PREDICTIVE MODELING OF DEVICE AND CIRCUIT RELIABILITY IN HIGHLY SCALED CMOS AND SIGE BICMOS TECHNOLOGY

A Dissertation
Presented to
The Academic Faculty

by

Kurt A. Moen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2012

PREDICTIVE MODELING OF DEVICE AND CIRCUIT RELIABILITY IN HIGHLY SCALED CMOS AND SIGE BICMOS TECHNOLOGY

Approved by:

Professor John D. Cressler, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Faramarz Fekri
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Shyh-Chiang Shen
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Samuel Graham
School of Mechanical Engineering
Georgia Institute of Technology

Professor David S. Citrin
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 10, 2012

To my dear wife,

Jennifer

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to Dr. John D. Cressler for his guidance and support as my advisor. Both in the lab and in the classroom, his enthusiasm for research and teaching, combined with his sincere care for others, has shown me how to excel as an engineer while being a blessing to those around you.

I would like to thank the other members of my dissertation committee. I would also like to sincerely thank all of the members of the SiGe research team at Georgia Tech for their support and collaboration: Stan Phillips, Partha Chakraborty, Jiahui Yuan, Ted Wilcox, Tushar Thrivikraman, Tom Cheng, Laleh Najafizadeh, Steven Horst, Ryan Diestelhorst, Marco Bellini, Aravind Appaswamy, Uppili Ragunathan, Duane Howard, Sachin Seth, Akil Sutton, Prabir Saha, Anuj Madan, and others. I am grateful to Dr. Guofu Niu from Auburn University for helpful discussions regarding my research.

I am very grateful for the support provided by the Graduate Fellowship Program of the Semiconductor Research Corporation, as well as that given by DTRA, NASA ETDP, and the SiGe BiCMOS team at IBM Microelectronics. I am particularly thankful for the close collaborations I have enjoyed with Marek Turowski and Ashok Raman of CFD Research Corporation in Huntsville, AL, and with Hiroshi Yasuda of Texas Instruments in Dallas, TX.

I am deeply indebted to my parents, Dale and Marghy Moen, for their love and care throughout my life. I am very grateful to my brother and his wife, Brent and Amy, for the sincere care and support that they have shown me. And most of all, I am beyond grateful to my wife Jennifer, who has made me a better person in so many ways, and who I have grown to love so dearly. Her constant care has sustained me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The advent of high-frequency silicon-based technologies has enabled the design of mixed-signal circuits that incorporate analog, RF, and digital circuit components to build cost-effective system-on-a-chip solutions. Emerging applications provide great incentive for continued scaling of transistor performance, requiring careful attention to mismatch, noise, and reliability concerns. If these mixed-signal technologies are to be employed within space-based electronic systems, they must also demonstrate reliability in radiation-rich environments. SiGe BiCMOS technology in particular is positioned as an excellent candidate to satisfy all of these requirements. The objective of this research is to develop predictive modeling tools that can be used to design new mixed-signal technologies and assess their reliability on Earth and in extreme environments. Ultimately, the goal is to illuminate the interaction of device- and circuit-level reliability mechanisms and establish best practices for modeling these effects in modern circuits. To support this objective, several specific areas have been targeted first, including a TCAD-based approach to identify performance-limiting regions in SiGe HBTs and support device optimization, measurement and modeling of carrier transport parameters that are essential for predictive TCAD simulations, and measurement of device-level single-event transients to better understand the physical origins and implications for device design. These tasks provide the foundation for the bulk of this research, which addresses circuit-level reliability challenges through the application of novel mixed-mode TCAD techniques. All of the individual tasks are tied together by a guiding theme: to develop a holistic understanding of the challenges faced by emerging broadband technologies by coordinating results from material, device, and circuit studies.

CHAPTER I

INTRODUCTION

In the past decade, there has been extraordinary growth in the global telecommunications market, driven largely by emerging broadband-communications applications, such as mobile communications (GSM/CDMA), WLAN, GPS, DSL, and satellite communications. At the core of this growth is the development of monolithic integrated circuit (IC) technologies that have made it possible to build complex integrated systems at reasonable cost. In part, this growth been enabled by the relentless scaling of core device-performance metrics; more importantly, however, technologies have been developed to support increasingly high levels of integration, which allow a wide variety of functionality to be defined together on the same chip, simplifying packaging and reducing total die count.

