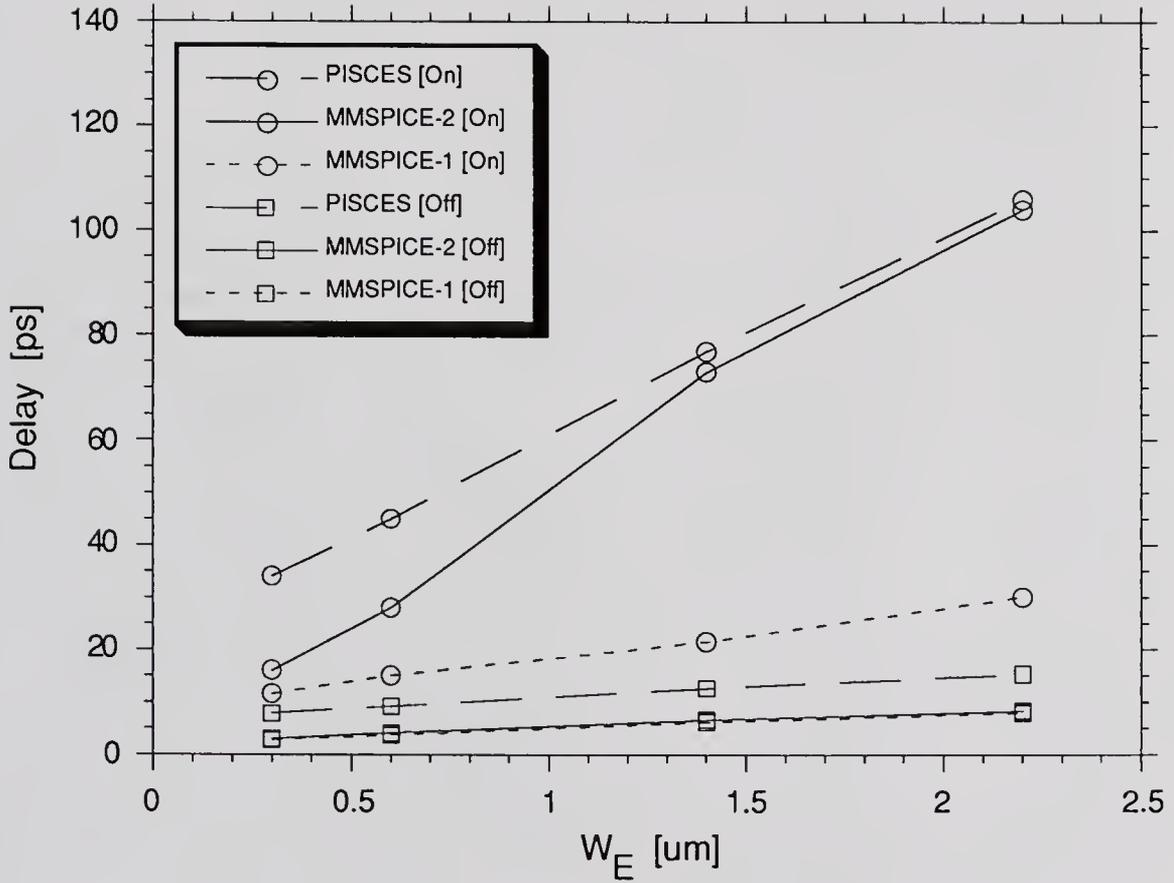


Fig. 3.10 PISCES simulations of the switch-on and switch-off delays versus (effective) W_E of the single transistor inverter, with corresponding MMSPICE simulations. $L_E=1\mu\text{m}$ for all simulations.

MMSPICE-1 simulations which are also shown. Some discrepancy in the submicron region is apparent. This could be due to a parasitic peripheral-region transistor unaccounted for in MMSPICE-2 simulations; or possibly to slightly different physical model parameters, e.g., mobility, assumed by PISCES and MMSPICE-2. In the switch-off case, the crowding is seen to be insignificant as implied previously. It can be inferred then that the reduction of switch-off delay of a scaled device is primarily caused by the reduced charge storage rather than the diminished crowding in the BJT.

Additional verification of the NQS crowding formalism in MMSPICE-2 is provided in Fig. 3.11 where switch-on transient collector currents predicted by PISCES, MMSPICE-2, and MMSPICE-1 are plotted. These currents were taken from the $W_E=1.4\mu\text{m}$ simulations of Fig. 3.10. Note the good correspondence in time between the PISCES and MMSPICE-2 currents, which are separated from the MMSPICE-1 current by a significant (added NQS) delay.

In MMSPICE-1, a semi-empirical accounting for current crowding can be effected by using a parameter which defines the intrinsic base resistance as a function of the current-dependent charge. Although the parameter could account for the current crowding for given device dimension, it is not applicable to other device dimensions since the parameter is neither scalable nor predictable. Hence it cannot yield a

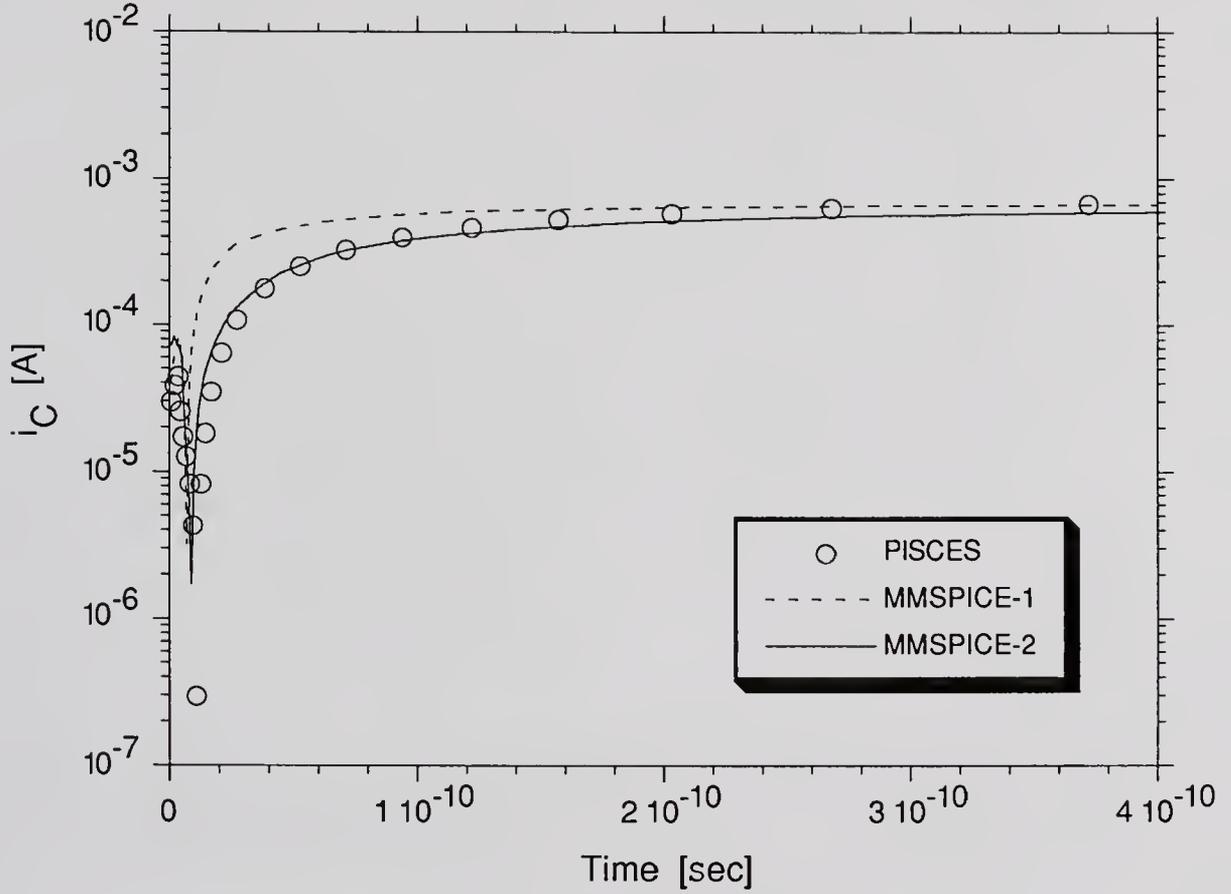


Fig. 3.11 Predicted switch-on transient collector currents taken from the PISCES, MMSPICE-2, and MMSPICE-1 simulations of Fig. 3.10 for $W_E=1.4\mu\text{m}$.

trend like Fig. 3.10.

Finally, to emphasize the mixed-mode NQS simulation capability of MMSPICE-2, transient simulations of an ECL inverter stage, the basic building block of high-speed digital circuits, were done. Fig. 3.12 shows the circuit diagram; the four nominal BJTs have $W_E=1.2\mu\text{m}$. The output voltage waveforms of the circuit predicted with and without (via MMSPICE-1) current crowding are plotted in Fig. 3.13. The effect of the NQS current crowding is apparent; the propagation delay is increased by almost 50%.

3.5 Summary

A novel NQS model for transient current crowding in advanced BJTs has been developed. The new model, based on the use of the previous time-step solution in the current time-step analysis, characterizes a time-dependent effective bias on the emitter-base junction for each circuit nodal iteration at each time-step in a semi-numerical analysis following Hauser [Hau64], but physically accounting for base conductivity modulation and the NQS nature of the crowding. The NQS model, implemented in MMSPICE-2, enables a semi-numerical, scalable, mixed-mode device/circuit simulation capability for application-specific TCAD. The tool is supported by numerical simulations of advanced BJT structures

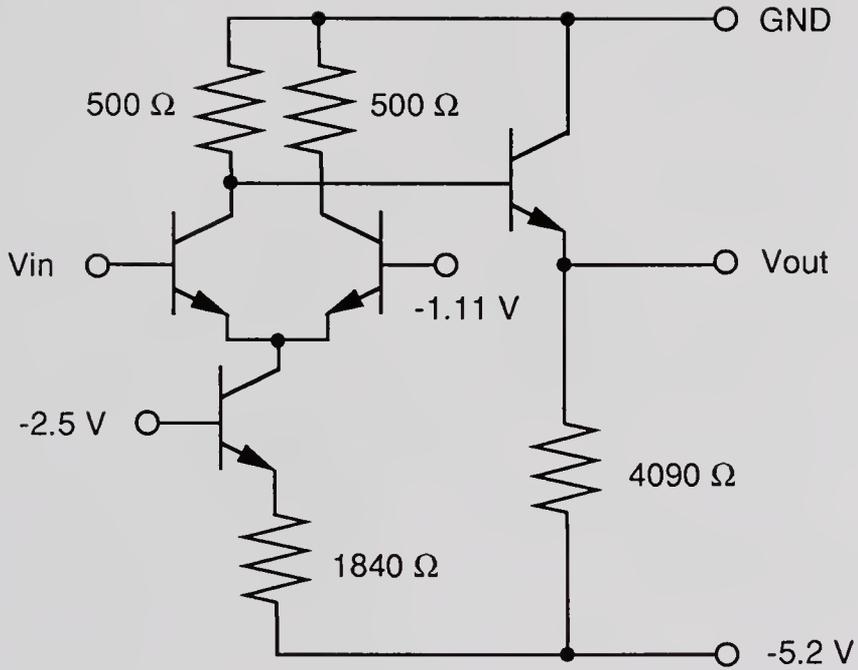


Fig. 3.12 An advanced-technology ECL inverter circuit. The four BJTs have $L_E/W_E=9.2\mu\text{m}/1.2\mu\text{m}$.

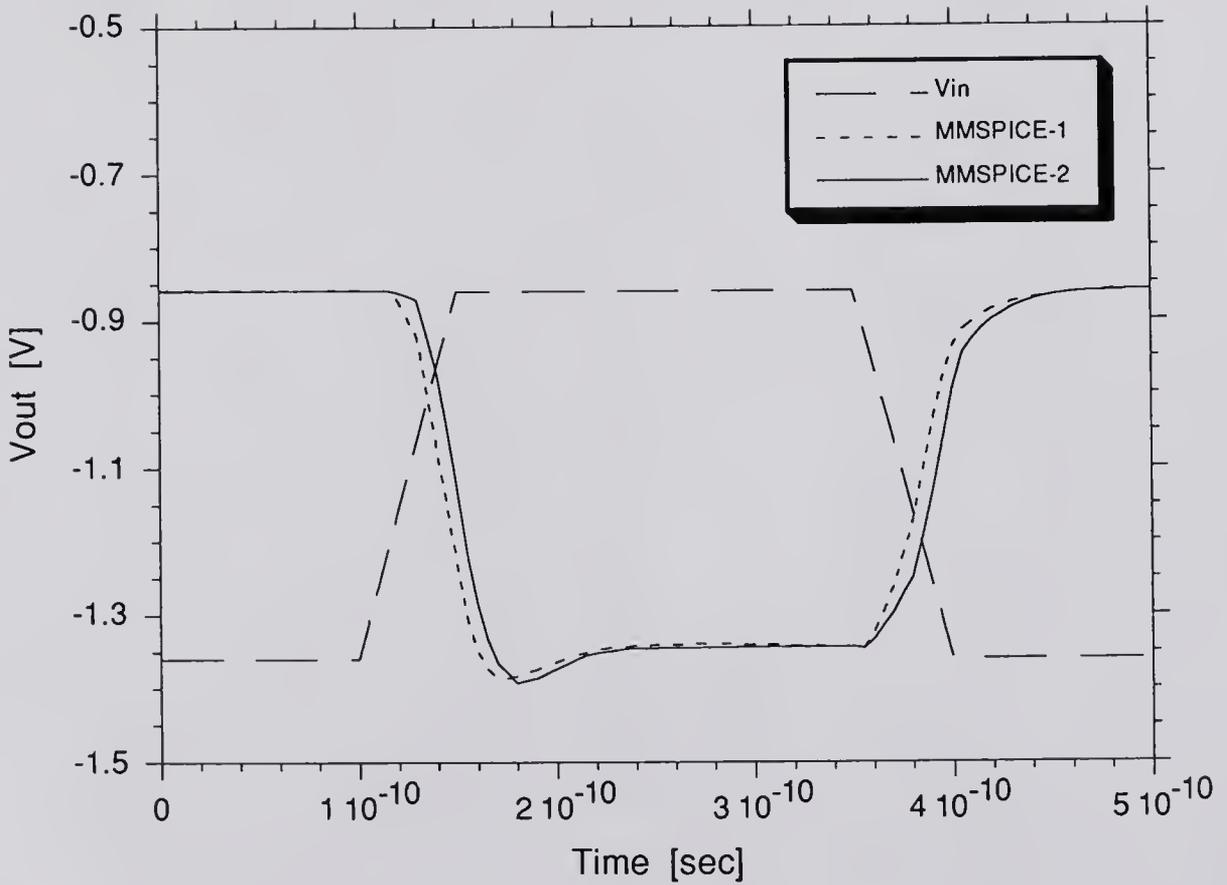


Fig. 3.13 Switching waveforms of the ECL inverter circuit simulated with (MMSPICE-2) and without (MMSPICE-1) the transient current crowding accounted for.

using PISCES. From the simulations of a representative BJT inverter circuit, the following conclusions were reached.

- (1) For the switch-on transient, peripheral-emitter crowding causes an added delay, and tends to become insignificant only when W_E is scaled to deep-submicron values.
- (2) For the switch-off transient, the added delay due to central-emitter crowding is negligible, at least for $W_E < 2\mu\text{m}$. Indeed the reduced delay of a scaled (W_E and A_E) device is due predominantly to the reduced charge storage in the BJT.

We note that the novel modeling/implementation involving use of the previous time-step solution to update the model for the current time-step analysis could be a viable means of accounting for general NQS behavior in semi-numerical transient device/circuit simulation. Such behavior must indeed be modeled to enable truly predictive mixed-mode simulation for TCAD.

CHAPTER 4
ANALYTIC ACCOUNTING FOR CARRIER VELOCITY OVERSHOOT

4.1 Introduction

In advanced silicon-based bipolar technology, the vertical as well as the lateral dimensions of the BJT are being scaled to deep-submicron values. Consequently, very high electric fields and field gradients are not uncommon in the scaled device. When the field increases rapidly over distances comparable to the energy-relaxation mean free path, carrier velocity can overshoot the value corresponding to the local electric field. This enhanced transport occurs because the carrier (kinetic) energy, which controls the collision time and hence limits the velocity, lags the field and remains relatively small [Ruc72]. Such a non-local effect has been recognized as significant in MOSFETs and MESFETs for years, but only now is its significance in advanced bipolar transistors (BJTs) becoming an issue [Lee89, Cra90].

Recent work [Fus92] has indicated that velocity overshoot in scaled silicon BJTs can be beneficial, and must be accounted for in the device and circuit design. The effect, however, has not yet been physically accounted for in any circuit simulator. Indeed, this phenomenon is not

accounted for in most device simulators because of the implied computational intensiveness. The conventional drift-diffusion current equation used in ordinary circuit and device simulators does not account for the non-local effect of an inhomogeneous electric field on the carrier velocity. It is based on the assumption that the drift velocity is a function of the local electric field, and ignores the actual dependence (of mobility) on carrier energy.

Non-local effects on carrier transport have been accounted for using different analyses, but with severe restrictions because of the accuracy/computational efficiency trade-off. Hence these analyses--which include rigorous Monte Carlo statistical treatments [Lee89], less complex solutions of the hydrodynamic equations involving the solution of the moments of Boltzmann transport equation (i.e., a set of equations describing conservation of particle number, momentum, and energy solved in conjunction with Poisson's equation) [Blo70], and even simpler solutions of the energy transport equations which, with some assumptions, can be derived from the hydrodynamic model [Bor91]--have limited utility for device simulation and virtually no use for circuit simulation. Alternatively, the so-called augmented drift-diffusion (ADD) transport model [Tho82], which retains most of the efficiency of the drift-diffusion equation but uses additional analytic terms to account for

the non-local effects, has been proposed as a way of efficiently extending the utility of drift/diffusion-based tools for scaled technologies.

In Section 4.2, a simple but physical analytic model for first-order accounting of the electron velocity overshoot in advanced silicon-based BJT "circuit simulation" is presented. The model, which characterizes the non-local electron velocity in the high-field collector space-charge regions (SCRs), is shown to be identical to the ADD formalism when the electron diffusion is negligible. The associated velocity relaxation, which is not accounted for in the ADD model, is characterized phenomenologically to be consistent with the overshoot analysis. In Section 4.3, the comparison of our model with the energy transport analysis is presented. In Section 4.4, the implementation of the model in MMSPICE is discussed. In the last section, device and circuit simulation results are presented to assess the significance of the electron velocity overshoot in advanced silicon bipolar and BiCMOS technologies, and to support the model. This is the first time that a non-local effect has been explicitly accounted for in a circuit simulator.

4.2 Model Development

4.2.1 Velocity Overshoot

When the randomly moving conduction-band electrons in a semiconductor encounter an electric field, they experience an increase in average (drift) velocity, and an increase in average kinetic energy which however tends to lag the drift velocity [Ruc72]. When the kinetic energy is important (i.e., when the electrons are not in thermal balance with the lattice), a phenomenological force acting on the electrons can be expressed in one dimension as

$$qE_{(\text{eff})} = \frac{d}{dx} (\mathcal{E}_c - \mathcal{W}) \quad (4.1)$$

where \mathcal{E}_c and \mathcal{W} are the (average) potential and kinetic energies of the electrons respectively. Note that \mathcal{E}_c and \mathcal{W} in (4.1) are "correlated" in accord with electron flow. When \mathcal{W} is small ($\approx 3kT/2$ where T is the lattice temperature), $E_{(\text{eff})}$ is the actual field, E , proportional to $d\mathcal{E}_c/dx$ as it is classically expressed.

Ballistic transport of the electrons, driven by E , would result in unlimited \mathcal{W} . However the electrons in a crystal lattice frequently collide with impurities and phonons, the result of which is to randomize their motion and limit their

(average) drift velocity, v , and hence their momentum. Effectively the collisions give rise to a retarding force proportional to the velocity, as characterized by the balance of momentum [Shu81]:

$$m^* \frac{dv}{dt} = -qE_{(\text{eff})} - m^* \frac{v}{\tau(\mathcal{W})} \quad (4.2)$$

where m^* is the effective mass of conduction (sub-)band electrons and $\tau(\mathcal{W})$ is an energy-dependent momentum relaxation time. Combining (4.1) and (4.2) yields

$$\begin{aligned} m^* \frac{dv}{dt} &= \left(-\frac{d\mathcal{E}_c}{dx} + \frac{d\mathcal{W}}{dx} \right) - m^* \frac{v}{\tau(\mathcal{W})} \\ &= \left(-qE + \frac{d\mathcal{W}}{dx} \right) - m^* \frac{v}{\tau(\mathcal{W})} \quad . \end{aligned} \quad (4.3)$$

For dc or quasi-static analysis, $dv/dt \approx 0$ in (4.3) and

$$v = \frac{\tau(\mathcal{W})}{m^*} \left(-qE + \frac{d\mathcal{W}}{dx} \right) \quad . \quad (4.4)$$

Note that when $d\mathcal{W}/dx$ is negligible, (4.4) becomes a well-known equation defining the electron mobility $\mu(\mathcal{W})$ ($=|v/E|$):

$$\mu(\mathcal{W}) \equiv \frac{q\tau(\mathcal{W})}{m^*} \quad . \quad (4.5)$$

The mobility is expressed as a function of \mathcal{W} to emphasize that it depends more on the local carrier energy than on the local electric field. Using (4.5) in (4.4) with the chain rule for differentiation gives

$$\begin{aligned} v &= -\mu(\mathcal{W})E \left[1 + \left(\frac{1}{q} \frac{d\mathcal{W}}{d|E|} \right) \left(\frac{1}{E} \frac{dE}{dx} \right) \right] \\ &\equiv v_0(E) \left[1 + \frac{L(E)}{E} \frac{dE}{dx} \right] \end{aligned} \quad (4.6)$$

where $v_0(E)$ is the conventional drift velocity defined by the local field, and $L(E) \equiv (d\mathcal{W}/d|E|)/q$ is a phenomenological length coefficient [Pri88], which describes to first-order the non-local effect of the electric field gradient on v . For $L(E) \neq 0$, a large dE/dx in (4.6) implies a possibly significant velocity overshoot, $|v| > |v_0(E)|$, in accord with the more rigorous physics underlying the electron transport. Note that (4.6) is identical with the ADD formalism [Tho82] when the diffusion of carriers is negligible [Kan91]. The field gradient in (4.6) was substituted with the quasi-Fermi level by other authors [Kiz89], to avoid inappropriate overshoot corrections in the presence of built-in electric field. However this would not be important in real applications, since the simulation of the equilibrium condition is not needed in most cases.

The length coefficient has been characterized via Monte Carlo analysis [Art88] by several investigators. However the results show some quantitative differences, possibly because of the different transport parameters and band structures used. Recently, Chen et al [Che91] derived an analytic formula for $L(E)$, but its utility is subject to uncertainties in the evaluation of some model parameters. Hence we suggest a simplified piecewise-linear representation of $L(E)$ for electrons in silicon at room temperature, based on Artaki's Monte Carlo simulations [Art88], which is illustrated in Fig. 4.1. In fact, $L(E)$ can be negative for low $|E|$, although the velocity undershoot thereby implied by (4.6) is generally not significant [Lun90] and will be neglected here.