One increasingly important segment of the broadband-communications market is high-speed communications with satellites in orbit around the earth. Extra-terrestrial electronics systems are required to operate in extremely harsh environments and are subjected to both particle radiation and cryogenic temperatures. Cryogenic temperatures induce significant changes in all aspects of device operation. The fundamental physical properties that drive these changes must be accurately characterized to provide a solid foundation for the development of robust devices and circuits. Furthermore, radiation effects introduce serious reliability concerns that must be addressed before a particular technology can become viable for extreme-environment applications.

Among the building blocks for broadband technologies are low-noise amplifiers, power amplifiers, and voltage-controlled oscillators. The key performance metrics for these blocks include low power consumption, high gain, high-frequency operation, high

dynamic range, good linearity, and low noise. However, these circuit-level metrics are necessarily coupled to corresponding device-level metrics; thus, an understanding of device-level performance is critical to achieving circuit and system-level performance gains. Any useful analysis of device performance depends on a solid understanding of the fundamental physical processes that operate within the device.

1.1 Broadband IC Technologies

The advent of high-frequency silicon-based technologies has enabled the design of mixed-signal ICs that incorporate analog and radio-frequency (RF) circuit components, including the requisite passive elements and interconnects, with highly-integrated digital circuit components. Leveraging the economy of scale provided by silicon IC manufacturing, these technologies allow designers to build cost-effective system-on-a-chip (SoC) or system-in-a-package (SiP) solutions for a variety of communications applications [26]. At present, there exist two families of high-frequency silicon-based mixed-signal technologies: highly-scaled RF complementary-metal-oxide-semiconductor (RF-CMOS) and silicon-germanium bipolar-CMOS (SiGe BiCMOS), with each possessing its own unique advantages.

Aggressive lithographic scaling, new materials, and process innovations such as strain engineering have enabled the integration of RF-optimized metal-oxide-semiconductor field-effect-transistors (MOSFETs) into traditional digital CMOS technology. Strain-engineered RF-CMOS on silicon-on-insulator (SOI) represents the leading edge of CMOS technology, possessing advantages over bulk RF-CMOS by minimizing parasitics, improving isolation, decreasing leakage, improving short-channel effects, and improving single-event upset (SEU) tolerance [58]. The combination of enhanced RF performance with state-of-the-art digital CMOS makes RF-CMOS on SOI an attractive technology for system-on-a-chip applications such as integrating RF front ends and baseband analog/digital circuitry on a single chip.

Similarly, SiGe BiCMOS technology is well suited for a wide variety of analog, RF, and high-speed digital circuits, because of its high-frequency operation, low broadband and 1/f noise, high transconductance per unit area, and compatibility with conventional CMOS fabrication. Modern SiGe technology is almost universally implemented with the high-frequency SiGe heterojunction bipolar transistor (HBT) as an add-on to a core digital CMOS technology (SiGe HBT + Si CMOS). Consequently, within a mixed-signal IC, this allows the SiGe HBT to be used where it is best suited, that is, within RF, microwave, analog, and high-speed digital circuit components, whereas Si CMOS can be used to its greatest advantage within lower-performance memory and digital circuit components [23]. Although InP HBTs have demonstrated greater current-gain and power-gain cutoff frequencies (f_T and f_{max}) at higher breakdown voltage (BV_{CEO}) than their SiGe HBT counterparts, scaling trends suggest that THz levels of performance could be achieved by SiGe HBTs at useful levels of breakdown, combining the enormous integration and cost advantages of silicon manufacturing with device performance comparable to III-V technologies [115].

Comparing SiGe BiCMOS to RF-CMOS, one key difference is that the frequency performance of the SiGe HBT is primarily determined by its vertical profile, whereas RF-optimized MOSFETs depend on the minimum feature size allowed at a particular lithography node. Consequently, SiGe HBTs enjoy roughly a two-generation lithographic scaling advantage over CMOS for fixed performance. Since lithography has increasingly become the largest fixed cost for IC manufacturing, this gives the SiGe HBT a significant cost advantage over RF-CMOS, outweighing the cost of the additional masks needed to define the SiGe HBT in the BiCMOS process [24]. Furthermore, CMOS transistors face increasingly difficult challenges with device-to-device matching as they are scaled because of larger relative variations in the lateral dimensions, further compounded by increased short-channel effects and the high-k dielectrics and metal gate used in the most advanced technology nodes. In contrast, matching in bipolar

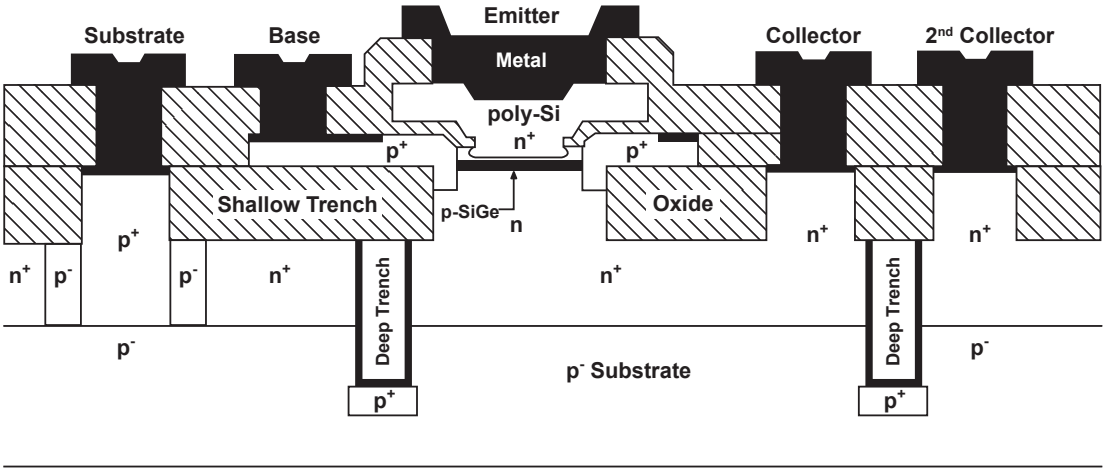

Figure 1: Cross-section of a representative first-generation SiGe HBT.

devices tends to improve with scaling as a result of the increase in doping levels. Highly-scaled RF-CMOS also presents serious challenges for circuit designers because of poor output conductance, high leakage currents, degraded low-frequency noise, and low breakdown voltages. For the SiGe HBT, breakdown voltage is becoming one of the major scaling bottlenecks, although breakdown voltages remain higher than CMOS for fixed frequency performance [115].

1.2 The SiGe HBT

Fundamentally, the SiGe HBT is very much the same as its Si bipolar-junction transistor (BJT) counterpart, except that in the HBT a graded Ge profile is introduced in the base layer, which allows device designers to exercise bandgap-engineering for the first time in silicon-based technology. The cross-section of a representative first-generation SiGe HBT is shown in Figure 1. Its corresponding doping and Ge profiles are given in Figure 2. In the resulting energy-band diagram (Figure 3), the Ge profile produces a graded offset that is primarily manifested in the conduction band. Although the inherent band offset caused by the Ge profile occurs in the valence band, it is effectively translated to the conduction band. With a constant p-type doping in the base, both the Fermi level and the energy difference between the Fermi level