Equation (4.5) implies that the classical mobility decreases with increasing electric field since the electrons gain kinetic energy which reduces the average (scattering) time between collisions. When the velocity imparted to an electron by the applied field is much less than the random thermal velocity, τ is however insensitive to E , implying a linear $v(E)$ dependence: $v_0 = -\mu_0 E$ where μ_0 is the low-field mobility. At high fields however, the drift velocity becomes comparable to the random thermal velocity, and τ is reduced. The drift velocity (magnitude) in this case, in the absence of a high gradient of E , approaches a limiting (saturated)

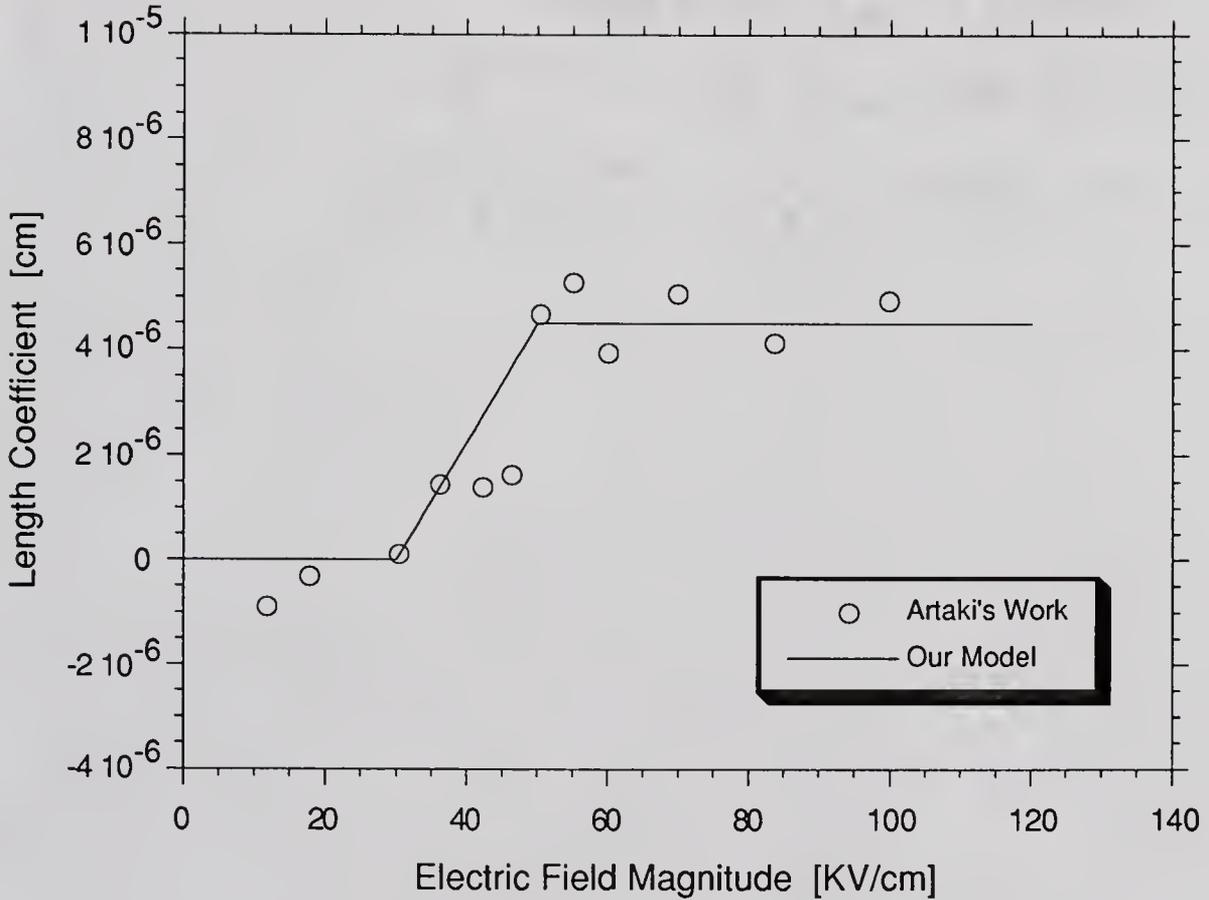


Fig. 4.1 The length coefficient versus electric field (magnitude) for silicon at room temperature. The points were derived from Monte Carlo simulations [Art88], and the piecewise-linear approximation is used in our model.

value v_s ($\approx 10^7$ cm/sec in silicon at room temperature), which can be empirically expressed as the product of μ_0 and a critical electric field (magnitude) E_s defining the onset of velocity saturation: $v_s = \mu_0 E_s$.

Hence depending on the magnitude of the electric field in a region with $d|E|/dx > 0$, the magnitude of the carrier drift velocity in (4.6) can be expressed as

$$|v| = \mu_0 |E| \left[1 + \frac{L(E)}{E} \frac{dE}{dx} \right] \approx \mu_0 |E| \quad \text{for } |E| < E_s \quad (4.7)$$

and

$$= v_s \left[1 + \frac{L(E)}{E} \frac{dE}{dx} \right] \quad \text{for } |E| > E_s. \quad (4.8)$$

The typical value of E_s for electrons in silicon at room temperature is less than 30KV/cm, and for $|E| < E_s$, $L(E)$ vanishes as shown in Fig. 4.1. Hence as indicated in (4.7), $|v| \approx \mu_0 |E|$ for this case, in accord with the conventional characterization. This simplification means that the velocity overshoot characterization is needed only when $|E| > E_s$ as in (4.8), and that otherwise the conventional drift-diffusion formalism with (4.7) is still applicable even though $d|E|/dx$ is high.

4.2.2 Velocity Relaxation

The analytic velocity overshoot characterization in (4.8) is strictly valid only when the magnitude of the electric field is increasing in the drift current direction. It would yield no overshoot when $d|E|/dx = 0$ or an undershoot when $d|E|/dx < 0$, independent of the history of the transport, and hence is non-physical for these cases. For example, a hot (high- \mathcal{W}) electron entering such a region where $d|E|/dx$ is not positive must travel a few mean free paths to reach the velocity corresponding to the local field, and hence would experience velocity overshoot. This relaxation can be neglected for MOSFETs and MESFETs because the only significant non-local effects occur under the gate where electrons are accelerated to the drain by a high field with $d|E|/dx > 0$ [Kiz89, Kan91]. However for the BJT, which contains significant (space-charge) regions with $d|E|/dx < 0$ adjacent to those with $d|E|/dx > 0$, the velocity relaxation following overshoot must be simulated. Details on various types of SCRs will be presented in next section.

To understand the velocity relaxation in the advanced BJT, consider a mental experiment. Fig. 4.2 shows the possible relaxation of the drift velocity in the collector side of the base-collector junction SCR where $|E|$ is

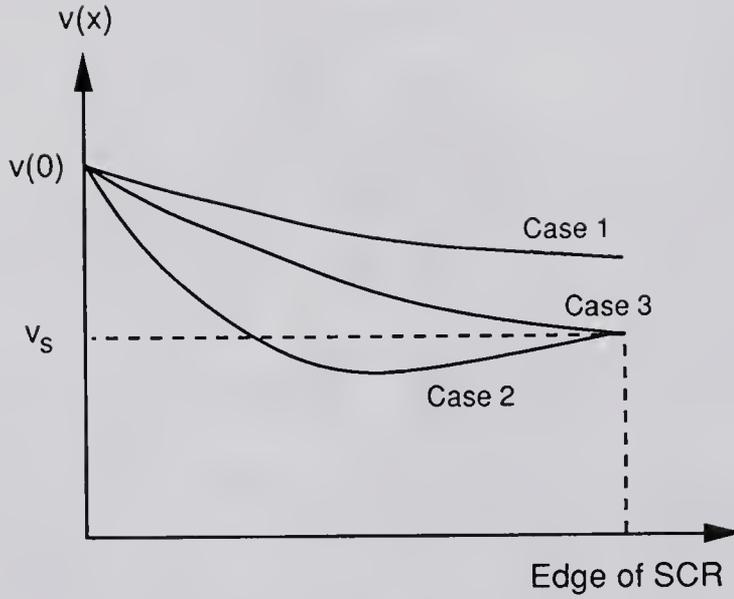


Fig. 4.2 Possible distributions of the drift velocity when $|E|$ is decreasing with distance. Note that the electric field magnitude at the edge of SCR is assumed to be E_s .

decreasing with distance (see Fig. 4.3(a)). Note that the electric field magnitude at the (nebulous) edge of the SCR is implicitly assumed to be E_s [Jeo89]. Normally when a hot electron leaves a high-field region, its velocity will decrease with distance due to the scattering by which it transfers its energy to the lattice (see Case 1 in Fig. 4.2). The relaxation however becomes somewhat different when the width of the SCR gets smaller. At a glance, it seems likely that the velocity would not decrease very much from its value at the junction because of the reduced scattering. But actually this tendency would be compensated by the velocity undershoot tendency [Lun90], which obtains when the electric field is decreasing very rapidly. The kinetic energy responds to fields more slowly than does the carrier velocity; hence immediately after the high- to low-field transition, the carrier's kinetic energy is still high, and thus its mobility is lower than that corresponding to thermal balance between the carrier and the lattice. After the electron has dissipated its excess energy, it would then have the velocity v_s (see Case 2). This is supported by the fact that $L(E)$ in (4.8) is 0 at the edge of the SCR because $|E|$ is assumed to be E_s . Taking these two conflicting phenomena into consideration, we assume that the velocity would decay monotonically with distance and finally reach v_s at the edge of the SCR (as described by Case 3).

Based on this insight, we use a phenomenological representation of the velocity relaxation in an SCR where $d|E|/dx < 0$ by simplifying (4.2) to

$$\frac{dv}{dt} \equiv \frac{dv}{dx} v \approx - \frac{v}{\tau} \quad ,$$

or

$$\frac{dv}{dx} \approx - \frac{1}{\tau} \equiv - \frac{v}{s} \quad (4.9)$$

where s is an average mean free path for velocity relaxation. The solution of (4.9) is

$$v(x) = v(0) \exp(-x/s) \quad (4.10)$$

where $v(0)$ is the velocity at the point where $|E|$ is maximum in the SCR. Since the velocity must be continuous, $v(0)$ is derived from the analysis of the velocity (overshoot) in the adjacent region where $d|E|/dx > 0$. To estimate s , we assume as discussed above that the carrier velocity reaches v_s at the edge of the SCR. (This assumption is consistent with a common designation of an SCR [Jeo89].) Thus

$$s \approx \frac{W_{RR}}{\ln\left(\frac{v(0)}{v_s}\right)} \quad (4.11)$$

where W_{RR} is the width of the relaxation region.

4.2.3 Effective Saturated Drift Velocity

To this point, we have modeled the hot-electron velocity in an SCR using either the length coefficient or the scattering mean free path, depending on the sign of $d|E|/dx$. To facilitate the implementation (discussed later) of the model into the bipolar device/circuit simulator MMSPICE, we define now an effective saturated drift velocity $v_{s(eff)}$ based on the actual transit time of electrons in the SCR being analyzed:

$$\int_{W_{SCR}} dt = \int_{W_{SCR}} \frac{dx}{v(x)} \equiv \frac{W_{SCR}}{v_{s(eff)}} \quad (4.12)$$

where $v(x)$ is given by (4.8) or (4.10), and W_{SCR} is the width of the SCR in which $|E|$ is greater than E_s .

For the advanced BJT, different operating conditions are distinguished by the charge conditions [Jeo89] in the epitaxial collector region, as reflected in Fig. 4.3. The electric field distributions shown are determined by the bias on the base-collector junction and the collector current.

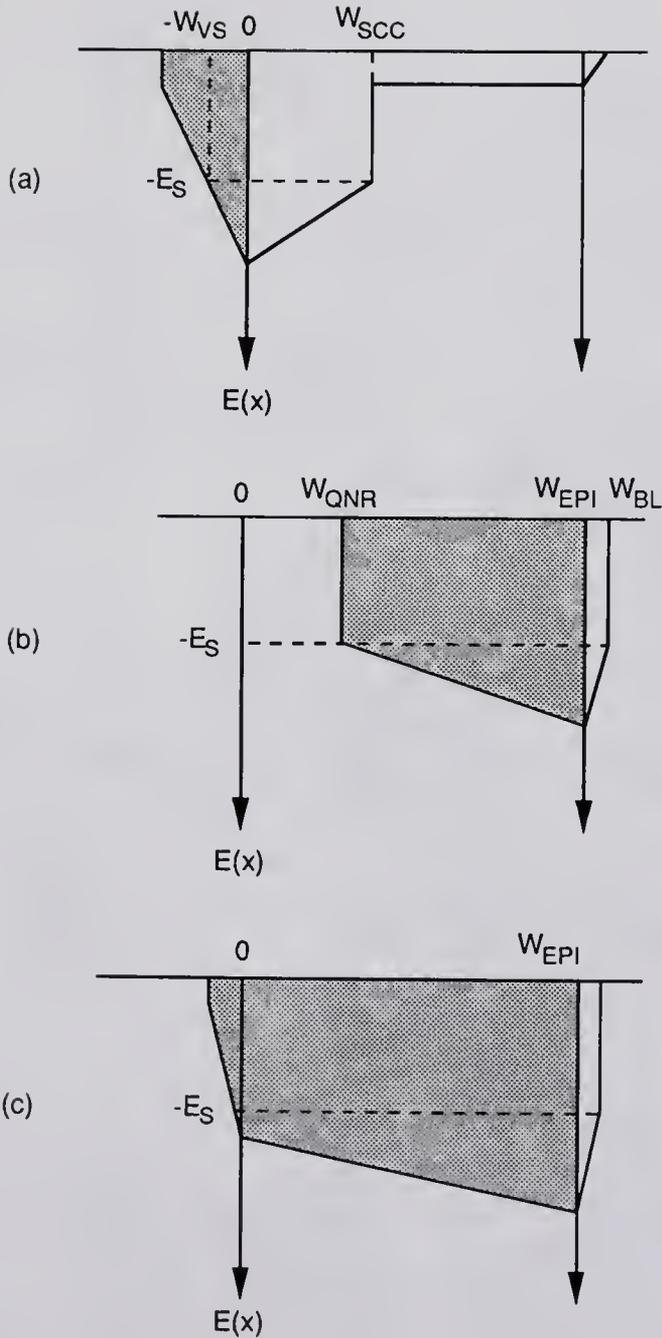


Fig. 4.3 Electric field distributions in a base-collector junction SCR (a), and a current-induced SCR (b) associated with non-ohmic quasi-saturation, i.e., base pushout. When either SCR expands, the entire epi layer can become space-charged (c).

Fig. 4.3(a) represents the conventional junction SCR at the base-collector junction under low-current conditions. For high-current conditions, when non-ohmic quasi-saturation (base pushout) prevails, the current-induced SCR exists in the epi-collector as denoted in Fig. 4.3(b). Note that the electric field is assumed to be $-E_s$ at the edge of the collector-side SCR in both cases; this assumption in fact defines the SCRs [Jeo89]. When either SCR expands, the entire epi layer can become space-charged, as shown in Fig. 4.3(c). We must consider the three SCR types in the BJT separately.

4.2.3.1 Junction SCR

When the SCR exists across the base-collector junction, as shown in Fig. 4.3(a), (4.12) applied to it yields

$$\int_{-W_{VS}}^0 \frac{dx}{v(x)} + \int_0^{W_{SCC}} \frac{dx}{v(x)} \equiv \frac{W_{VS} + W_{SCC}}{V_{s(eff)}} \quad (4.13)$$

where W_{VS} and W_{SCC} are the widths of the base and collector sides of the SCR respectively. The carrier velocity $v(x)$ is evaluated depending on the sign of $d|E|/dx$. In the base side, the velocity is characterized via (4.8), using the

depletion approximation coupled to a first-pass ($v_{s(\text{eff})} \rightarrow v_s$) MMSPICE simulation to describe $E(x)$ and W_{VS} :

$$\begin{aligned} \frac{dE}{dx} &= - \frac{q}{\epsilon} [N_A(x) + n] \\ &\approx - \frac{q}{\epsilon} N_A(x) \\ &= - \frac{q}{\epsilon} N_{AO} \exp\left[\frac{-\eta}{W_{BM}}(x + W_{BM})\right] \end{aligned} \quad (4.14)$$

where the assumed exponential doping profile is consistent with the base-transport analysis of the BJT model [Jeo89] in MMSPICE; W_{BM} is the metallurgical base width. Thus

$$E(x) = \frac{qW_{BM}}{\epsilon\eta} N_{AO} \exp\left[\frac{-\eta}{W_{BM}}(x + W_{BM})\right] + C \quad (4.15)$$

The integration constant C can be easily evaluated from the electric field at the junction ($x=0$), which is available from the output of the BJT model routine in MMSPICE. $E(x)$ and dE/dx are then substituted into (4.8) to give $v(x)$ for the first integral in (4.13). The validity of using the depletion approximation here will be discussed in Appendix D.

In the collector side, (4.10) is used directly for the second integral in (4.13), with $v(0)$ being equated to that derived from the analysis of base side. Both integrals in

(4.13) are evaluated by a numerical method to give $v_{s(\text{eff})}$. Strictly, the value of E_s in the base side tends to be greater than that in the collector side because the electron mobility (μ_0) in the base is lower due to the higher doping concentration. However because the (compensated) doping is generally not known precisely and because this variation in E_s is only a second-order effect, we neglect it.

4.2.3.2 Current-induced SCR

When the current-induced SCR exists, as illustrated in Fig. 4.3(b), (4.12) applied to it yields

$$\int_{W_{\text{QNR}}}^{W_{\text{EPI}}} \frac{dx}{v(x)} \equiv \frac{W_{\text{EPI}} - W_{\text{QNR}}}{V_{s(\text{eff})}} \quad (4.16)$$

where W_{QNR} is the extended width of the pushed-out (quasi-neutral) base region. The transit time across the portion of the SCR in the adjacent buried layer of the BJT structure is neglected since the heavy doping there implies only a negligibly thin depletion-region width, $W_{\text{BL}} - W_{\text{EPI}}$.

From the first-pass MMSPICE simulation [Jeo89], the electric field in the SCR and W_{QNR} are obtained in accord with

$$\frac{dE}{dx} = - \frac{q}{\epsilon} \Delta n$$

$$= - \frac{q}{\epsilon} \left(\frac{I_C}{qAv_s} - N_{EPI} \right) \quad (4.17)$$

where N_{EPI} is the doping concentration of the epi-collector layer and Δn is the excess electron density in the SCR, which is assumed to be spatially constant since the current I_C is constant. From (4.17),

$$E(x) = - \frac{q}{\epsilon} \left(\frac{I_C}{qAv_s} - N_{EPI} \right) (x - W_{QNR}) - E_s \quad (4.18)$$

Equations (4.17) and (4.18) are substituted into (4.8) to yield $v(x)$, and $v_{s(eff)}$ is evaluated from (4.16). In the vicinity of the boundary between the SCR and the quasi-neutral region in the epi-collector, the electric field gradient is very large. However this transitional region can be ignored because the length coefficient is, as shown in Fig. 4.1, assumed to be 0 when $|E| < E_s$.

4.2.3.3 Special case

In previous sections, the velocity overshoot effect was characterized via regional analyses depending on the sign of the field gradient. There is a special case for the BJT however where the overshoot effect would not be properly accounted for in this manner. This is the case where the epi

layer is completely space-charged, and the magnitude of the electric field is still increasing with distance due to non-ohmic quasi-saturation, as shown in Fig. 4.3(c). (Note that when the entire epi layer is space-charged, but $|E|$ is decreasing with distance, the overshoot analysis for the junction SCR is still applicable.)

According to our formalism, the same overshoot analysis would be applied in the collector side as in the base side. Of course, this is adequate if the field gradient is relatively large. When the electric field is increasing slightly however, the direct application of our model would tend to exaggerate the overshoot effect since velocity relaxation is ignored. In fact, the carrier velocity would decrease with distance in the epi-collector. In order to cope with this deficiency of our formalism, we empirically combine the overshoot model with the relaxation model for this case as follows:

$$v(x) \approx v_{rel}(x) + v_{ov}(x) \exp\left(\frac{-1}{f \cdot \frac{|dE|}{dx}}\right) + v_{offset} \quad (4.19)$$

where $v_{rel}(x)$ and $v_{ov}(x)$ are the velocity distributions characterized by the relaxation and overshoot models respectively, and f is an empirical weighting factor. When the field gradient is very small, (4.19) reduces to (4.10),

implying that velocity relaxation would be predominant in the collector side. When the gradient becomes large, $v(x)$ is given as the sum of $v_{rel}(x)$ and $v_{ov}(x)$ with the empirical factor chosen to ensure a smooth transition from velocity relaxation to velocity overshoot. The offset velocity, v_{offset} in (4.19) is used to make the velocity at the junction continuous.

4.3 Comparisons with Energy Transport Model

One way to characterize the velocity overshoot effect is to solve the energy transport equation [Bor91]. Such a solution can provide support for our simple analytic model. In this section, we will numerically solve the energy transport equation Goldsman *et al* presented [Gol88], and contrast it with our model. By assuming the electron energy as entirely thermal, they derived the steady-state momentum equation from the Boltzmann transport equation as

$$v = \frac{\tau_p(w)}{m^*} \left(-qE - \frac{2}{3} \frac{dw}{dx} - \frac{2w}{3n} \frac{dn}{dx} \right) \quad (4.20)$$

where w is the average electron energy ($\equiv \frac{1}{2} m^* v^2 + \frac{3}{2} kT_e \approx \frac{3}{2} kT_e$ where T_e is the electron temperature), $\tau_p(w)$ is the energy-dependent momentum relaxation time, and n is the electron concentration. Combining (4.20) (with $dn/dx=0$) with the

steady-state energy equation, Goldsman et al derived an equation for average electron energy that includes the effect of velocity overshoot:

$$\frac{dw}{dx} = -\frac{21}{20} qE - \frac{9}{20} \left[\frac{40}{9} \frac{m^*}{\tau_p \tau_w} (w-w_0) + q^2 E^2 \right]^{1/2} \quad (4.21)$$

where $\tau_w(w)$ is the energy relaxation time, and w_0 is the thermal energy of the lattice ($=3kT/2$).

In order to solve these equations, both τ_p and τ_w must be known as functions of the electron energy. Although Goldsman et al evaluated the relaxation times by Monte Carlo simulations in homogeneous fields, we use simple functions to empirically approximate the parameters they derived:

$$\tau_p(w) = c_0 + \frac{c_1}{w} \quad , \text{ and} \quad (4.22)$$

$$\tau_w(w) = d_0 + d_1 w + d_2 w^2 + d_3 w^3 \quad (4.23)$$

where c_n and d_n denote empirical constants. In Fig. 4.4, the discrete points represent the momentum and energy relaxation times Goldsman et al have derived, and the solid lines which best fit the data are given by (4.22) and (4.23). Then the energy dependent carrier velocity can be numerically evaluated from (4.20) and (4.21), since those equations are a function of the single variable w .

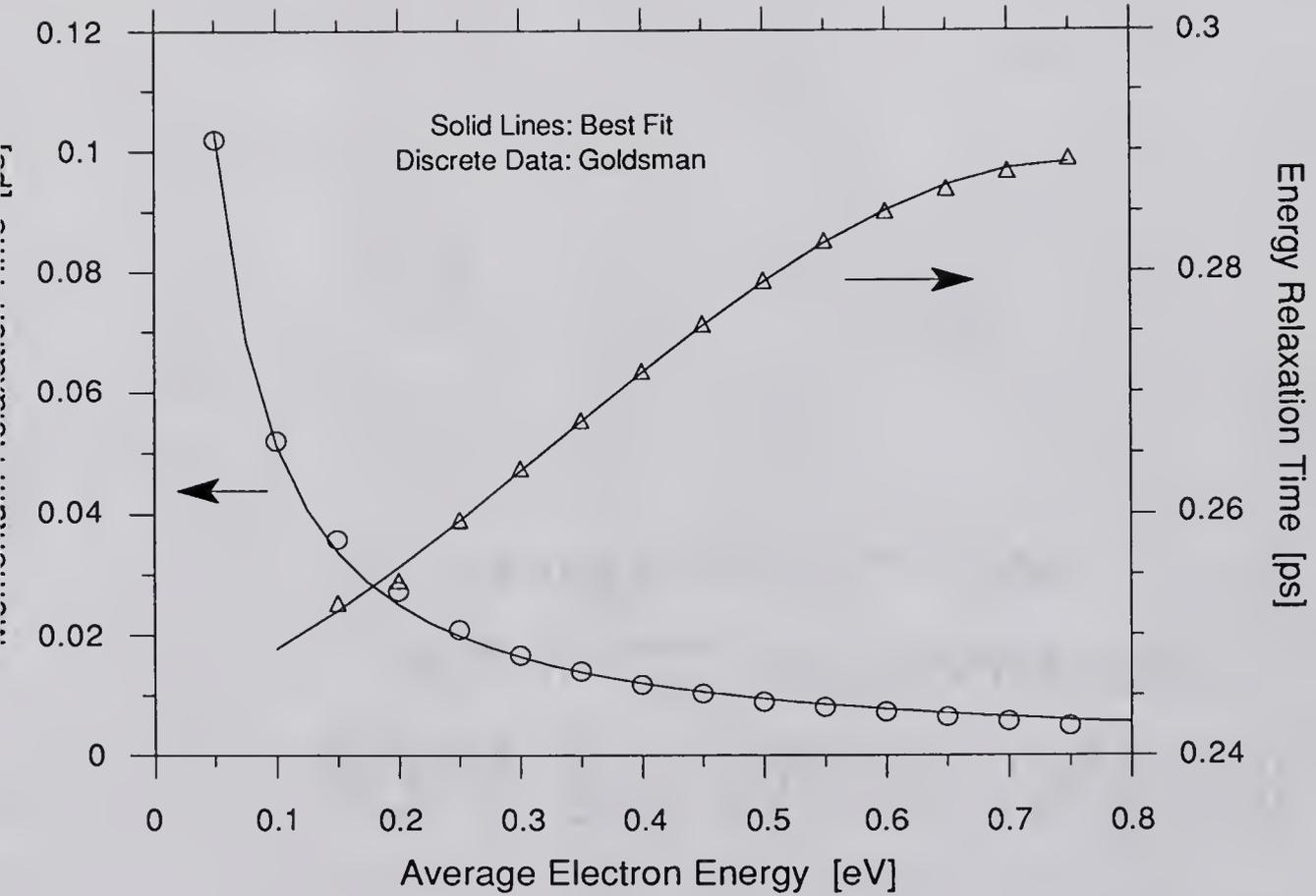


Fig. 4.4 Momentum and energy relaxation times as functions of energy.

For comparisons, we evaluated the velocity distributions for the typical advanced BJT, when the junction or the current-induced SCR exists, using our model and that of Goldsman *et al.* Fig. 4.5(a) shows the predicted velocity distributions in the junction SCR when $V_{BE}=0.7V$ and $V_{BC}=0.0V$ are applied to the terminals of the device. (For the effective mass of conduction sub-band electron, $m^*=0.26m_0$ was used, where m_0 is the rest mass [Mul89]. The electric field used as inputs for both the models was available from the output of MMSPICE.) As described before, our overshoot analysis is done when the magnitude of the electric field is increasing ($x<0$). In accord with our piecewise-linear $L(E)$ model, the carrier velocity reaches its peak value v_{peak} when the length coefficient is at its maximum value at $E=50KV/cm$ (see Fig. 4.1). Note that the location of v_{peak} is about the same as that predicted by the energy transport model. When reverse bias is applied on the base-collector junction ($V_{BC}=-2.0V$), v_{peak} increases as shown in Fig. 4.5(b), because the gradient of the electric field also increases. Figs. 4.6(a) and (b) illustrate the velocity distributions in the current-induced SCR (for $V_{BE}=1.0V$, $V_{BC}=0.0V$ or $V_{BC}=-2.0V$).

We note in the above figures that our model predicts a higher peak overshoot velocity than that yielded by either the energy transport model or Monte Carlo simulations (not

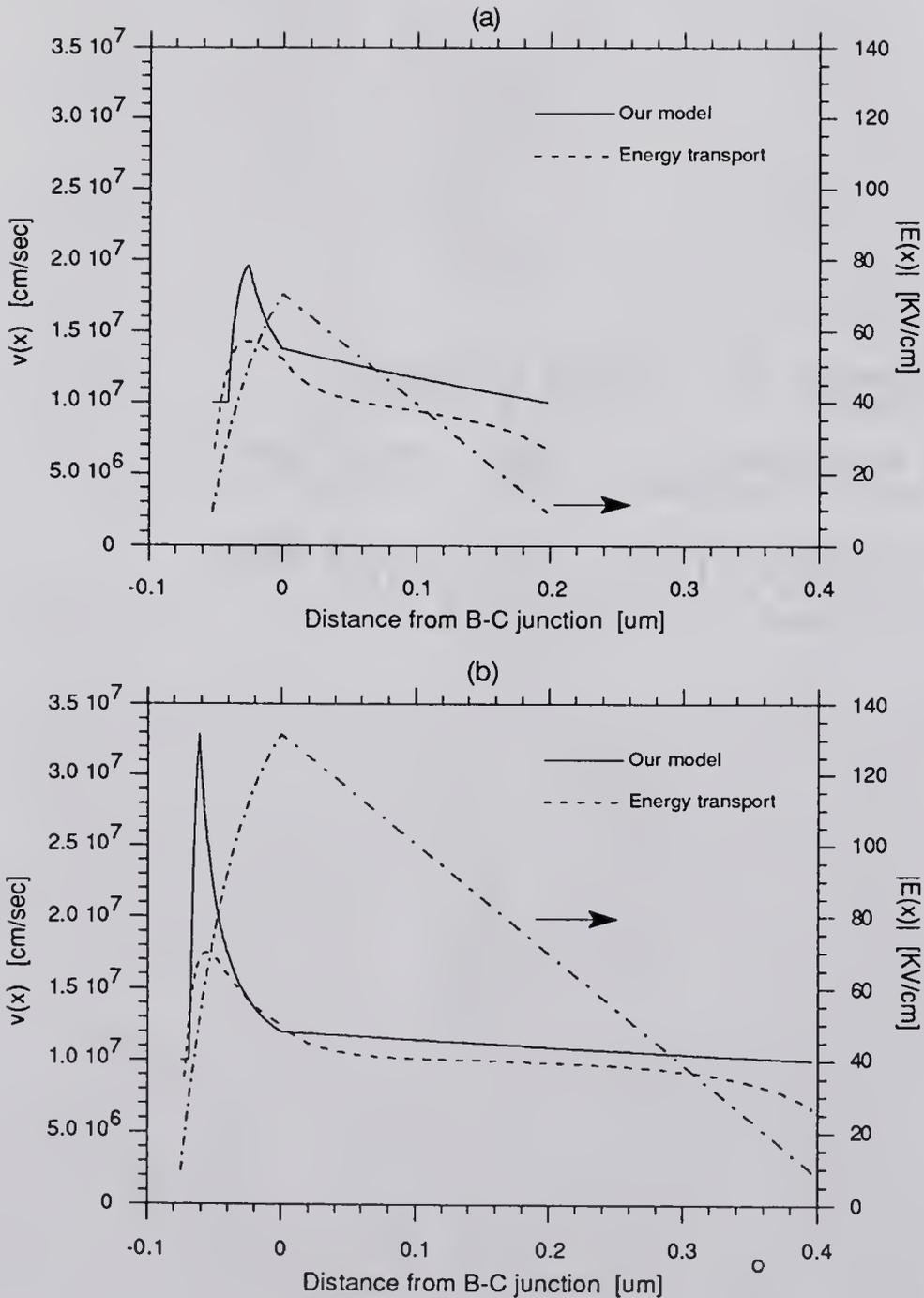


Fig. 4.5 Drift velocity and electric field in junction SCR:
 (a) $V_{BE}=0.7\text{V}$ and $V_{BC}=0.0\text{V}$; (b) $V_{BE}=0.7\text{V}$ and $V_{BC}=-2.0\text{V}$.

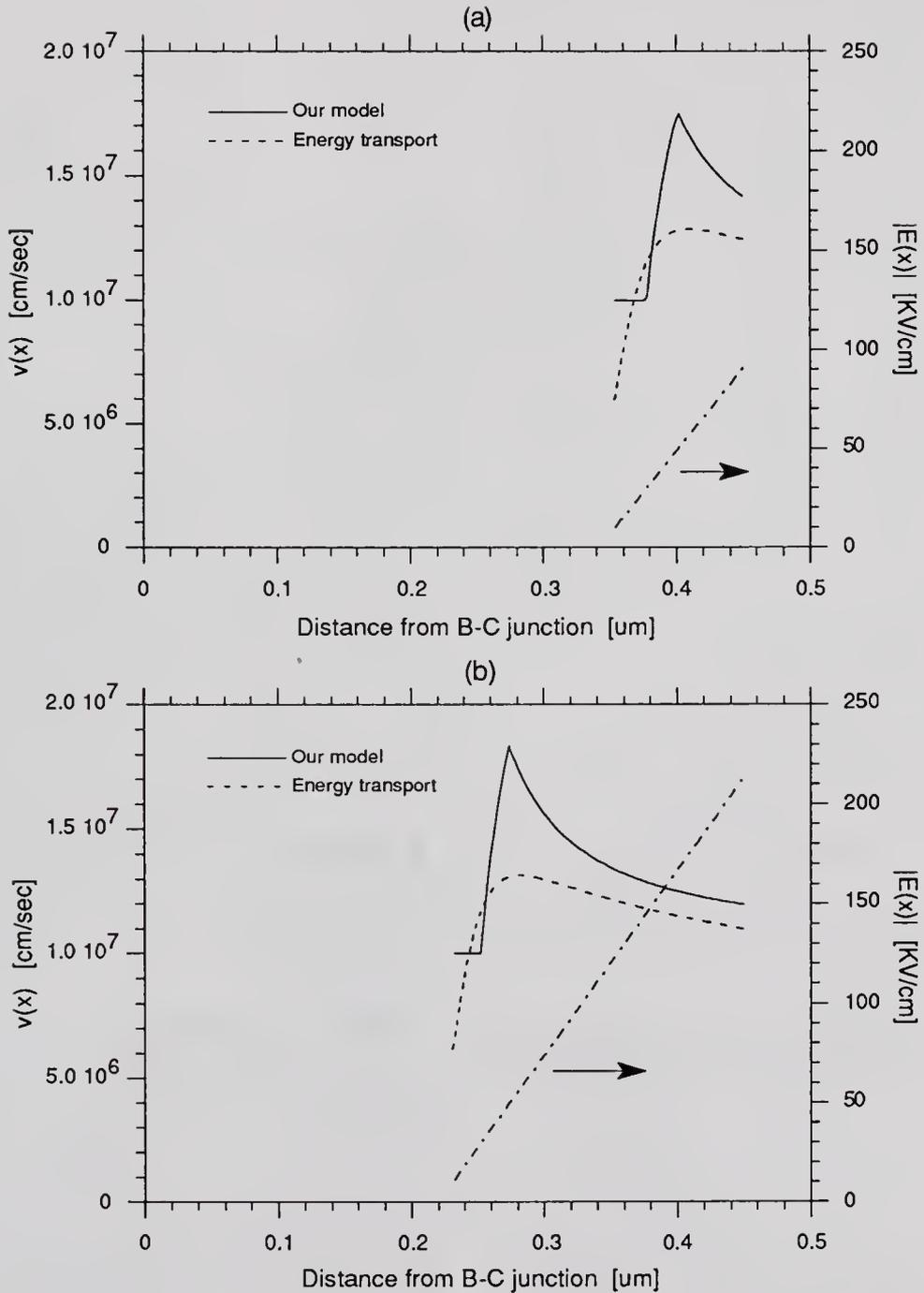


Fig. 4.6 Drift velocity and electric field in current-induced SCR: (a) $V_{BE}=1.0\text{V}$ and $V_{BC}=0.0\text{V}$; (b) $V_{BE}=1.0\text{V}$ and $V_{BC}=-2.0\text{V}$.

shown in the figures) [Prof. M. Lundstrom of Purdue University, private communication, 1991]. This discrepancy could mean that the length coefficient [Art88] which we used might be erroneous. Indeed when the high- $|E|$ saturation value of $L(E)$ in Fig. 4.1 is reduced from $4.5 \times 10^{-6} \text{cm}$ to $2.0 \times 10^{-6} \text{cm}$, which has been suggested [Art88], the MMSPICE-predicted velocity overshoot is in better agreement with that predicted by the energy transport and Monte Carlo analyses. This uncertainty in $L(E)$ can be attributed to the different set of transport parameters used. However we stress that the terminal characteristics of advanced BJTs predicted by our model, which will be shown later, agree quite well with results [Fus92] of measurements and numerical simulations based on a hydrodynamic model for energy transport. Conversely then, we note that the energy transport model has several uncertainties as well. It is based on several equivocal assumptions. For example, it assumes that the electron energy is entirely thermal. Also, the results depend on the degree of the energy transport equation, and there are still some uncertainties in the evaluation of the model parameters such as m^* , τ_p , and τ_w . Monte Carlo analysis is not unequivocal either. For example, detailed and accurate information about the numerous scattering parameters as well as needed details of the energy-band structure are lacking. With these deficiencies then, our model is

reasonable for first-order accounting of the electron velocity overshoot in circuit simulation, which has never been done before.

4.4 Implementation

The implementation of the electron velocity overshoot model in MMSPICE is based on a single iteration of the existing (conventional) model routine [Jeo89] for the (n^+pnn^+) BJT, as illustrated in Fig. 4.7. The analysis is done for each iteration of the circuit nodal analysis at each time step. With V_{BE} and V_{BC} passed in from the nodal analysis, the (one-dimensional) BJT model routine, which assumes a saturated drift velocity v_s (no overshoot) in the high-E epi-collector SCRs, is called to solve the conventional ambipolar transport, and characterize $E(x)$. Thus unlike empirical circuit models, the MMSPICE BJT model is susceptible to an extension to account for the augmented non-local carrier velocity distribution. From the predicted $E(x)$, combined with the length coefficient $L(E)$, the carrier velocity is evaluated depending on the SCR type (see Fig. 4.3). Then from $v[E(x)]$, the effective saturated drift velocity, $v_{s(eff)} > v_s$, is evaluated as described in Section 4.2.3.

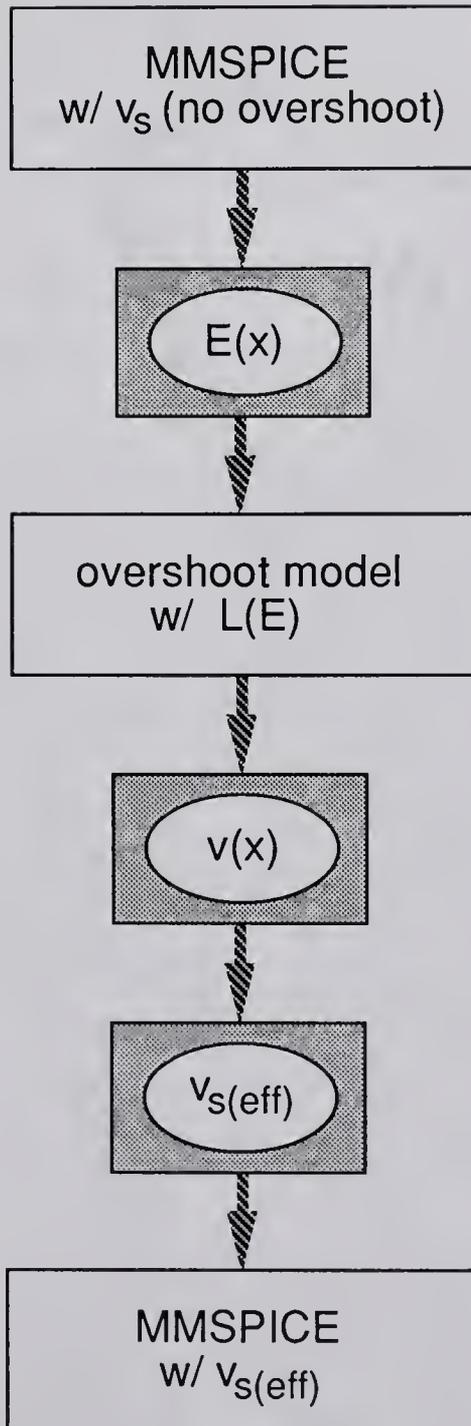


Fig. 4.7 Algorithm for implementation of velocity overshoot model in MMSPICE.

Once $v_{s(\text{eff})}$ is characterized, E_s is correspondingly updated as well to $E_{s(\text{eff})} \equiv v_{s(\text{eff})} / \mu_0$, which is higher than the preliminary value. Fig. 4.8 illustrates the resulting velocity-field model in the epi-collector SCRs used in MMSPICE, which we believe is suitable for first-order accounting of the velocity overshoot in circuit simulation. Note that the $v(E)$ slope (i.e., the low-field mobility μ_0) is not changed. Hence the solution obtained in regions where $|E| < E_s$ (reflected by μ_0) is not perturbed at all, which is physically appropriate.

With $v_{s(\text{eff})}$ and $E_{s(\text{eff})}$, the MMSPICE BJT model routine is called once more to effect the first-order accounting for the non-local transport in the predicted device currents and charges. The accounting for velocity overshoot, which is done here in a circuit simulator for the first time, is computationally efficient, and enables representative mixed-mode simulation for advanced bipolar technologies as we now demonstrate.

4.5 Simulations

In this section, MMSPICE device and circuit simulation results are presented to assess significance of the velocity overshoot effects in advanced silicon-based bipolar

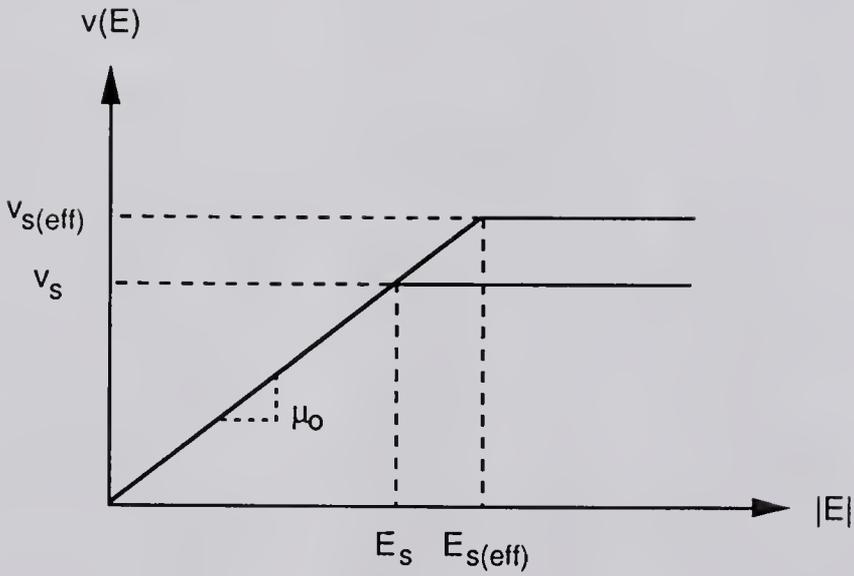


Fig. 4.8 Piecewise-linear velocity-field model. Effective saturated velocity is larger than the classical value because of overshoot. Note that the low- $|E|$ slope (μ_0) is not changed.

technologies, as well as to lend credibility to our modeling. The assumed BJT model parameters characterize a contemporary advanced silicon device structure: the actual emitter area is $3.2 \times 1.2 \mu\text{m}^2$; the peak base doping density is $1.5 \times 10^{18}/\text{cm}^3$; the metallurgical base width is $0.15 \mu\text{m}$; the epi doping density is $2.0 \times 10^{16}/\text{cm}^3$; and the width of the epi-collector is $0.45 \mu\text{m}$.

The MMSPICE-predicted results in Fig. 4.9 illustrate how the electric field and the corresponding electron velocity in the collector-region SCRs of the advanced BJT vary as V_{BE} is increased. These $v[E(x)]$ dependences define $v_{s(\text{eff})}$, as described in Section 4.2.3, for the final call of the model routine; and they underlie device/circuit performance simulations exemplified below. When the device is operating at low current in the forward-active mode, the base-collector junction SCR (Fig. 4.3(a)) exists as indicated for Case 1 ($V_{BE}=0.70\text{V}$, $V_{BC}=-2.0\text{V}$) in Fig. 4.9. The electric field (magnitude) is high, increasing abruptly near the metallurgical junction. Significant overshoot is predicted. As V_{BE} is increased, the electrons constituting the collector current modulate the SCR, decreasing $d|E|/dx$ on the collector-side, and the onset of quasi-saturation [Jeo89] (base pushout) is approached. The velocity overshoot is diminished because the (positive) gradient of the electric field is decreased as the collector-side SCR expands. As

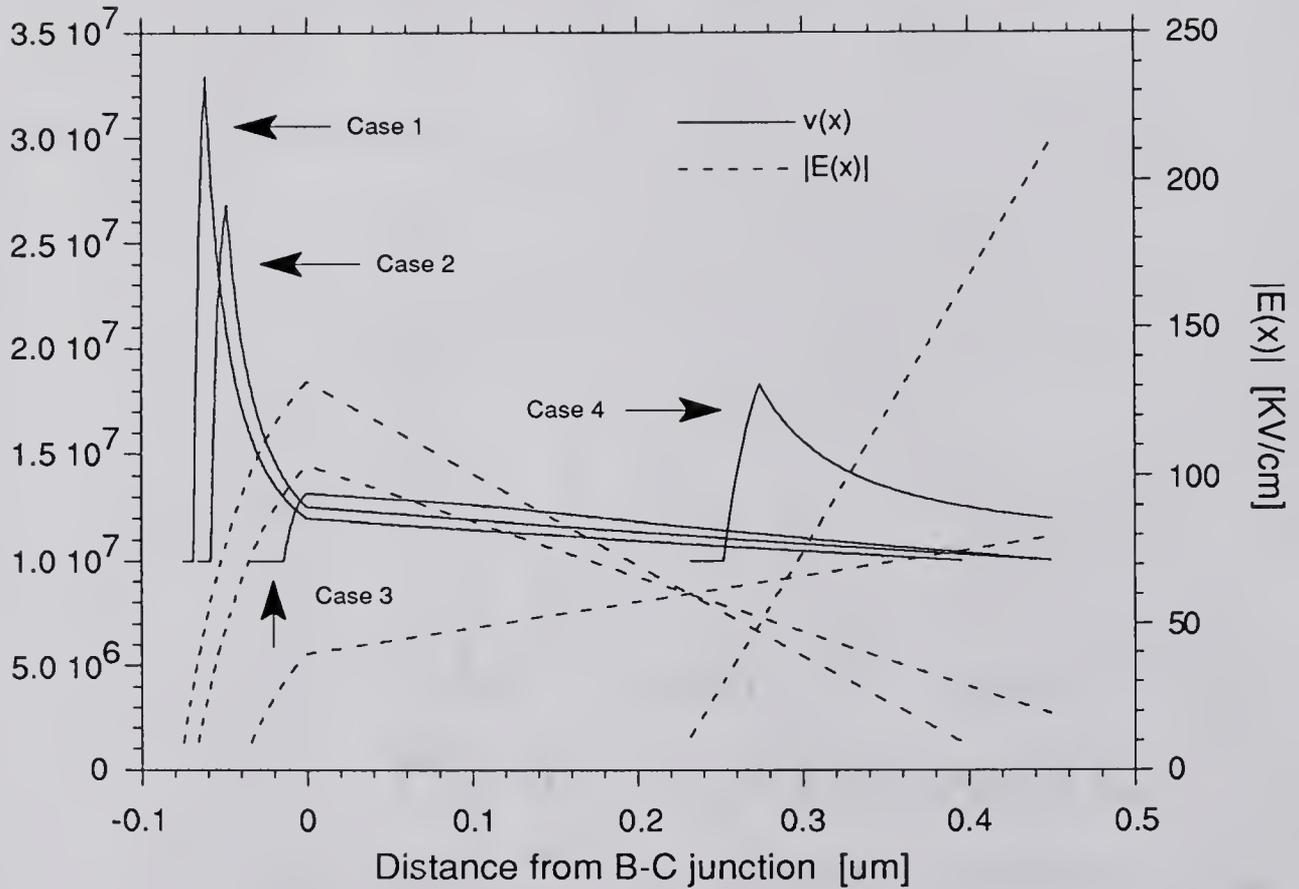


Fig. 4.9 MMS-PICE-predicted electron drift velocity and electric field variations in the epi-collector of an advanced BJT for $V_{BC} = -2.0V$. For each case, V_{BE} and $v_{s(eff)}$ are: 0.70V and 1.15×10^7 cm/sec (Case 1); 0.85V and 1.16×10^7 cm/sec (Case 2); 0.90V and 1.15×10^7 cm/sec (Case 3); 1.0V and 1.32×10^7 cm/sec (Case 4).

shown by Case 2 ($V_{BE}=0.85V$), eventually the SCR covers the entire epi-collector. Case 3 ($V_{BE}=0.90V$) shows that the sign of $d|E|/dx$ on the collector-side is reversed since the electron density is larger than N_{EPI} (see Fig. 4.3(c)). If V_{BE} is increased further, as in Case 4 ($V_{BE}=1.0V$), a high- $|E|$ current-induced SCR (Fig. 4.3(b)) is formed, forcing non-ohmic quasi-saturation (i.e., base pushout), and significant velocity overshoot begins to recur, with the region of overshoot pushed out toward the buried layer. These predictions do indeed agree in principle with results of numerical simulations, including the hydrodynamic equations, of advanced BJTs [Fus92]. One point should be mentioned; our other simulation results, although not shown here, reveal that $v_{s(eff)}$ is, in accord with Das' work [Das91], decreased with the reverse bias on the base-collector junction, because the electric field profile changes due to the increase in the SCR width, which reduces the importance of the initial velocity overshoot.