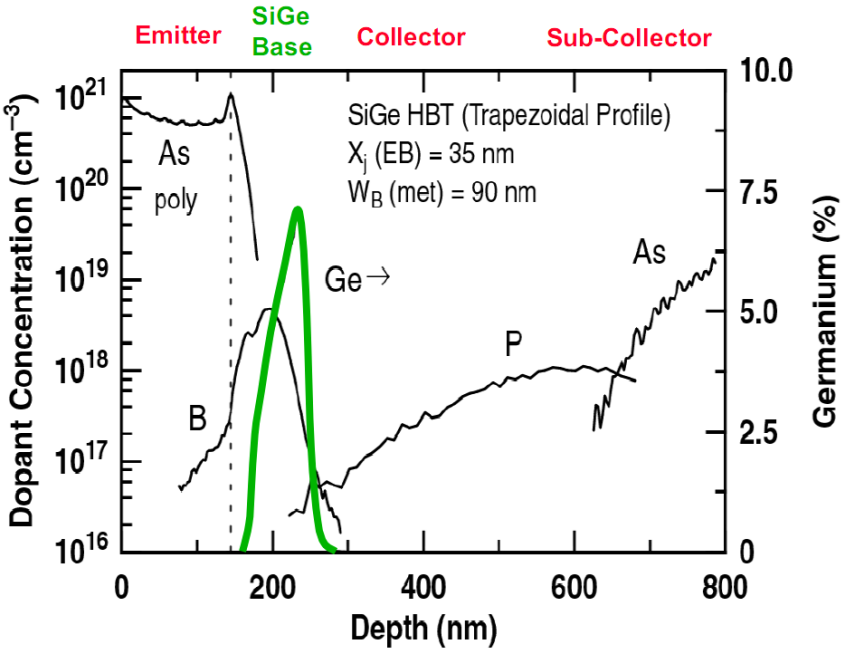

Figure 2: Measured doping profile of a representative first-generation SiGe HBT.

and valence band are fixed; the Ge grading induces a valence band offset, but because the Fermi level must remain constant in equilibrium, it must decrease in energy along with the conduction-band edge.

For DC operation, one fundamental impact of the graded conduction-band offset is to enhance minority-electron transport across the base by inducing a drift field. In addition, the Ge content at the emitter-base (EB) junction will reduce the potential barrier for electron injection from the emitter to the base, yielding exponentially greater electron injection for the same applied V_{BE} (i.e. higher current gain). Finally, a finite Ge content at the collector-base (CB) junction will positively influence the output conductance of the transistor (i.e. higher Early voltage), since the smaller base band gap near the CB junction effectively weights the base profile so that back side depletion of the neutral base with increasing V_{CB} is suppressed [24].

For AC operation, the Ge grading-induced drift field will intuitively lead to a reduced base transit time, which typically is the limiting transit time that determines

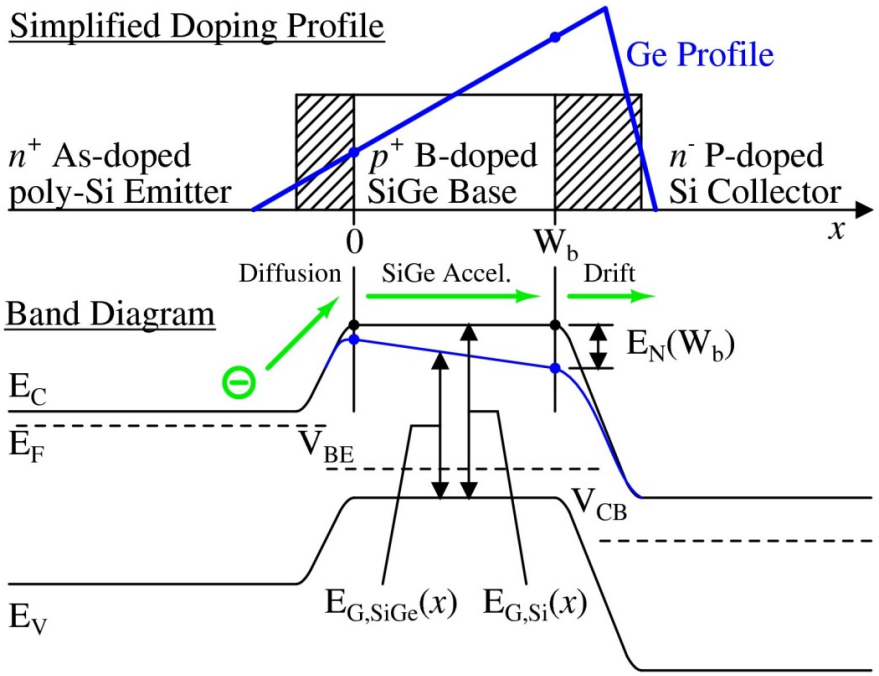

Figure 3: Energy-band diagrams for a Si BJT and a graded-base SiGe HBT, biased in forward active mode at low-injection.

performance metrics such as the maximum operating frequency. In addition, the Ge-enhanced injection of electrons from the emitter into the base dynamically produces a back-injection of holes from the base into the emitter. This reduces the emitter charge-storage delay time, which is reciprocally related to the AC current gain of the transistor [24].