Fig. 4.10 shows the MMSPICE-predicted dc current gain ($\beta(V_{BE})$) of the same BJT, simulated with and without the velocity overshoot accounted for. Even though the overshoot effect becomes noticeable with increasing reverse bias on the base-collector junction at high V_{BE} , it is relatively insignificant with regard to β for this BJT. Simulated high-

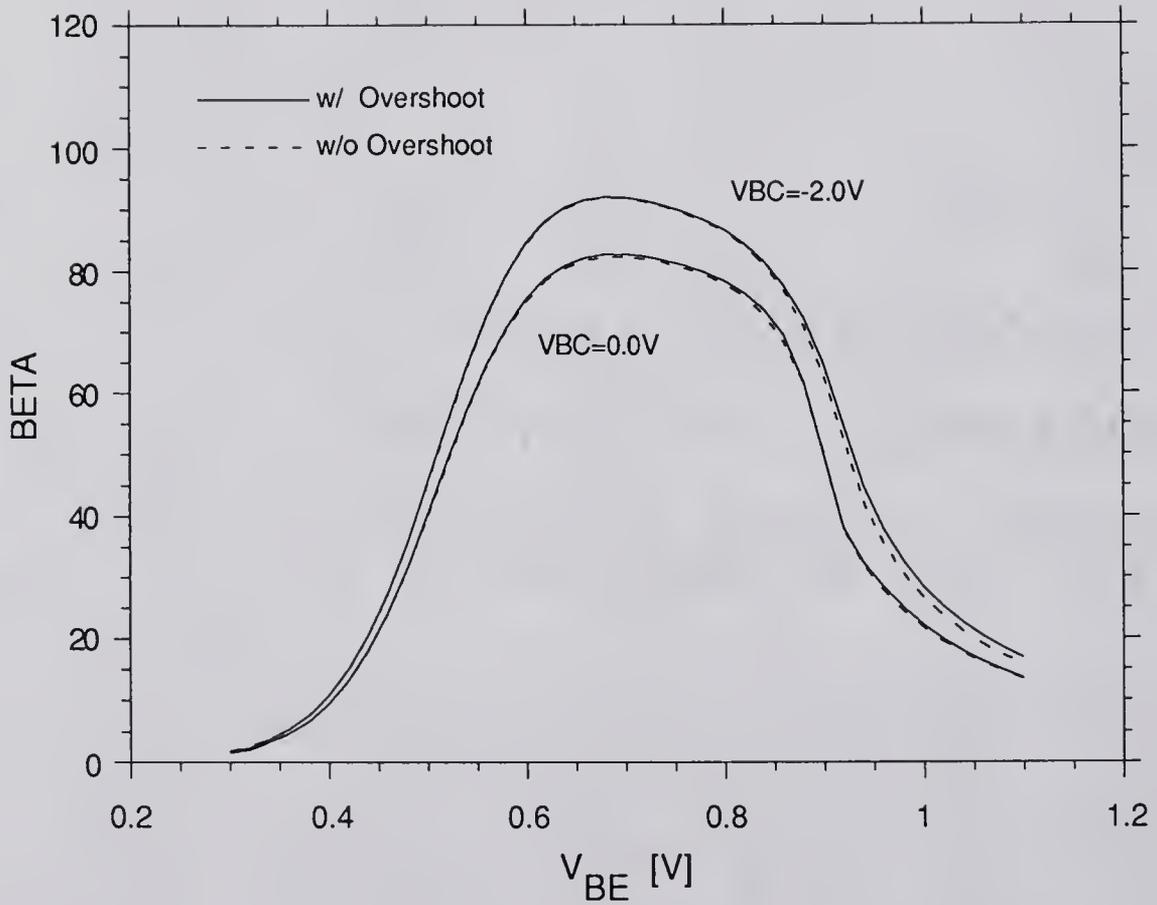


Fig. 4.10 MMSPIICE-predicted dc current gain versus V_{BE} for the BJT for $V_{BC} = 0.0V$ and $-2.0V$, with and without velocity overshoot accounted for.

injection collector current densities ($J_C(V_{CE})$ versus I_B) are shown in Fig. 4.11. These predictions of the velocity-overshoot effect on J_C show excellent agreement with results of measurements and numerical simulations [Fus92]. For the relatively low $I_B=20\mu A$, the overshoot effect is not noticeable for low V_{CE} because ohmic quasi-saturation [Jeo89] prevails in this case; viz., the electric field is low. Although the overshoot effect does become apparent in this case at higher V_{CE} when the entire epi-collector is space-charged, it is not a strong effect since the field gradients are not extremely high. For higher fixed I_B though, the overshoot effect becomes more significant with increasing V_{CE} . In this case, the prevalent current-induced SCR (Fig. 4.3(b)) is being enlarged and concomitantly the electric field gradient is increasing, thereby increasing $v_{s(eff)}$ and counteracting the J_C limitation that normally results from quasi-saturation and base pushout [Jeo89]. For fixed V_{CE} where the current-induced SCR exists, the overshoot effect becomes more noticeable as I_B increases because of the increased $d|E|/dx$ in the SCR which increases $v_{s(eff)}$ and delays the base pushout (see Fig. 4.9). For the high $I_B=160\mu A$ at $V_{CE}=3.0V$, J_C is underestimated by about 5% when the velocity overshoot is not accounted for.

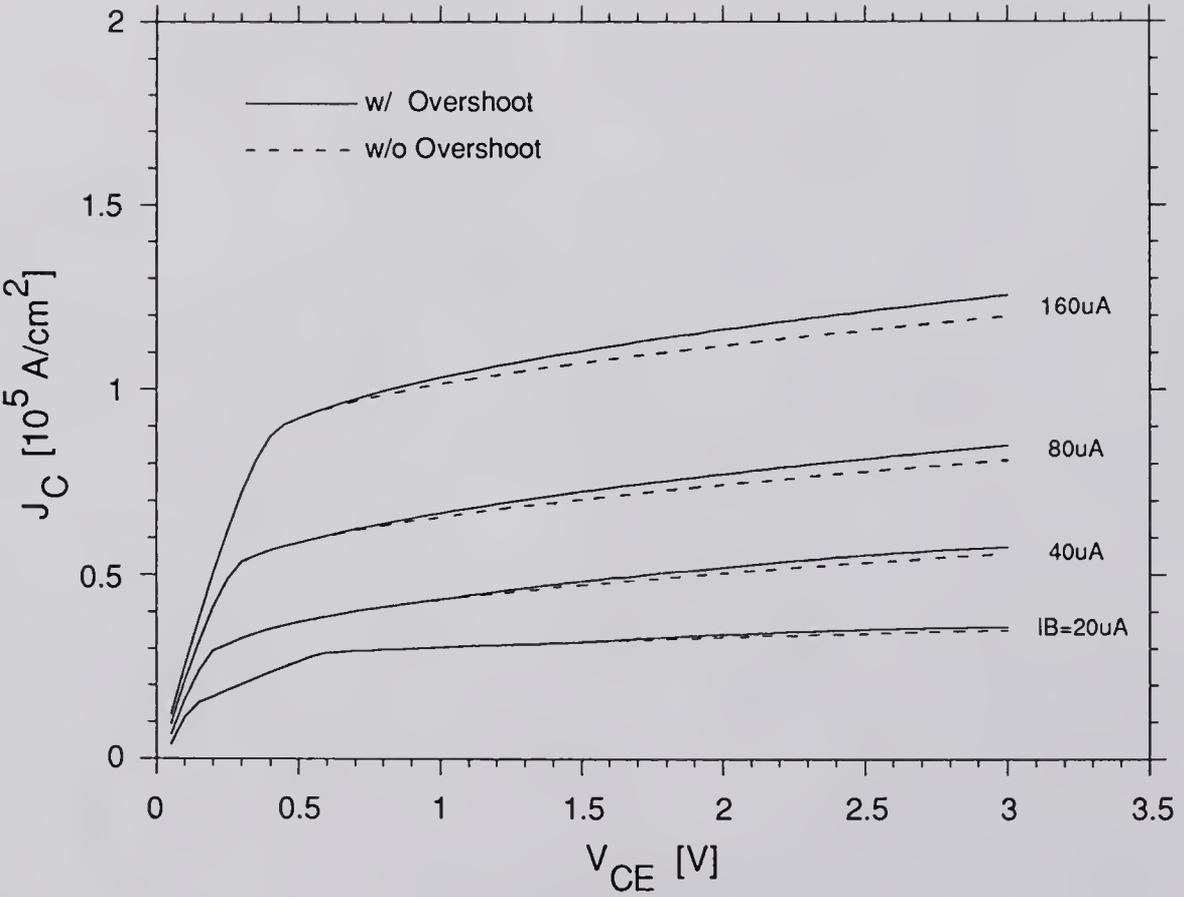


Fig. 4.11 MMSPICE-predicted current-voltage characteristics of the BJT, with and without velocity overshoot accounted for.

From the previous results, it can be inferred that the velocity overshoot would reduce the signal delay by suppressing base pushout [Fus92]. In Fig. 4.12, the base pushout factors are plotted versus the collector current density. The base pushout factor is defined as the ratio of the widened base width to the nominal base width. Obviously, base pushout was suppressed with the overshoot effect; when $J_C=0.5 \times 10^5 \text{A/cm}^2$ for example, the conventional model predicts about 25% of base pushout, but according to the overshoot model, the (nominal) quasi-neutral base region is not widened yet. Since base pushout is suppressed by the overshoot effect, it can be exploited to enhance the performance of the scaled devices. In fact, it can lower the otherwise high collector doping concentration, if the other parameters are carefully optimized.

The influence of the velocity overshoot on performance is illustrated in Fig. 4.13 where the MMSPIICE-predicted cutoff frequency (f_T) versus J_C is plotted for $V_{CE}=3.0\text{V}$. The velocity overshoot increases the onset value of J_C for base pushout, thereby yielding higher f_T at high currents. For this contemporary BJT, this effect would allow about an 80% higher operating current, or a lower epi doping density [Fus92] (for lower collector-base junction capacitance and higher breakdown voltage), for the same performance as that

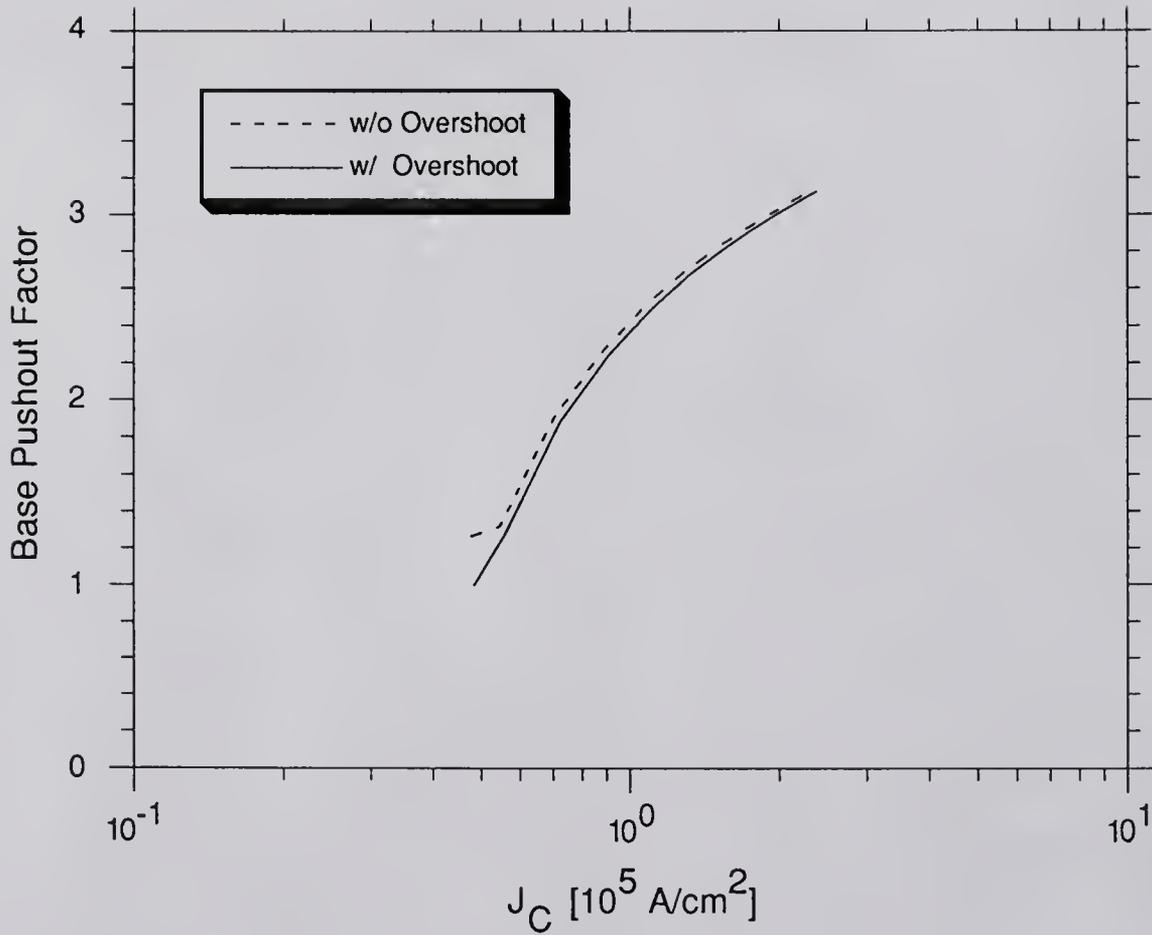


Fig. 4.12 MMSPICE-simulated base pushout factor as a function of J_C when $V_{BC}=-1.5\text{V}$, with and without velocity overshoot accounted for.

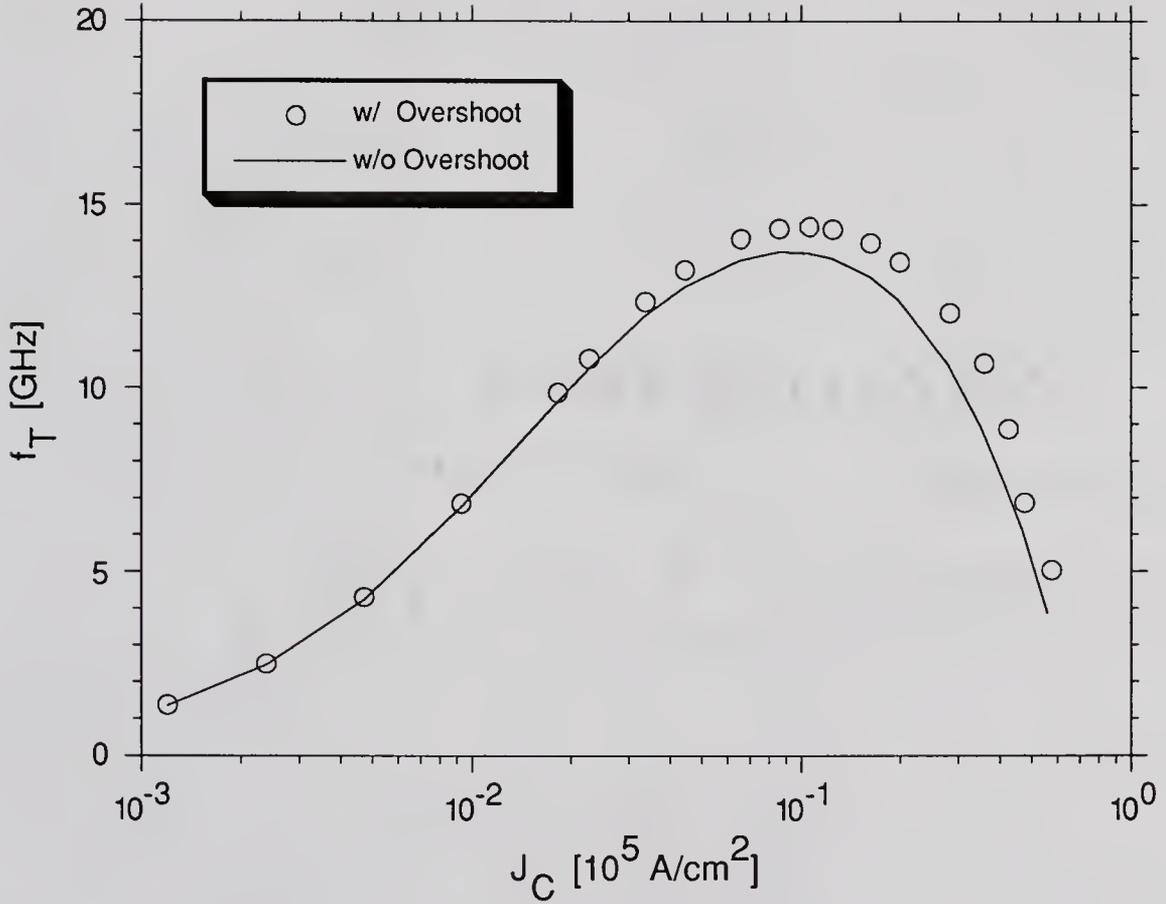


Fig. 4.13 MMSPIICE-predicted cutoff frequency versus current density for the BJT for $V_{CE}=3.0V$, with and without velocity overshoot accounted for.

predicted without velocity overshoot.

In order to assess the significance of the electron velocity overshoot in the BJT in actual circuit performance, MMSPICE transient simulations of a single-stage ECL gate comprising four BJTs (see Fig. 3.12 for the circuit diagram), each characterized as previously described, were done. For a (high) gate current (I_{gate}) of 0.9mA ($J_C=0.2 \times 10^5 \text{A/cm}^2$) and a logic swing of 0.5V, switching voltage transients of the circuit predicted with and without the overshoot effect are plotted in Fig. 4.14(a). With the overshoot accounted for, the ECL propagation delay is reduced because of less base pushout. However the benefit is small; the velocity overshoot produces only about a 5% speed enhancement. Fig. 4.14(b) illustrates average propagation delay versus the gate current. As high-current conditions prevail, the delay is increased due to base pushout. If the velocity overshoot is accounted for, the delay is decreased as discussed before. However the benefit of overshoot is reduced as the region of overshoot is pushed out toward the buried layer.

To further investigate the influence of the overshoot in circuit applications, a BiCMOS driver was also simulated. Fig. 4.15(a) shows the circuit diagram of the driver; the gate length of the nominal MOSFETs is $1.0\mu\text{m}$ and two BJTs have $1.2\mu\text{m}$ emitter widths. The switching characteristics of the

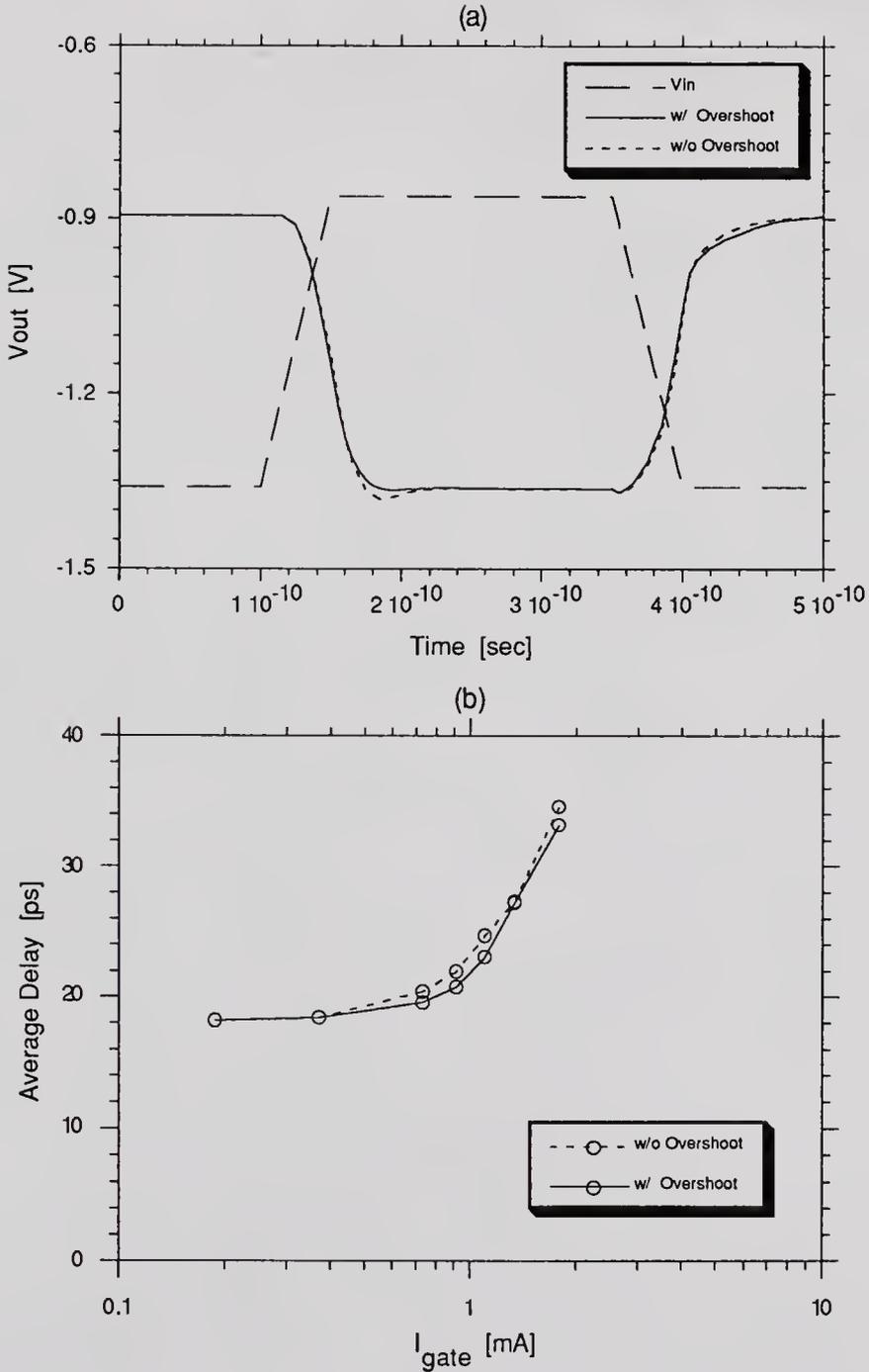


Fig. 4.14 MMSPIICE-simulated switching voltage transients of an ECL gate ($I_{gate}=0.9\text{mA}$) in (a) and average propagation delay as a function of I_{gate} in (b), with and without velocity overshoot accounted for.

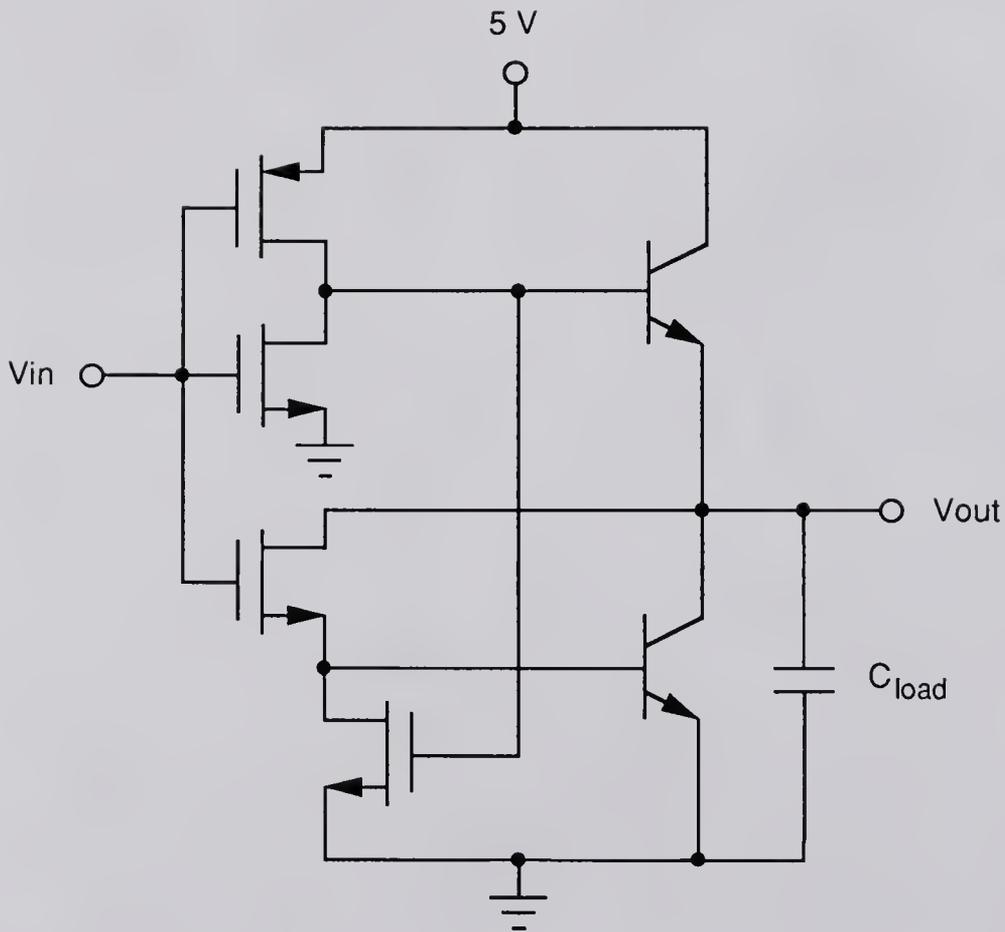


Fig. 4.15(a) Circuit diagram of a BiCMOS driver; the gate length of the nominal MOSFETs is $1.0\mu\text{m}$.

circuit with a 5.0V supply voltage, simulated with and without the overshoot accounted for, are shown in Fig. 4.15(b). With a capacitive load (C_{load}) of 2pF, the average propagation delay is reduced by about 4% by the electron velocity overshoot in the epi-collector of the BJTs. The average propagation delays versus C_{load} are plotted in Fig. 4.15(c). With the overshoot effect, the propagation delay is decreased by about 4% on the average. (For these simulations the overshoot effect was accounted for only in the BJTs.)

In the transient simulations discussed, the transient field (dE/dt) dependence of the electron drift velocity [Bla90], which might be significant during fast transients, was not accounted for. (In fact, this non-stationary effect in scaled BJTs could be accounted for via the methodology described in [Jin92a].) Because of the usual insignificance of non-quasi-static effects in contemporary bipolar digital circuits, we do not believe that this transient overshoot is as significant as the quasi-static one which we modeled.

The velocity overshoot model, implemented in MMSPICE, increases the circuit/device simulation time mainly because of the additional call of the BJT model routine, but also because of the numerical integration needed to evaluate the effective saturated drift velocity and higher iteration counts required for convergence to time-point solutions of

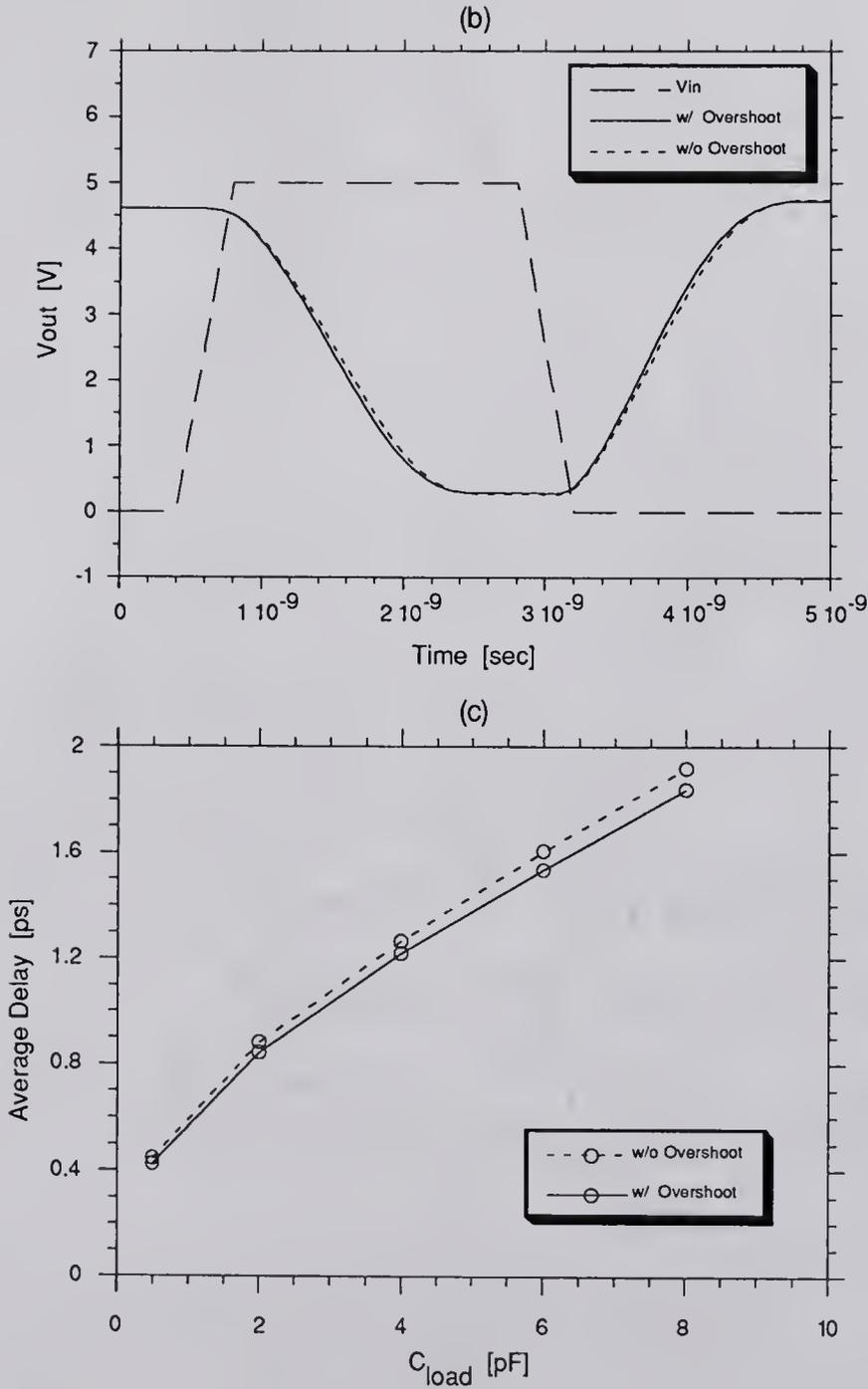


Fig. 4.15 MMSPICE-simulated switching voltage transients of the BiCMOS driver ($C_{load}=2\text{pF}$) in (b) and average propagation delay as a function of C_{load} in (c), with and without velocity overshoot accounted for.

the circuit nodal analysis. For the ECL transient simulations of Fig. 4.14, the run-time is increased by about 40%. Nonetheless the tool seems to effect a very good trade-off between accuracy and computational intensiveness in accounting for velocity overshoot in bipolar transistor and circuit simulation, which has not been done previously.

4.6 Summary

A physical model for electron velocity overshoot in advanced silicon BJTs has been developed and implemented into a circuit simulator (MMSPICE) for the first time. The model, which is based on an augmented drift-velocity formalism, assumes a piecewise-linear length coefficient derived from Monte Carlo simulations. The associated relaxation of the carrier velocity is characterized phenomenologically to be consistent with the overshoot analysis. The model characterizes an effective saturated drift velocity, larger than the classical value due to overshoot, for each iteration of the circuit nodal analysis at each time step. This first-order semi-numerical accounting for non-local velocity overshoot was shown to be amenable to representative and computationally efficient circuit and device simulation. Indeed it can be especially useful for predictive mixed-mode

simulation for bipolar and BiCMOS VLSI technology/manufacturing CAD.

From MMSPICE simulation results, we conclude that electron velocity overshoot effects in contemporary silicon BJTs produce only small performance enhancements, but can be exploited to optimize design trade-offs in scaled technologies.

CHAPTER 5 MMSPICE-2 DEVELOPMENT

5.1 Introduction

In the preceding chapters, the modeling of multi-dimensional currents, current crowding, and velocity overshoot in advanced BJTs has been presented. To utilize the models for mixed-mode device/circuit simulation, we implemented them into MMSPICE [Jeo90], which resulted in the creation of "MMSPICE-2". In this chapter, the issues related to the implementation are discussed.

The previously developed MMSPICE (version 1) is a physical alternative to Gummel-Poon/SPICE, containing a charge-based semi-numerical BJT model based on one-dimensional regional analyses involving technological parameters that pertain to doping profile and to physical models that depend on doping. The one-dimensional model physically accounts for high-current effects, impact ionization, and non-reciprocal (trans)capacitances. Although this physics-based model is sufficient for many applications, the utility of MMSPICE is enhanced herein by adding user options to account for the multi-dimensional currents, the dc/transient current crowding, and the electron velocity

overshoot. Furthermore, the extrinsic collector-base junction capacitance (charge) is added. In fact, this capacitance has a predominant effect on the performance of scaled devices because the extrinsic base region is not reduced proportionally as the intrinsic device is scaled down. The collector-substrate capacitance (charge) is also included.

In Section 5.2, the new features of MMSPICE-2 are presented, with descriptions of the new parameters. The models are implemented into MMSPICE so that users may activate any combination of them by option. In Section 5.3, discussion of the evaluation of the new parameters is given. In Section 5.4, the modifications of MMSPICE source code, which are needed to produce MMSPICE-2, are discussed with the aid of the software-structure flowcharts. Finally in Section 5.5, MMSPICE-2 simulation results exemplifying the new features are presented to demonstrate the utility of the simulator in bipolar TCAD applications.

5.2 New Features

5.2.1 Multi-dimensional Currents

As described in Chapter 2, the most important peripheral current in the advanced BJT structure is the component of

base current due to lateral injection of electrons from the emitter sidewall. The peripheral base current I_{BP} was modeled as [eq. (2.2)]

$$I_{BP} \equiv 2 \cdot JEOP \cdot (WE + LE) \left[\exp\left(\frac{V_{BE}}{NEBP \cdot V_T}\right) - 1 \right] \quad (5.1)$$

where the new parameters JEOP and NEBP represent the peripheral saturation current density (per unit length) and the peripheral emission coefficient respectively. If JEOP is not specified on the model card, the lateral injection current is neglected.

5.2.2 Current Crowding

In MMSPIICE-1, a semi-empirical accounting for current crowding was effected by defining the intrinsic base resistance as a function of the current-dependent base charge. This was done through the parameter CRBI, the intrinsic base resistance coefficient [Jeo90]. However in order to characterize the crowding properly, we have to use the physical model developed in Chapter 3. The new model uses the majority-hole mobility in the intrinsic base region, which is needed to evaluate the specific base resistivity. Hence we introduce a new model parameter, UPBASE, the effective hole mobility (for the npn device).

To activate the dc/transient crowding analysis in MMSPICE-2, CRBI is used as a flag. If $CRBI > 0$, the old accounting is accessed. If $CRBI = 0$ is specified, then the new crowding analysis is activated, using either the user-specified UPBASE or a default value ($= 230 \text{ cm}^2 / \text{V} \cdot \text{sec}$). In Chapter 3, it was revealed that the switch-off transient crowding is usually negligible for contemporary BJTs. Hence in MMSPICE-2, a sub-option to skip the crowding analysis during the switch-off transient is employed for reduction of computation time and for potential improvements in convergence. If $CRBI < 0$ on the model card, $v_{BE(\text{eff})} = v_{BE}$ is assumed during the switch-off transient, while the regular crowding analysis is done during the switch-on transient. Note that when CRBI is not specified in the model parameters, a nonzero default value ($= 2 \times 10^{-4} \text{ V} \cdot \text{sec} / \text{cm}^2$) is assumed, and the original (crude) analysis is used.

One point should be mentioned regarding the device-card parameters. As described in Chapter 3, the crowding analysis is done along the shorter dimension of the rectangular emitter because the predominant base current flow is along that direction. Therefore, WE on the device card should be less than LE ($= AE / WE$) for proper accounting of the crowding effect.

5.2.3 Velocity Overshoot

Accounting for the velocity overshoot effects in simulations does not require any additional parameters. Instead, the conventional saturation velocity V_S is used as a flag. If $V_S=0$ on the model card, the overshoot analysis is activated to evaluate the effective saturated drift velocity $v_{s(\text{eff})}$ as described in Chapter 4. The initial value needed to evaluate $v_{s(\text{eff})}$ is assumed to be $1.0 \times 10^7 \text{cm/sec}$. If $V_S \neq 0$ is specified, the velocity overshoot effect is not accounted for at all, and V_S is used in the model analyses. Note that when V_S is not specified, a default value of V_S ($=1.0 \times 10^7 \text{cm/sec}$) is assumed, and the classical analysis is done without the overshoot effect accounted for.

5.2.4 Extrinsic Collector-base Capacitance

In order to account for the parasitic collector-base capacitance $C_{cb(\text{ext})}$, the depletion charge $Q_{JC(\text{ext})}$ in the extrinsic collector-base region is modeled. In some cases this charge in the advanced BJT can be approximated as

$$Q_{JC(\text{ext})} = q \cdot A_{LBE} \cdot NEPI \cdot X_d \quad (5.2)$$

where A_{LBE} is the area of the extrinsic base (p^+) region and X_d is the width of the epi-collector-side depletion region. We assume that any additional charge stored in the p^+ - n sidewall junction is implicitly included in the above expression. In the case of a rectangular emitter, A_{LBE} is expressed as

$$\begin{aligned} A_{LBE} &= (WE + 2LBE) \cdot (LE + 2LBE) - WE \cdot LE \\ &= 2(WE + LE) \cdot LBE + (2LBE)^2 \end{aligned} \quad (5.3)$$

where LBE is a new model-card parameter representing the peripheral width of the extrinsic base region, which is assumed constant all around the emitter periphery. If LBE is not specified or is given as 0, $C_{cb(ext)}$ is not accounted for.

Assuming a one-sided step junction, we approximate

$$X_d \approx \left[\frac{2\epsilon_s}{q \cdot NEPI} (PC - V_{BC}) \right]^{1/2} \quad (5.4)$$

where PC is the built-in potential of the (intrinsic) base-collector junction, which is an original model parameter of MMSPICE-1, as is NEPI. In a strict sense, the built-in potential at the extrinsic base-collector junction is, due to the higher base doping density, greater than PC, but this discrepancy is not significant and is ignored.

To this point, we have evaluated $Q_{JC(ext)}$ assuming that X_d is less than the width of the epi-collector W_{EPI} . In realistic situations however, the entire epi layer (under the extrinsic base), which is relatively thin due to the deep p^+ junction, can become completely space-charged in normal operation of the BJT. In this case, the depletion charge stored in the buried layer, Q_{BL} , must also be taken into account because of the charge neutrality in the $p^+ - n - n^+$ region:

$$Q_{JC(ext)} = q \cdot A_{LBE} \cdot N_{EPI} \cdot W_{EPI} + Q_{BL}(V_{BC}), \quad \text{for } V_{BC} < V_{BCWEPI} \quad (5.5)$$

where the critical voltage V_{BCWEPI} at which the entire epi region becomes space-charged can be approximated from $X_d \approx W_{EPI}$ in (5.4) as

$$V_{BCWEPI} \approx PC - \frac{q \cdot N_{EPI} \cdot W_{EPI}^2}{2\epsilon_s} \quad (5.6)$$

The depletion charge Q_{BL} is evaluated from Gauss' law:

$$Q_{BL} = \epsilon_s \cdot E_{BL} \cdot A_{LBE} \quad (5.7)$$

where E_{BL} is the magnitude of the electric field at the epi-buried-layer junction. With the assumption that the integral of the electric field across the entire base-collector region

equals $PC - V_{BC}$, E_{BL} can be expressed in terms of V_{BC} by applying Poisson's equation to the junction SCR:

$$E_{BL} = \frac{PC - V_{BC} - \frac{q}{2\epsilon_s} NEPI \cdot WEPI^2}{WEPI} \quad . \quad (5.8)$$

Note then that (5.7) is identical to the induced charge on a parallel plate capacitor when the voltage is incremented by $V_{BC} - V_{BCWEPI}$:

$$Q_{BL} = \frac{\epsilon_s}{WEPI} A_{LBE} |V_{BC} - V_{BCWEPI}| \quad . \quad (5.9)$$

Once $Q_{JC(ext)}$ is evaluated by either (5.2) for $V_{BC} \geq V_{BCWEPI}$ or (5.5), it is added algebraically to Q_{BC} in MMSPICE-2:

$$Q_{BC} \rightarrow Q_{BC} - Q_{JC(ext)} \quad , \quad (5.10)$$

which then implicitly accounts for the effect of $C_{cb(ext)}$. Note that Q_{BC} includes components of the total hole charge in the base that neutralize electron charge communicating with the collector [Jeo89], e.g., $Q_{JC(ext)}$.

5.2.5 Substrate Capacitance

The charge storage associated with the collector-substrate junction is simply accounted for via a lumped capacitance C_{sub} , as in the SPICE diode model, that depends on the substrate-collector bias V_{SC} :

$$C_{\text{sub}} = \frac{C_{\text{JS}} \cdot \text{ABL}}{\left(1 - \frac{V_{\text{SC}}}{\text{PS}}\right)^{\text{MS}}} \quad (5.11)$$

where C_{JS} is the zero-bias collector-substrate capacitance, PS is the built-in potential at the substrate junction, and MS is the junction grading coefficient, all of which are old model parameters. The new parameter ABL represents the area of the buried (n^+) layer in the BJT structure.

5.3 Parameter Evaluation

The new parameters of MMSPICE-2 are categorized into two groups according to their nature, as summarized in Tables 5.1 and 5.2. Note that ABL is a device-card parameter related to the layout whereas NEBP , JEOP , UPBASE , and LBE are model-card parameters dependent on the technology. The device parameter can be evaluated directly from layout. The model parameters

TABLE 5.1
NEW DEVICE PARAMETER OF MMSPICE-2

Name	Description	Units	Default
ABL	Area of buried (n ⁺) layer	m ²	AC

TABLE 5.2
NEW MODEL PARAMETERS OF MMSPICE-2

Name	Description	Units	Default
NEBP	Peripheral Emission Coefficient	-	2.0
JEOP	Peripheral Sat. Current Density	A/m	0.0
UPBASE	Hole mobility in Intrinsic Base	cm ² /V·sec	230
LBE	Width of Extrinsic Base	m	0.0

associated with the peripheral base current are easily extracted as described in Chapter 2 and Appendix A.

However, the evaluation of UPBASE, which is a key parameter in the crowding analysis, is not straightforward because base conductivity modulation is not uncommon, implying that the effective hole mobility should be an implicit function of biases. One possible choice is to define a (constant) average hole mobility valid at all injection levels:

$$\text{UPBASE} = \overline{\mu_p} \equiv \frac{\int_0^{\text{WBM}} \mu_{pb}(x) p_b(x) dx + \mu_{pc} \int_{\text{WBM}}^{\text{WBM} + \text{WQNR}} p_c(x) dx}{\int_0^{\text{WBM}} p_b(x) dx + \int_{\text{WBM}}^{\text{WBM} + \text{WQNR}} p_c(x) dx} \quad (5.12)$$

where $p_b(x)$ and $p_c(x)$ are the majority- and minority-hole concentrations in the metallurgical and widened base regions respectively; $\mu_{pb}(x)$ and μ_{pc} are the corresponding hole mobilities as functions of the doping density, which are available from common sources (μ_{pc} in the epi-collector is assumed to be a constant, neglecting carrier-carrier scattering.); W_{QNR} is the width of the extended base region.

Since the hole distributions are available from the MMSPICE model solution, the average hole mobility can also be

evaluated. A study on this area is suggested as a future work.

5.4 Model Implementation

The new features were implemented in MMSPICE-2 via some modifications of the source code. The modifications comprise two new model subroutines, CROWD and OVERSHOOT, and minor changes in seven of the original routines (ELPRNT, ERRCHK, MODCHK, QBBJT, QBCT, READIN, and TMPUPD) in MMSPICE-1 to accept the new models. (The changes in ELPRNT, ERRCHK, READIN, and TMPUPD are trivial, and therefore they are not discussed here. Details regarding these minor changes are in "Programmer's Reference Manual of MMSPICE-2" [Jin92b].)

5.4.1 Subroutine Modifications

5.4.1.1 Subroutine MODCHK

MODCHK, which performs a pre-processing of device model parameters, is modified for the velocity overshoot analysis. As described before, the overshoot effect is accounted for only when $VS=0$ on the model card. In this case, the secondary model parameters which are functions of VS cannot be defined. To resolve this problem, MODCHK is modified as

indicated in Fig. 5.1. The highlight of the modification is to define a flag (IFLAG_VS): when VS=0, the list element NODPLC(LOC+3), which is transferred to QBBJT and QBCT as IFLAG_VS, is set to 1 to activate the velocity overshoot analysis. Also, VS is set to the default value of 1.0×10^7 cm/sec to evaluate the VS-dependent parameters. Then, the ordinary analysis follows.

5.4.1.2 Subroutine OBBJT

All the new features of MMSPICE-2 are implemented so that users may activate any combination of them by option. Two subroutines are added for the current crowding and velocity overshoot models. In order to link these new routines to the nodal analysis, the subroutine QBBJT is modified as shown in Fig. 5.2. First, $v_s(\text{eff})$ and $v_{BE}(\text{eff})$ are set to VS and v_{BE} respectively to make the routine generally applicable. Then, the value of CRBI on the model card is checked. If $CRBI \leq 0$, the crowding analysis (in CROWD) is activated to define $v_{BE}(\text{eff})$ on the emitter-base junction for the given v_{BE} ; otherwise, the analysis is skipped. Similarly, IFLAG_VS specified in MODCHK is monitored. If IFLAG_VS=1, the overshoot analysis (in OVERSHOOT) is turned on to evaluate $v_s(\text{eff})$ for a given effective junction bias $v_{BE}(\text{eff})$.

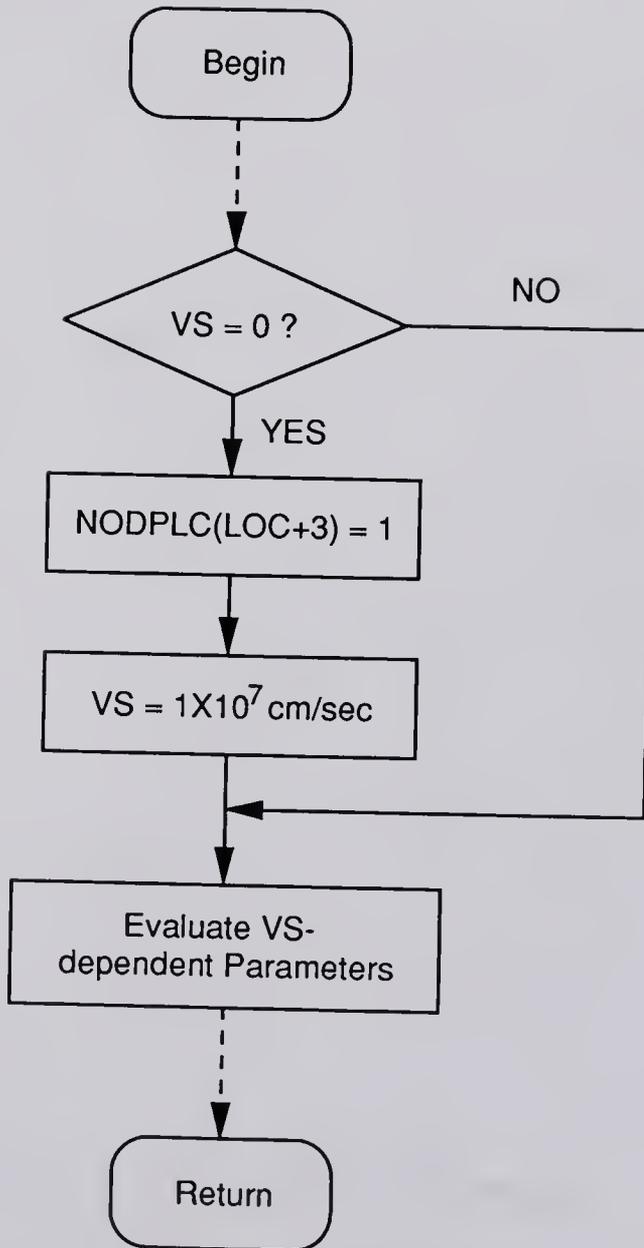


Fig. 5.1 Velocity overshoot algorithm of subroutine MODCHK.

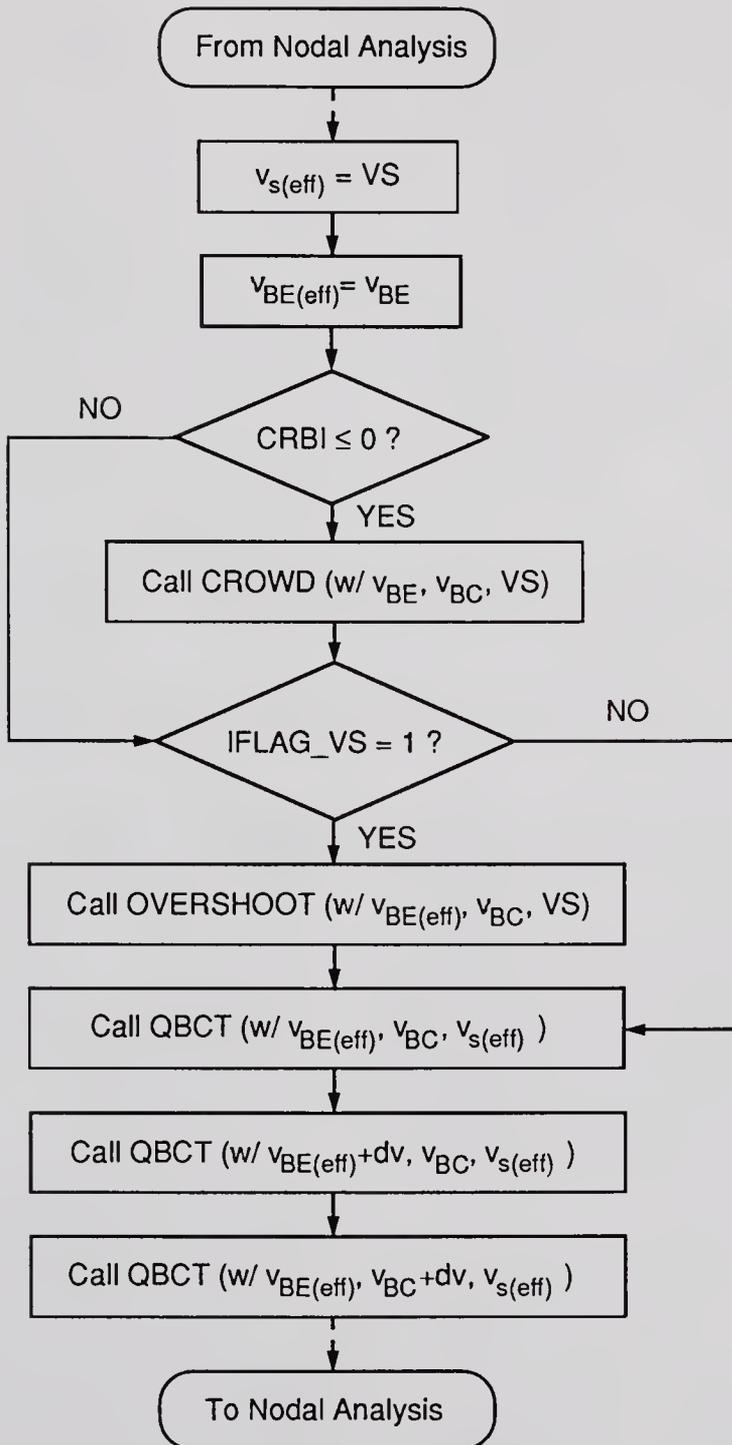


Fig. 5.2 Subroutine QBBJT modified for current crowding and velocity overshoot.

Note that $v_{BE(eff)}$ would be just v_{BE} if the crowding analysis is not done. If $IFLAG_VS \neq 1$, the overshoot analysis is skipped; i.e., $v_{s(eff)} = VS$ for the QBCT calls.

With $v_{BE(eff)}$ and $v_{s(eff)}$, the MMSPICE model routine (in QBCT) is called again to obtain the nominal bias-point solution, with or without the current crowding and/or velocity overshoot effects accounted for. Note that the evaluation of $v_{BE(eff)}$ and $v_{s(eff)}$ is done only at the nominal bias point; they are not updated, although $v_{BE(eff)}$ and v_{BC} are perturbed (for subsequent calls of QBCT) to calculate the derivatives of the currents and charges (via difference approximations), which are needed in the circuit nodal analysis. Then, the ordinary analysis follows, including the calculation of the substrate capacitance in terms of ABL.

If $CRBI \leq 0$ and $VS = 0$ on the model card, both the current crowding and velocity overshoot effects are accounted for. According to the current implementation however, the crowding analysis is done at first assuming no overshoot, and then the velocity overshoot analysis is activated with a new $v_{BE(eff)}$. In fact, there need not be a specific order in doing these analyses; what is most important is both effects are accounted for in each iteration of the circuit nodal analysis.

5.4.1.3 Subroutine QBCT

The main model routine QBCT, which calculates the operating-point currents and regional charges of the BJT, is also modified so that it may properly process $v_{BE(eff)}$ and $v_{s(eff)}$ passed in from the QBBJT routine. Fig. 5.3 illustrates the highlight of the modification. Previously, the electric field was evaluated only when the impact ionization analysis is needed (i.e., when $AI \neq 0$ and $BI \neq 0$). But now, when $IFLAG_VS=1$, the field is evaluated to characterize the overshoot effect. Note that the saturated velocity-dependent parameters are also updated when the overshoot analysis is on. Also, the peripheral base current and extrinsic collector-base junction charge are evaluated using the new parameters JEOP, NEBP, and LBE, although not indicated in the figure.

5.4.2 Subroutine Additions

5.4.2.1 Subroutine CROWD

The new routine CROWD, which does the dc/transient current crowding analysis, is activated only if $CRBI \leq 0$. As detailed in Chapter 3, the time-dependent variable J_Q , which

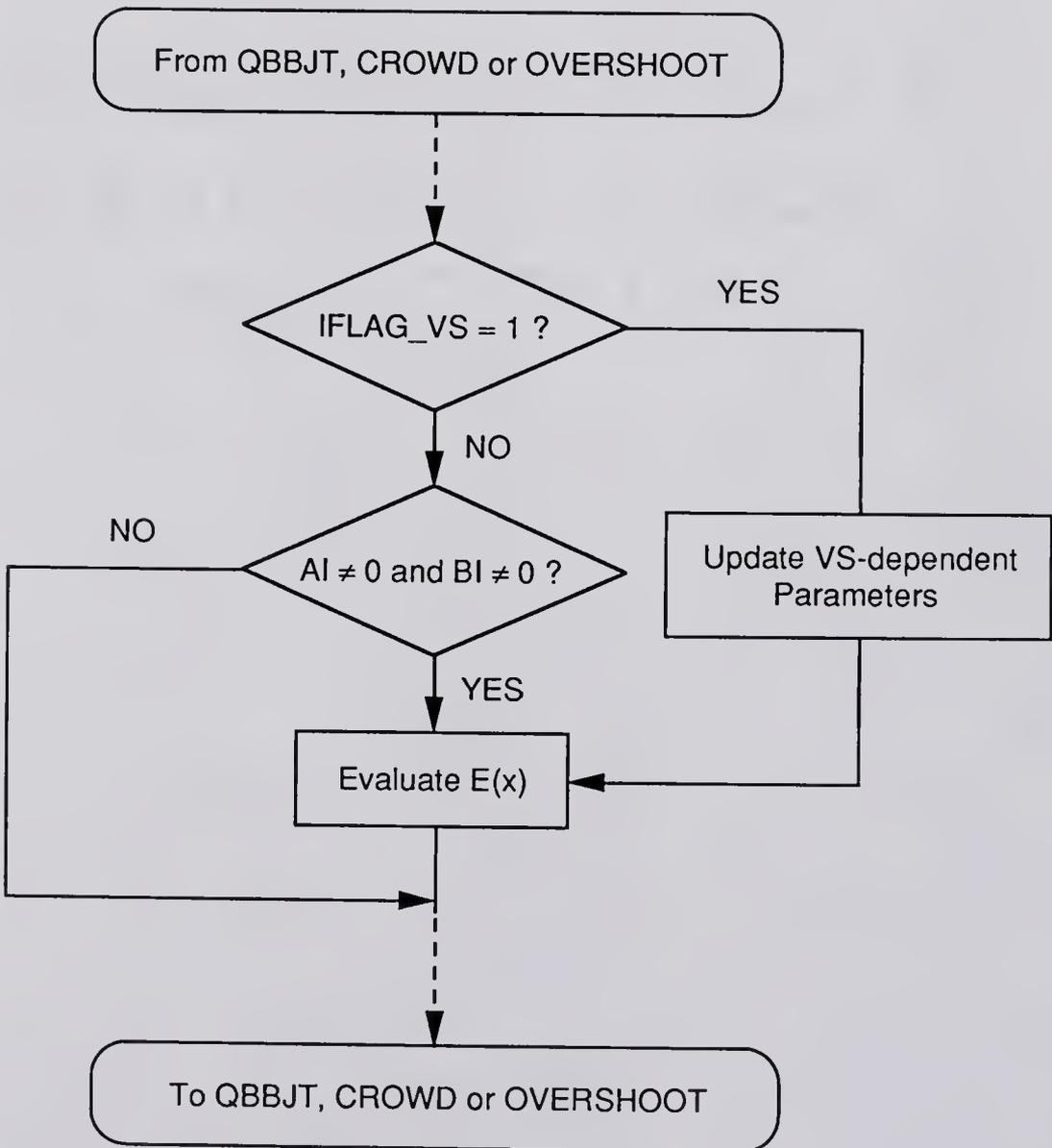


Fig. 5.3 Subroutine QBCT modified for MMSPICE-2.

is the transient counterpart to J_{EO} , is calculated from the previous time-step solution for use in the current time-step analysis. This calculation is done in the first iteration at each time step, as illustrated in Fig. 5.4. Then the value of CRBI is monitored. If CRBI=0, the crowding analysis in Chapter 3 is activated to derive the effective bias on the emitter-base junction. Otherwise (i.e., if CRBI<0), the current crowding is accounted for in a different way, depending on the sign of J_Q ; in switch-on case ($\rightarrow J_Q > 0$), the crowding analysis is done as before, but in switch-off case ($\rightarrow J_Q < 0$), $v_{BE(eff)}$ is assumed to be v_{BE} without any further calculations, based on the recognition that the switch-off transient current crowding is not significant for contemporary devices as discussed in Chapter 3. When only the dc crowding analysis is needed, all these steps are skipped; the crowding model is activated directly with $J_Q=0$.

5.4.2.2 Subroutine OVERSHOOT

The new routine OVERSHOOT does the velocity overshoot analysis, if VS=0 on the model card. For given biases $v_{BE(eff)}$ and v_{BC} , the electric field distribution is first determined (by calling QBCT), using a saturated drift velocity VS. As discussed before, $v_{BE(eff)}$ would be the

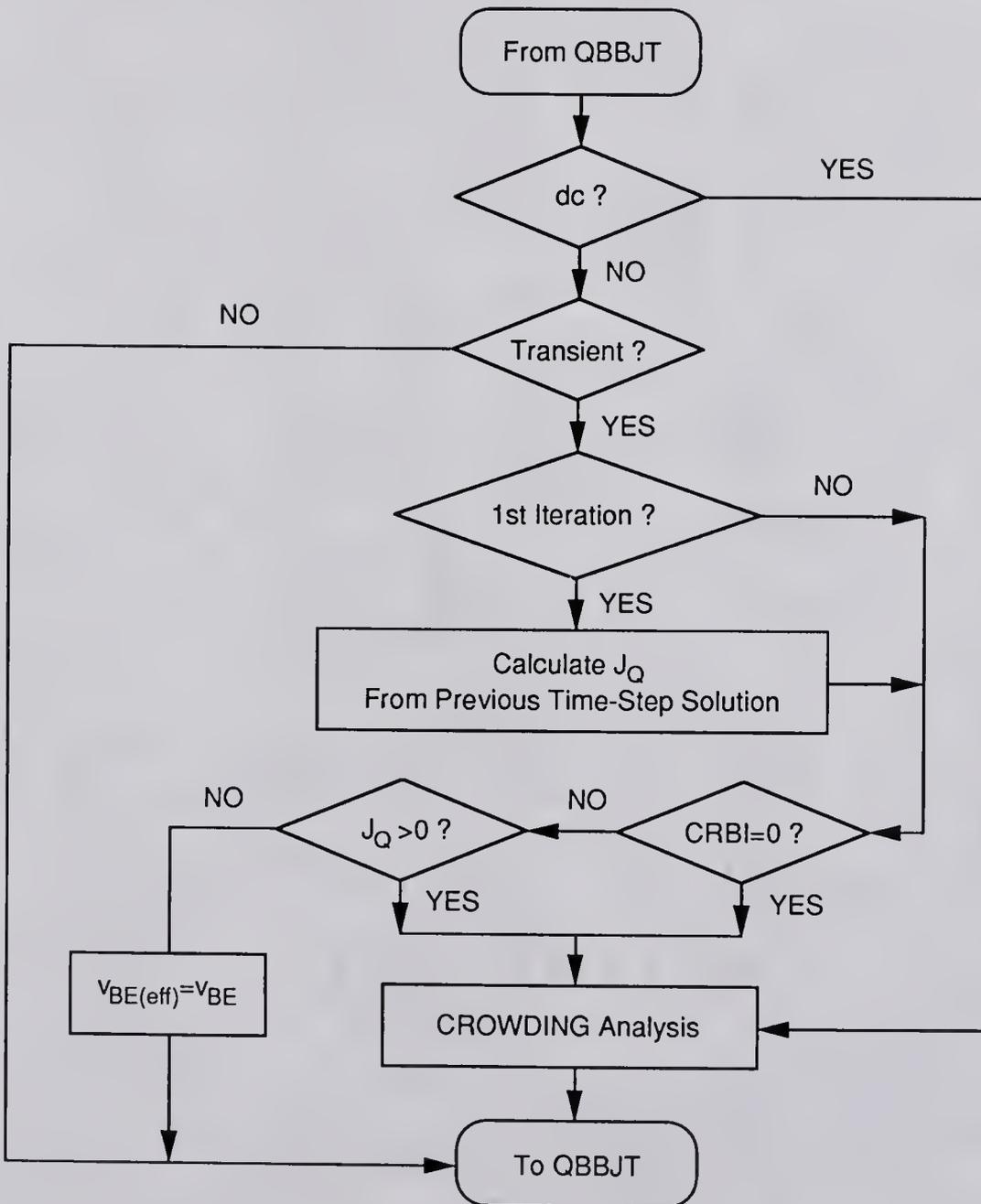


Fig. 5.4 Subroutine CROWD.

actual bias if the crowding option is not on. Then, depending on the sign of the electric field gradient, two different kinds of analyses are activated separately as shown in Fig. 5.5.

Where $d|E|/dx > 0$, which corresponds to the base-side of the junction SCR or the current-induced SCR, the overshoot analysis is done; the length coefficient $L(E)$ is calculated at first, and then the carrier velocity in the SCR is evaluated via the augmented drift-diffusion formalism described in Chapter 4. Where $d|E|/dx < 0$, which corresponds to the collector-side of the junction SCR, the associated velocity relaxation is characterized phenomenologically to be consistent with the overshoot analysis. Once the velocity distribution is known over the entire SCR region, the effective saturated velocity $v_{s(eff)}$ is derived using a numerical integration method. Then, this velocity is returned to QBBJT to be used in the nodal analysis.

5.5 Demonstration

Examples of simulations by MMSPICE-2 were presented together with corroborating measurements in previous chapters describing the individual developments of the new features. In this section, the utility of MMSPICE-2 is demonstrated by

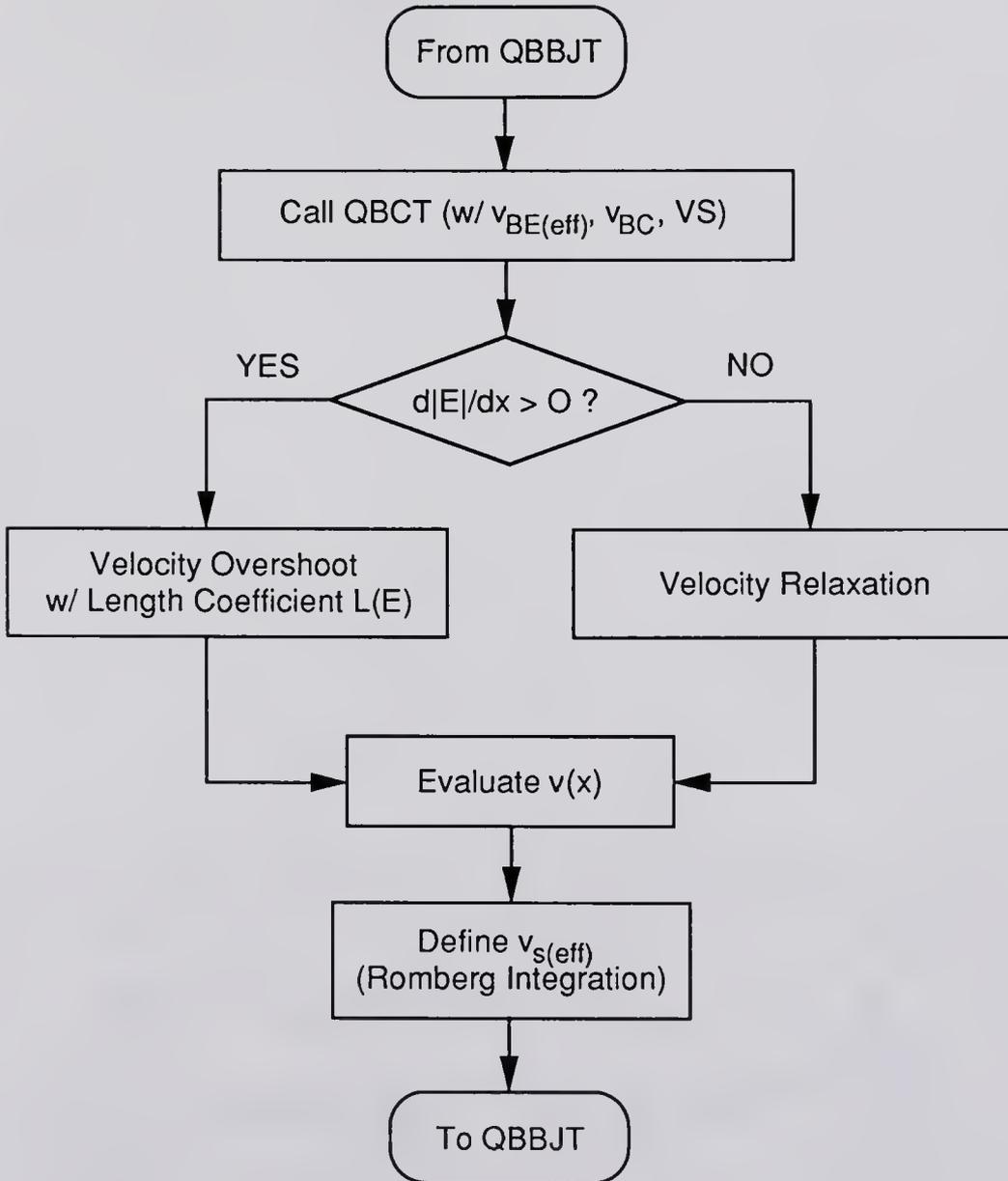


Fig. 5.5 Subroutine OVERSHOOT.

using it in an actual IC TCAD exercise, emphasizing how the user options can be exploited to effect computational efficiency. A comparison of computation times involving various combinations of the new features is given to benchmark the simulator.

The technology chosen for this exercise is ECL. The four-transistor ECL inverter, which was also used in Chapters 3 and 4, is the TCAD vehicle. The circuit diagram is the same as before, except that the four BJTs have $A_E=10\mu\text{m}^2$ ($W_E=1.2\mu\text{m}$) in this example. Fig. 5.6 shows a typical MMSPICE-2 input file for transient simulation of the circuit. Note that CRBI and VS in the model line are set to 0 to activate both the crowding and overshoot analyses, and the extrinsic collector-base and substrate capacitances are also accounted for by specifying LBE and ABL respectively. The peripheral component of base current is neglected since JEOP (and NEBP) are not specified. Note that MMSPICE-2 is structured such that if none of the new options are used, the MMSPICE-1 simulation is done. In fact, MMSPICE-1 input file can be used directly with MMSPICE-2 provided CRBI>0, or is defaulted.

As inferred from the previous chapters, the current crowding and velocity overshoot have opposite effects on the propagation delay of the circuit, as is well illustrated by

AN ECL INVERTER FOR MMSPICE-2 TRANSIENT SIMULATION

```

Z1 2 1 4 9 NPNMOD AE=10P AC=10P WE=1.2U ABL=10P IC=0.7 1.0
Z2 3 5 4 9 NPNMOD AE=10P AC=10P WE=1.2U ABL=10P IC=0.7 1.0
Z3 4 6 7 9 NPNMOD AE=10P AC=10P WE=1.2U ABL=10P IC=0.7 1.0
Z4 0 2 8 9 NPNMOD AE=10P AC=10P WE=1.2U ABL=10P IC=0.7 1.0

RC1 0 2 500
RC2 0 3 500
REE 7 9 1840
RL 8 9 4090
VREF 5 0 -1.11
VCS 6 0 -2.5
VEE 9 0 -5.2

VIN 1 0 PULSE(-1.36 -0.86 100P 50P 50P 200P .5N)
.TRAN 5P 500P
.PRINT TRAN V(1) V(8)
.PLOT TRAN V(8) V(1) (-1.5,-0.5)

.MODEL NPNMOD QBNPN
+ (UNEPI=1069 NEPI=1.998E16 WEPI=0.466E-6 WBM=0.151E-6 ETA=3.66
+ NAO=1.455E18 TC=8.335E-6 TB=4.424E-6 TE=2.590E-10 JEO=0.85E-8
+ JSEO=2.0E-4 NEB=2.0 WSEO=2.32E-8 PE=0.978 ME=0.824 PC=0.791
+ DNB=9.790 CJS=1.45E-4 PS=0.689 MS=0.5 CIF=1.0 CIR=1.0 RC=70 RB=70 RE=9.5
+ FB=0.5 FC=0.5 AI=0.0 BI=0.0 UPBASE=230 CRBI=0.0 VS=0.0 LBE=1.5E-6)

.END

```

Fig. 5.6 MMSPICE-2 input file for transient simulation of an ECL inverter.

the transient simulation results in Fig. 5.7(a). When only the (transient) current crowding is accounted for, the delay is increased by about 45% over the delay when neither effect is accounted for. Conversely, the velocity overshoot reduces the delay by about 5%. Hence they tend to compensate each other in the actual device; the propagation delay is increased by approximately 40% for this circuit when both effects are modeled. This implies that the crowding effect is more predominant than the overshoot effect in contemporary devices. These simulations were done without the extrinsic collector-base capacitance accounted for. When $C_{cb(ext)}$ is included, the overall delay is also increased as shown in Fig. 5.7(b). For the assumed width of the extrinsic base region of $1.5\mu\text{m}$, which is typical in advanced BJTs, the actual delay is, due to $C_{cb(ext)}$ only, lengthened by about 110% over the delay when none of these effects is accounted for. If both $C_{cb(ext)}$ and current crowding are included in the simulations, the overall delay is obviously increased further, but the relative significance of the crowding is lessened by the presence of the extrinsic capacitance as shown in the figure; the transient crowding in this case lengthens the propagation delay by approximately 25%, in contrast to the 45% noted above.

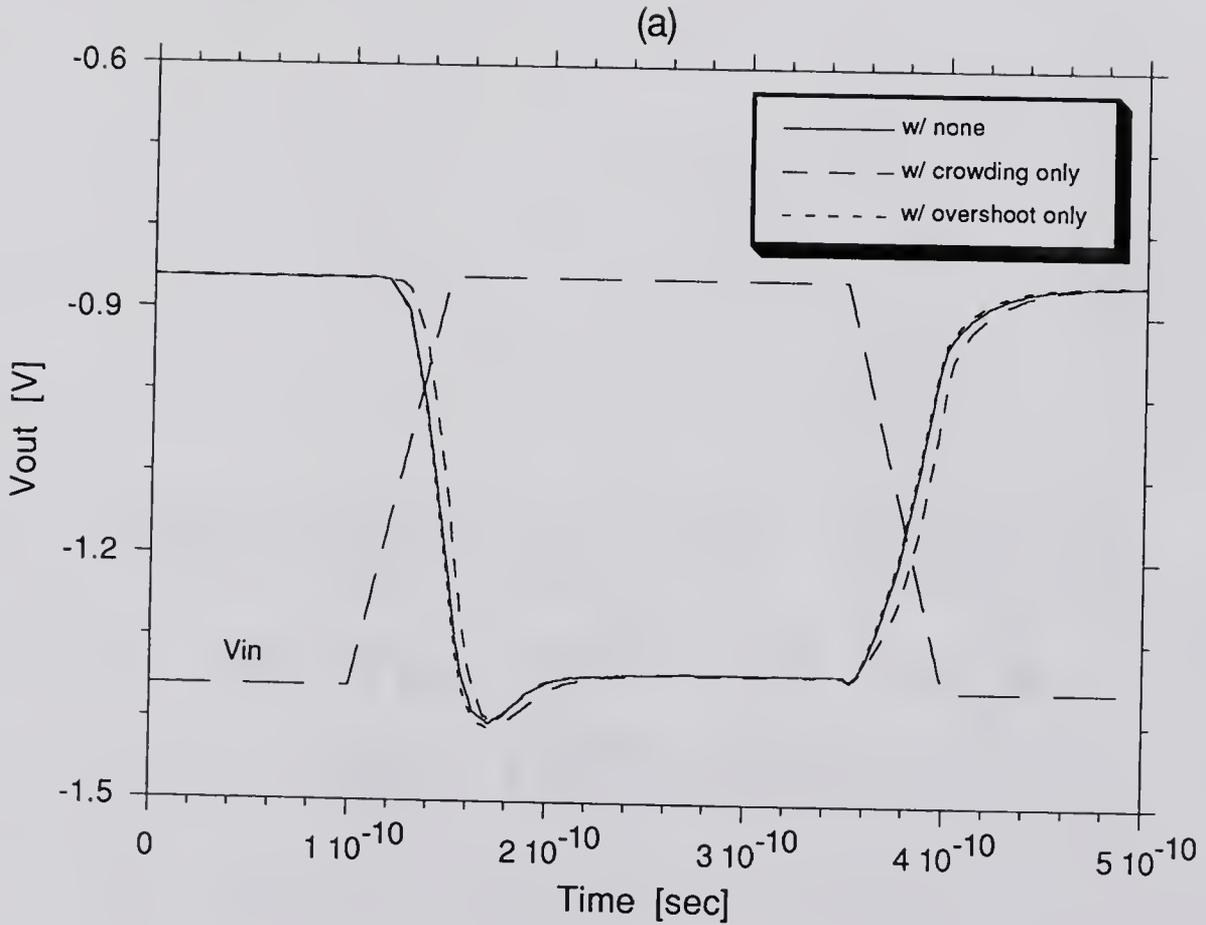


Fig. 5.7(a) Switching waveforms of the ECL inverter circuit simulated with either the transient current crowding or velocity overshoot accounted for. The effect of $C_{cb(ext)}$ is not included.

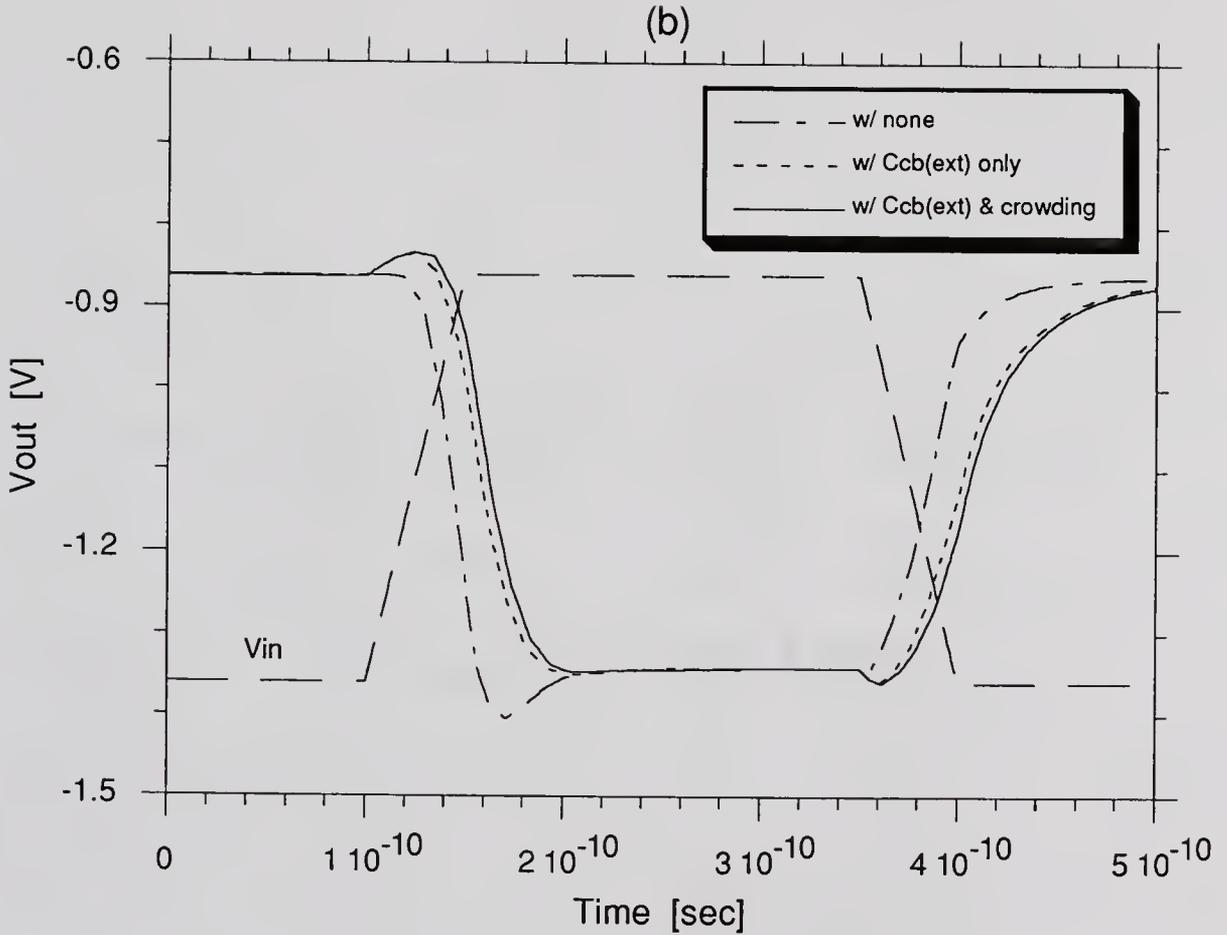
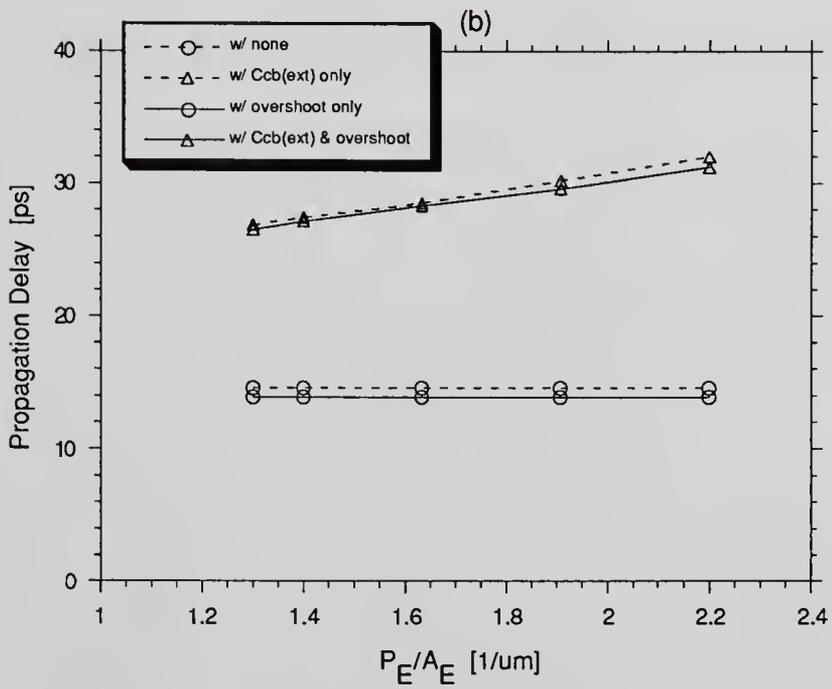
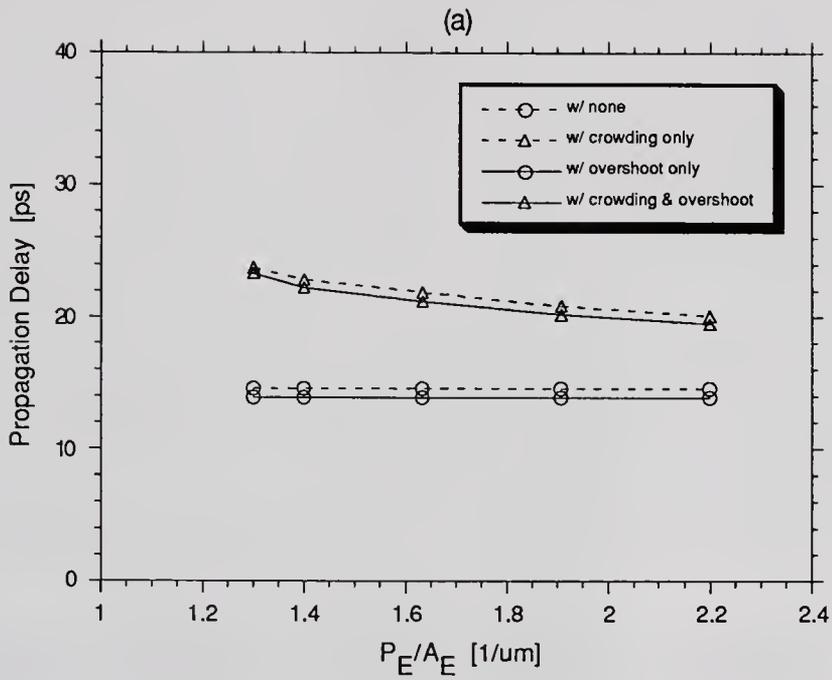


Fig. 5.7(b) Switching waveforms of the ECL inverter circuit, simulated with and without $C_{cb}(\text{ext})$ accounted for. The overall delay is increased with the inclusion of the current crowding in the simulation, but its relative significance is lessened by the presence of $C_{cb}(\text{ext})$.

The previous example reveals that for contemporary bipolar technologies, $C_{cb(ext)}$ is most predominant. For more advanced bipolar integrated circuits however, the aforementioned effects can occur simultaneously, all with relative significance. Then the device/circuit design will have to be done in such a way that all the combined effects are accounted for to yield the optimal performance. An example of such a design trade-off, which can also be applicable to future technologies, is now exemplified.

The goal in this example is to determine the optimum value of P_E/A_E for the BJT (with $A_E=10\mu m^2$) that minimizes the propagation delay of the ECL gate. First, let us consider the current-crowding effect. As characterized in Chapter 2, the current crowding is, as shown in Fig. 5.8(a), diminished as P_E/A_E increases since the voltage drops in the intrinsic base region are reduced with decreasing W_E . The simulations when neither $C_{cb(ext)}$, current crowding, velocity overshoot, nor multi-dimensional currents are accounted for are also shown in the figure for comparison. In this case, the delay is a constant since the amount of charge stored in the BJT is a constant, as implied by the fixed emitter area. If the velocity overshoot is also accounted for in the simulations, the delays are reduced because of less base pushout, and hence all the curves are shifted downward. However the trend



is not changed because the overshoot effect is almost independent of the device geometry.

The nature of $C_{cb(ext)}$ is quite opposite to that of the current crowding, as shown in Fig. 5.8(b). The effect of $C_{cb(ext)}$ becomes significant with increasing P_E/A_E since the area of the extrinsic base region increases. Based on this recognition, it can be inferred that there is an optimum value of P_E/A_E which minimizes the propagation delay by compensating both effects. This value can be determined by accounting for both effects in the simulations. Fig. 5.8(c) shows that MMSPICE-predicted optimum value of P_E/A_E is 1.4, which is equivalent to $L_E/W_E=5\mu\text{m}/2\mu\text{m}$. Note that velocity overshoot still does not change the trend.

Although this kind of design optimization is crude, it reveals not only the versatility of MMSPICE-2, but also its potential utility in TCAD applications. This could not be achieved by any existing simulator because of prohibitive computation time. In fact, MMSPICE-2 is more efficient and thus offers a viable alternative to purely numerical simulation when the simulation time is at a premium.

To stress the computational efficiency, run-time performance of MMSPICE-2 is considered for various combinations of the individual models. Comparisons are made against MMSPICE-1, which is typically only about a factor of

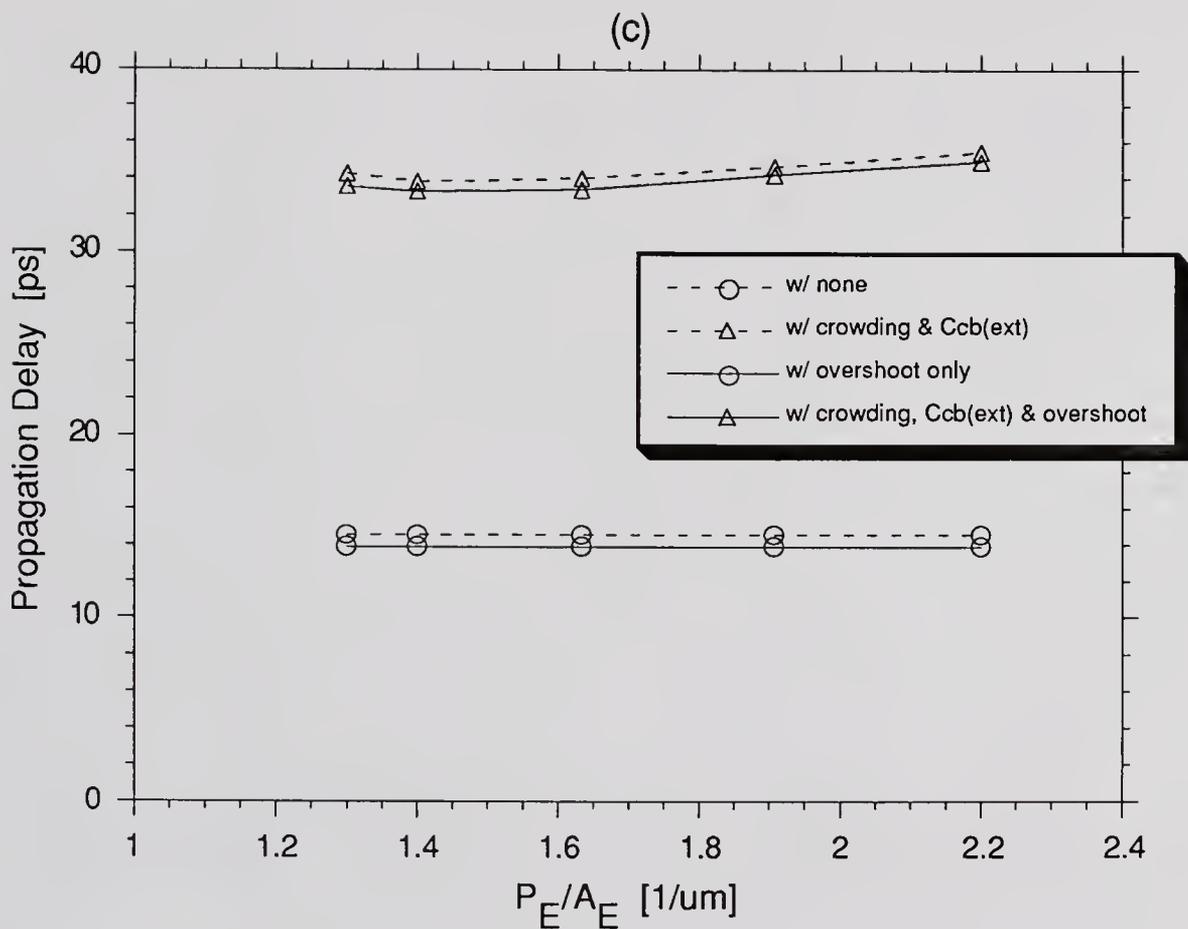


Fig. 5.8 Average propagation delays of the ECL gate versus P_E/A_E (with fixed $A_E=10\mu\text{m}^2$) simulated with only current crowding (a), with only $C_{cb}(\text{ext})$ (b), and with both effects accounted for (c). For each case, results are shown with and without the velocity overshoot included.

two slower than SPICE2 [Jeo90, Hon91]. In the dc case, I-V characteristics like in Fig. 4.11 were simulated for 60 bias points. In the transient case, the previous ECL gate and the single-stage RTL inverter used in Chapter 3 were simulated for 120 and 200 time points respectively. Assumed model parameters are identical to those described in Fig. 5.6. Table 5.3 lists the total number of iterations and execution times for each circuit. Computation time was counted on a SUN4 SPARC station, with $ABSTOL=5 \times 10^{-12}$, $VNTOL=5 \times 10^{-6}$, and $RELTOL=5 \times 10^{-3}$.

With either the current crowding or velocity overshoot accounted for, the execution time is increased by at least 33% theoretically, since an additional QBCT-call is needed for each analysis. (In MMSPICE-1, QBCT is accessed three times.) The run time is also lengthened due to the numerical nature of the both models, and to higher iteration counts required for convergence. When only the current crowding is accounted for, the execution time is increased by about 60% whereas the overshoot analysis increases the run time by about 40% on the average. Table 5.3 implies that MMSPICE-2 tends to be computationally less efficient especially when the transient crowding analysis is being done. This is to be expected since the NQS nature of the transient crowding makes the convergence of the solution at each time-point harder.

TABLE 5.3
 MMSPICE-2 RUN-TIME COMPARISONS
 (all run times in sec)

Models	dc analysis (I-V curve)		transient analysis			
			RTL inverter		ECL gate	
	iter.no	time	iter.no	time	iter.no	time
w/ none	365	3.13	146	1.97	550	4.95
w/ crowding	391	4.25	234/203*	3.30/2.82*	831/787*	8.51/8.30*
w/ overshoot	377	4.14	138	2.75	548	7.64
w/ both	390	5.13	190/191*	3.94/3.84*	806/780*	12.98/12.37*

* Only the switch-on current crowding is accounted for.

When only the switch-on crowding is accounted for (by specifying a negative CRBI on the model card), the number of iterations and thus the computation time of course decrease, although the reduction is modest for the circuits under consideration. However this simplified crowding analysis, the error for which is typically small based on the full-scale crowding analysis, does improve the convergence and is recommended for simulations of large circuits where numerical efficiency is desired.

5.6 Summary

The models developed in previous chapters were successfully implemented into MMSPICE as options, creating MMSPICE-2 so that users may activate any combination of the new models. The algorithm for implementation was discussed in detail. The new features of MMSPICE-2 were demonstrated by the simulations of the ECL gate. Then an example of a design trade-off was exemplified, which revealed that MMSPICE-2 could be a powerful tool for mixed-mode device/circuit simulation for bipolar TCAD applications. The computation time of MMSPICE-2 was also examined for different combinations of the new models, and it appeared that accounting for all the feature models of MMSPICE-2 would increase the execution time by about a factor of two.

CHAPTER 6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK

In this dissertation, modeling of non-quasi-static, non-local, and multi-dimensional current effects in advanced BJTs has been presented. First, a simple analytic way of accounting for the lateral injection of the base current, which was shown to be the most important multi-dimensional effect, was described based on the separation of the current into internal and peripheral components. Second, a new model for transient current crowding was derived. The model, which characterizes a time-dependent effective bias on the emitter-base junction, accounts for base conductivity modulation and the NQS nature of crowding. The modeling/implementation was based on the use of the previous time-step solution in the current time-step analysis. Third, an analytic model for electron velocity overshoot in advanced silicon BJTs was presented. The model, which characterizes an effective saturated drift velocity in the collector SCRs, is based on a non-local augmented drift-velocity formalism that involves a length coefficient derived from Monte Carlo simulations. All the new models, including both the extrinsic collector-base and collector-substrate capacitances, were implemented in MMSPICE so that the user may activate any combination of the

new features by option. The resulting tool could enable truly predictive, scalable, mixed-mode device/circuit simulation for bipolar (and BiCMOS) VLSI technology CAD. The final version of the code, MMSPICE-2.0, evolved from this dissertation, is available at the University of Florida. The following tasks are suggested as future work to make the simulator more useful.

- (1) The transient current-crowding model needs to be verified through small-signal s-parameter and large-signal transient measurements.
- (2) Numerical aspects associated with the no-solution problem in switch-off crowding analysis need to be investigated further for potential improvements in convergence and numerical efficiency.
- (3) A study on the evaluation of an average hole mobility needed in the current-crowding analysis is recommended.
- (4) A study on the NQS modeling of ac crowding will be a worthwhile task.
- (5) The evaluation and modeling of the length coefficient, which is crucial in the characterization of velocity overshoot analysis, have to be refined.
- (6) An assessment of the significance (benefit) of velocity overshoot in future scaled bipolar devices and circuits could be done with MMSPICE-2.

- (7) For the accurate analysis of the electron transport, the transient field (dE/dt) dependence of the electron drift velocity, which might be significant during fast transients, must be accounted for in overshoot analysis.
- (8) SUMM [Gre90] should be expanded for MMSPICE-2.
- (9) More comprehensive verification of the models and the tool, based on purely numerical device simulations and on experimental measurements, should be done to check for robustness and stability of MMSPICE-2.

APPENDIX A
EVALUATION OF J_{SEO} , n_{EB} , J_{EOP} AND n_{EBP}

The evaluation of the model parameters associated with the peripheral base current is quite straightforward. In the low current region, equation (2.3) can be approximated as a straight line:

$$\frac{I_B}{A_E} = \frac{I_{BA} + I_{BP}}{A_E} \approx \alpha \left(1 + \beta \frac{P_E}{A_E} \right) \quad (A.1)$$

where

$$\alpha \equiv J_{SEO} \exp\left(\frac{V_{BE}}{n_{EB}V_T}\right) \quad , \quad (A.2)$$

$$\beta \equiv \frac{J_{EOP}}{J_{SEO}} \exp\left[\left(\frac{1}{n_{EBP}} - \frac{1}{n_{EB}}\right) \frac{V_{BE}}{V_T}\right] \quad . \quad (A.3)$$

When V_{BE} is fixed, α is found by extrapolation, and β is derived from the slope of the straight line.

On the other hand, α and β can be rearranged as

$$\ln(\alpha) = \ln(J_{SEO}) + \frac{V_{BE}}{n_{EB}V_T} \quad , \quad (A.4)$$

$$\ln(\beta) = \ln\left(\frac{J_{EOP}}{J_{SEO}}\right) + \left(\frac{1}{n_{EBP}} - \frac{1}{n_{EB}}\right) \frac{V_{BE}}{V_T} \quad . \quad (A.5)$$

Similarly, J_{SEO} , n_{EB} , J_{EOP} and n_{EBP} can be evaluated by determining α and β at different values of V_{BE} , and by plotting $\ln(\alpha)$ and $\ln(\beta)$ versus V_{BE} . For the advanced BJTs fabricated at Texas Instruments, the extracted parameter values are $J_{SEO}=1.03 \times 10^{-5} \text{A/m}^2$, $n_{EB}=1.79$, $J_{EOP}=1.21 \times 10^{-11} \text{A/m}$, and $n_{EBP}=1.53$, which are similar to those previously reported [Cha91].

APPENDIX B
DISCUSSION ON J_Q

The time-dependent variable J_Q , which is a transient counterpart to J_{EO} , can be defined by two different ways:

$$J_Q^{m+1} \approx J_Q^m \equiv \frac{\frac{dQ_{BE}^m(0)}{dt}}{L_E W_{E(\text{eff})}^m \left[\exp\left(\frac{V_{BE}}{V_T}\right) - 1 \right]} \quad (\text{B.1})$$

or

$$\equiv \frac{\frac{dQ_{BE}^m(0)}{dt}}{L_E W_E \left[\exp\left(\frac{V_{BE(\text{eff})}^m}{V_T}\right) - 1 \right]} \quad (\text{B.2})$$

However the investigation reveals that using effective emitter width $W_{E(\text{eff})}$ is not physical; in the switch-on case where $W_{E(\text{eff})}$ would be less than W_E , the smaller $W_{E(\text{eff})}$ can not properly account for the deactivated emitter region because, in reality, the current is still flowing along the deactivated region. Also, $W_{E(\text{eff})}$ that results when $J_Q \gg J_{EO}$ often reaches values that are less than actual W_E by about two orders of magnitude during the switch-on transient. This may be unrealistic, and often causes convergence problems in

the circuit nodal analysis.

On the contrary, using $v_{BE(\text{eff})}$ is not only physical but also stable in the numerical sense, since it is the argument of the exponential function in (B.2). Consequently, $v_{BE(\text{eff})}$ will be used as a time-dependent variable to account for the transient current crowding non-quasi-statically.

APPENDIX C
LIMITING $J_{EO(eff)}$ IN THE SWITCH-OFF SIMULATION

The no-solution problem in the switch-off analysis, which arises from extremely large negative $J_{EO(eff)}$, can be avoided by limiting $J_{EO(eff)}$. To do this, we first combine (3.19) and (3.21) to express the magnitude of $J_{EO(eff)}$ as a function of A:

$$f(A) = |J_{EO(eff)}| = \frac{KA^2}{W_E L_E \exp\left(\frac{V_{BE}}{V_T}\right) \cosh^2(KA)} \quad (C.1)$$

where $K \equiv \rho W_E / (4V_T)$. We want to determine the largest value of $|J_{EO(eff)}|$ for which (3.7), (3.19), and (3.21) have a real solution. This is simply the maximum of the function $f(A)$, which we assume occurs at $A=A_{crit}$ where

$$\frac{df(A_{crit})}{dA} = 0 \quad (C.2)$$

Using (C.1) in (C.2) yields the condition

$$\tanh(KA_{crit}) - \frac{1}{KA_{crit}} = 0 \quad (C.3)$$

The solution A_{crit} of this equation is obtained numerically by the Newton-Raphson method, and then substituted into (C.1) to give the desired limit of $J_{\text{EO}(\text{eff})}$.

APPENDIX D
VALIDITY OF THE DEPLETION APPROXIMATION

In order to solve Poisson's equation in SCRs more accurately, the free carrier concentration should possibly be included. Then (4.14) should be written as

$$\begin{aligned} \frac{dE}{dx} &= - \frac{q}{\epsilon} [N_A(x) + n] \\ &= - \frac{q}{\epsilon} N_{A0} \exp\left[\frac{-\eta}{W_{BM}}(x + W_{BM})\right] - \frac{|I_C|}{\epsilon A v_s} \quad . \end{aligned} \quad (D.1)$$

Here we assume that electrons would travel at a constant velocity v_s in the SCR.

Fig. D.1(a) illustrates the electric fields and the velocity distributions (based on (4.8)) in the base-side of the junction SCR with and without the electron concentration accounted for, when $V_{BE}=0.7V$ and $V_{BC}=0.0V$. As shown in the figure, the error in the evaluation of the electric field is very small; the maximum error is only 0.24%. Fig. D.1(b) illustrates the case when a reverse bias of $V_{BC}=-2.0V$ is applied; the maximum error is 0.33%. Therefore, the use of the depletion approximation is quite valid.

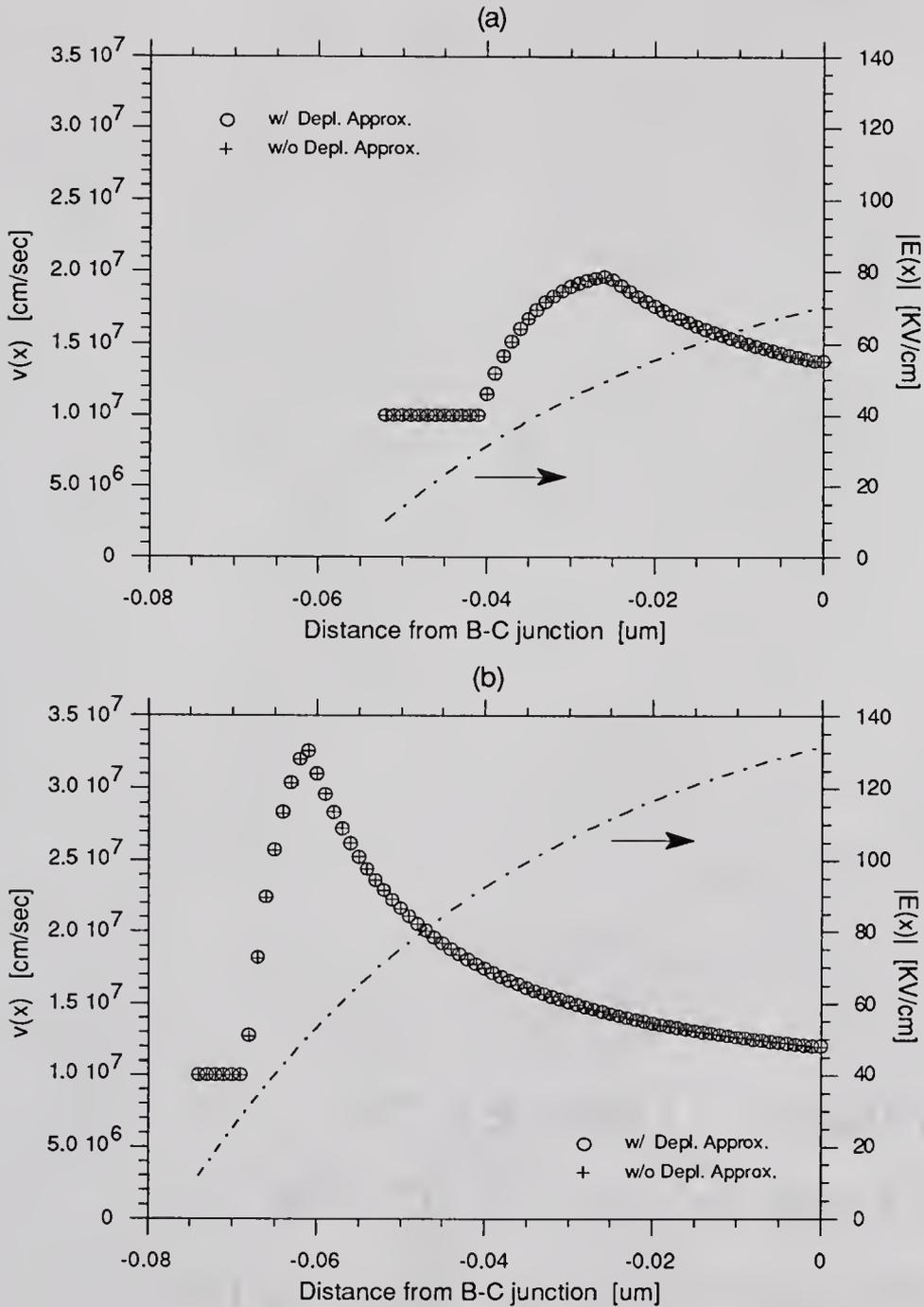


Fig. D.1 Drift velocity and electric field in the base-side of the junction SCR when $V_{BE} = 0.7V$: (a) $V_{BC} = 0.0V$; (b) $V_{BC} = -2.0V$.

REFERENCES

- Art88 M. Artaki, "Hot-Electron Flow in an Inhomogeneous Field," Appl. Phys. Lett., vol. 52, pp. 141-143, Jan. 1988.
- Bla90 P. A. Blakey, X. -L. Wang, and C. M. Maziar, "A Generalized Formulation of Augmented Drift-Diffusion Transport Suitable for Use in General Purpose Device Simulators," Tech. Digest IEEE NUPAD-III Workshop, pp. 37-38, May 1990.
- Blo70 K. Blotekjaer, "Transport Equations for Electrons in Two-Valley Semiconductor," IEEE Trans. Electron Devices, vol. ED-17, pp. 38-47, 1970.
- Bor91 T. Bordelon, X. -L. Wang, C. M. Maziar, and A. F. Tasch, "An Evaluation of Energy Transport Models for Silicon Device Simulation," Solid-State Electron., vol. 34, pp. 617-628, June 1991.
- Cha91 A. Chantre, G. Festes, G. G.-Matlakowski, and A. Nouailhat, "An Investigation of Nonideal Base Currents in Advanced Self-Aligned Etched-Polysilicon Emitter Bipolar Transistors," IEEE Trans. Electron Devices, vol. 38, pp. 1354-1361. June 1991.
- Che91 D. Chen, E. C. Kan, and U. Ravaioli, "An Analytical Formulation of the Length Coefficient for the Augmented Drift-Diffusion Model Including Velocity Overshoot," IEEE Trans. Electron Devices, vol. 38, pp. 1484-1490, June 1991.
- Chu87 C. T. Chuang, D. D. -L. Tang, G. P. Li, and E. Hackbarth, "On the Punchthrough Characteristics of Advanced Self-Aligned Bipolar Transistors," IEEE Trans. Electron Devices, vol. ED-34, pp. 1519-1524, July 1987.
- Cra90 E. F. Crabbe, J. M. C. Stork, G. Baccarani, M. V. Fischetti, and S. E. Laux, "The Impact of Non-Equilibrium Transport on Breakdown and Transit Time in Bipolar Transistors," Tech. Digest International Electron Device Meet., pp. 463-466, 1990.

- Das91 A. Das and M. Lundstrom, "Does Velocity Overshoot Reduce Collector Delay Time in AlGaAs/GaAs HBT's?" IEEE Electron Device Lett., vol. 12, pp. 335-337, June 1991.
- Dej88 J. L. de Jong, R. H. Lane, J. G. de Groot, and G. W. Conner, "Electron Recombination at the Silicided Base Contact of an Advanced Self-Aligned Poly-Silicon Emitter," Proc. IEEE Bipolar Circuits and Technology Meet., pp. 202-205, 1988.
- Fus92 T. Fuse, T. Hamasaki, K. Matsuzawa, and S. Watanabe, "A Physically Based Base Pushout Model for Submicrometer BJTs in the Presence of Velocity Overshoot," IEEE Trans. Electron Devices, vol. 39, pp. 396-403, Feb. 1992.
- Gol88 N. Goldsman and J. Frey, "Efficient and Accurate Use of the Energy Transport Method in Device Simulation," IEEE Trans. Electron Devices, vol. 35, pp.1524-1529, Sept. 1988.
- Gre90 K. R. Green and J. G. Fossum, "SUMM: A SUPREM-3/MMSPICE-1 Integrator for Bipolar Technology CAD," M. S. Thesis, Dept. of Electrical Eng., Univ. of Florida, Gainesville, May 1990.
- Ham88 T. Hamasaki, T. Wada, N. Shigyo, and M. Yoshimi, "Lateral Scaling Effects on High-Current Transients in Submicrometer Bipolar Transistors," IEEE Trans. Electron Devices, vol. 35, pp. 1620-1626, Oct. 1988.
- Hau64 J. R. Hauser, "The Effects of Distributed Base Potential on Emitter-Current Injection Density and Effective Base Resistance for Stripe Transistor Geometries," IEEE Trans. Electron Devices, vol. ED-11, pp. 238-242, May 1964.
- Hon91 G. -B. Hong and J. G. Fossum, "Enhancement of MMSPICE: Version 1.3," VLSI TCAD Group, Dept. of Electrical Eng., Univ. of Florida, Gainesville, July 1991.
- Hur87 G. A. M. Hurkx, "On the Sidewall Effects in Submicrometer Bipolar Transistors," IEEE Trans. Electron Devices, vol. ED-34, pp. 1939-1946, Sept. 1987.

- Hwa87 B. Y. Hwang, P. J. Zdebel, R. J. Balda, G. G. Sweeney, and V. de la Torre, "Lateral Scaling Effects of Double-Polysilicon Advanced Self-Aligned Transistors," Proc. IEEE Bipolar Circuits and Technology Meet., pp. 176-179, 1987.
- Jaf92 J. Jaffee, "Improved RF Modeling of Bipolar Junction Transistors," M. S. Thesis, Dept. of Electrical Eng., Univ. of Florida, Gainesville, May 1992.
- Jeo89 H. Jeong and J. G. Fossum, "A Charge-Based Large-Signal Bipolar Transistor Model for Device and Circuit Simulation," IEEE Trans. Electron Devices, vol. 36, pp. 124-131, Jan. 1989.
- Jeo90 H. Jeong, J. G. Fossum, and D. K. FitzPatrick, "MMSPICE: A Semi-Numerical Mixed-Mode Device/Circuit Simulation for Advanced Bipolar Technology CAD," Solid-State Electron., vol. 33, pp. 1283-1291, Oct. 1990.
- Jin92a J. Jin and J. G. Fossum, "Non-Quasi-Static Modeling/Implementation of BJT Current Crowding for Seminumerical Mixed-Mode Device/Circuit Simulation," IEEE Trans. Computer-Aided Design, vol. 11, pp. 759-767, June 1992.
- Jin92b J. Jin, J. G. Fossum, and D. K. FitzPatrick, "Programmer's Reference Manual of MMSPICE-2," VLSI TCAD Group, Dept. of Electrical Eng., Univ. of Florida, Gainesville, May 1992.
- Jo90 M. Jo and D. E. Burk, "An Intrinsic Base Resistance Model for Low and High Currents," IEEE Trans. Electron Devices, vol. 37, pp. 202-209, Jan. 1990.
- Kan91 E. C. Kan, U. Ravaioli, and T. Kerkhoven, "Calculation of Velocity Overshoot in Submicron Devices Using an Augmented Drift-Diffusion Model," Solid-State Electron., vol. 34, pp. 995-999, Sept. 1991.
- Kiz89 I. C. Kizilyalli and M. Artaki, "Simulation of a GaAs MESFET Including Velocity Overshoot: An Extended Drift-Diffusion Formalism," IEEE Electron Device Lett., vol. 10, pp. 405-408, Sept. 1989.
- Kne90 R. W. Knepper, "Modeling Advanced Bipolar Devices for High Performance Applications," Tech. Digest International Electron Device Meet., pp. 177-180, 1990.

- Lee89 W. Lee, S. E. Laux, M. V. Fischetti, and D. D. Tang, "Monte Carlo Simulation of Non-Equilibrium Transport in Ultra-Thin Base Si Bipolar Transistors," Tech. Digest International Electron Device Meet., pp. 473-476, 1989.
- Li88 G. P. Li, C. T. Chuang, T. -C. Chen, and T. H. Ning, "On the Narrow-Emitter Effect of Advanced Shallow-Profile Bipolar Transistors," IEEE Trans. Electron Devices, vol. 35, pp. 1942-1950, Nov. 1988.
- Lun90 M. Lundstrom and S. Datta, "Physical Device Simulation in a Shrinking World," Circuit and Devices, pp. 32-37, July 1990.
- Mul89 R. S. Muller and T. I. Kamins, Device Electronics for Integrated Circuits. New York: Wiley, 1989.
- PIS84 "PISCES-II: Poisson and Continuity Equation Solver," Dept. of Electrical Eng., Stanford University Technical Report, Palo Alto, CA, Sept. 1984.
- Pri88 P. Price, "On the Flow Equation in Device Simulation," J. Appl. Phys., vol. 63, pp. 4718-4722, May 1988.
- Rei84 H. -M. Rein, "A Simple Method for Separation of the Internal and External (Peripheral) Currents of Bipolar Transistors," Solid-State Electron., vol. 27, pp. 625-631, 1984.
- Rey69 G. Rey, "Effets de la Defocalisation (c.c. et c.a.) sur le Comportement des Transistors a Jonctions," Solid-State Electron., vol. 12, pp. 645-659, 1969.
- Ruc72 J. G. Ruch, "Electron Dynamics in Short Channel Field-Effect Transistors," IEEE Trans. Electron Devices, vol. ED-19, pp. 652-654, May 1972.
- Saw88 S. Sawada, "Perimeter Effect in Advanced Self-Aligned Bipolar Transistor," Proc. IEEE Bipolar Circuits and Technology Meet., pp. 206-209, 1988.
- Shu81 M. S. Shur and L. F. Eastman, "Near Ballistic Electron Transport in GaAs Devices at 77°K," Solid-State Electron., vol. 24, pp. 11-18, Jan. 1981.

- Sto83 J. M. C. Stork, R. D. Isaac, "Tunneling in Base-Emitter Junction," IEEE Trans. Electron Devices, vol. ED-30, pp. 1527-1534, Nov. 1983.
- SUP88 "SUPREM-3: One-Dimensional Process Analysis Program," Technology Modeling Associates, Inc., Palo Alto, CA, Dec. 1988.
- Tan85 D. D. Tang, "Switch-On Transient of Shallow-Profile Bipolar Transistors," IEEE Trans. Electron Devices, vol. ED-32, pp. 2224-2226, Nov. 1985.
- Tho82 K. K. Thornber, "Current Equations for Velocity Overshoot," IEEE Electron Device Lett., vol. EDL-3, pp. 69-71, March 1982.
- Ver87 D. P. Verret, "Two-Dimensional Effects in the Bipolar Polysilicon Self-Aligned Transistor," IEEE Trans. Electron Devices, vol. ED-34, pp. 2297-2303, Nov. 1987.

BIOGRAPHICAL SKETCH

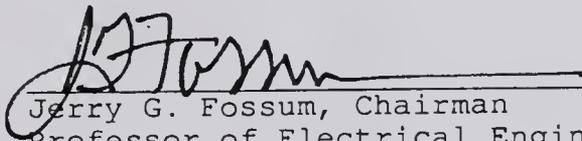
Joohyun Jin was born in Seoul, Korea, in 1958. He received the B. S. degree in electronic engineering from the Seoul National University in 1981 and the M. S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1984. Since 1987, he has been working toward the Ph.D. degree in electrical engineering at the University of Florida, Gainesville.

From 1984 to 1987, he was with Samsung Semiconductor and Telecommunication Co. Ltd., where he was involved in the design of high-speed CMOS devices. He also worked in the characterization of CMOS processes.

His current research interests are in the area of bipolar device modeling for TCAD.

He is a member of IEEE.

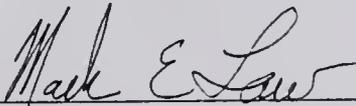
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Jerry G. Fossum, Chairman
Professor of Electrical Engineering

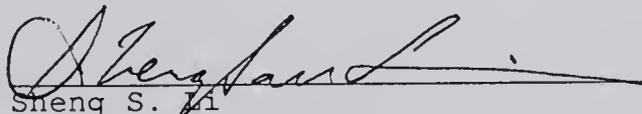
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Dorothea E. Burk
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Mark E. Law
Assistant Professor of Electrical
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

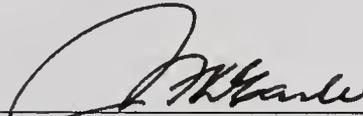

Sheng S. Li
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Timothy J. Anderson
Professor of Chemical Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1992


for _____

Winfred M. Phillips
Dean, College of Engineering

Madelyn M. Lockhart
Dean, Graduate School