These DC and AC effects are dependent on the profile of the Ge content, especially the mole fraction at the EB junction and the degree of grading across the neutral base. However, trade-offs in profile design exist because of the fact that SiGe film stability limits the total Ge content that can be present. Consequently, different Ge profiles can be designed to achieve specific performance goals. For example, a triangular profile beginning at the EB junction and peaking just inside the CB space-charge region would maximize the frequency performance and Early voltage while providing little improvement to the current gain. A box-shaped profile that is flat across the base would maximize the DC current gain, but would not enhance electron transport across

the base. Alternately, a trapezoidal profile or a profile as illustrated in Figure 3 would simultaneously improve all performance metrics, albeit to a lesser extent.

At present, state-of-the-art npn SiGe HBTs have been demonstrated with peak f_T and f_{max} above 400 GHz at room temperature [20, 117]. Great motivation for continued performance scaling exists as a result of increasing performance requirements for existing RF through mm-wave applications, as well as emerging applications such as mm-wave to sub-mm-wave radars and sensors for security, automotive, and medical applications. With the SiGe HBT breakdown voltage becoming a key challenge to performance scaling as a result of the inherent tradeoff between peak f_T and breakdown voltage, the collector doping profile and Ge retrograde in the CB junction must be carefully designed. Moreover, the development of next-generation SiGe HBTs will require implementation of new structures to minimize base resistance (R_B) and collector-base capacitance (C_{BC}) [116]. Developing and enhancing effective technology-computer-aided-design (TCAD) techniques will remain a key tool in addressing these scaling challenges by identifying the limiting factors during iterative optimization of new device designs.

1.3 Extreme Environment Electronics

Radiation fields result from the magnetosphere and proton and electron belts surrounding the earth. Solar wind particles trapped in the earth's magnetic field result in the "Van Allen" radiation belts, which are particularly concerning for the orbital paths of satellites [60, 61]. As a result of operating in this extreme environment, electronic systems often suffer from degraded performance or altogether fail after a length of time. Radiation-induced damage is of particular concern as technologies scale, since changes to the device structure and fabrication process can potentially lead to increased radiation sensitivity. Device and circuit performance degradation can be attributed to three primary mechanisms: displacement damage, ionization damage,

and single-event effects (SEE). The first two mechanisms are typically addressed together as total-ionizing-dose (TID) damage. SEE can be divided by the various types of errors that can be caused within a circuit. For example, temporary errors include single-event transients (SET), single-event upset (SEU), and multiple-bit upset (MBU), whereas permanent errors include single-event latchup (SEL), single-event burnout (SEB), and single-event gate rupture (SEGR).

1.3.1 Radiation Effects in CMOS

CMOS technologies suffer from increased off-state leakage as a result of TID radiation. The primary cause of leakage in modern CMOS platform is traps that are created along the shallow trench isolation (STI) sidewalls at each end of the transistor, creating leakage paths between the source and drain terminals [29]. TID radiation hardness appears to improve with device scaling, but studies show considerable variability between different manufacturers and even different fabrication lots of the same IC [51]. In SOI CMOS, considerable complexity is introduced with the addition of the SOI buried oxide, in which positive trapped charge can result in TID-induced back-channel leakage. The most effective mitigation technique to achieve multi-Mrad(SiO_2) TID hardness is to employ an annular MOSFET geometry; however, the penalties of this approach include increased layout area and parasitics, limitations in width to length ratio, a lack of existing compact models, and the inherent device asymmetry [29].

Continued scaling of CMOS has led to an increased SEE sensitivity both in SOI and bulk platforms. Although SOI CMOS is inherently immune to latchup and more resistant to SEE because of the elimination of charge collection from the substrate [28], its SEE sensitivity increases nonetheless as circuit switching speeds increase and the amount of charge that represents stored information is reduced [29]. As a result, SEE have become a significant reliability challenge not only for space-based CMOS technologies, but even for advanced Earth-based CMOS technologies. SET may set

fundamental limits on the operating speed of radiation-hardened ICs, and studies have shown that many newer ICs suffer complex failure modes such as single-event functional interrupt (SEFI) that may require a device reconfiguration or power cycle for recovery [29]. Thus, SEE mitigation will be critical to the design of of any reliable circuit operating in terrestrial, high-altitude, or space environments. To date, a variety of mitigation techniques have successfully been implemented, including device-level hardening through cross-coupled feedback resistors [88], circuit-level hardening through internally-redundant storage elements [9], and system-level hardening through error detection and correction circuitry.

1.3.2 Radiation Effects in SiGe HBTs

Due to its inherent tolerance to multi-Mrad(SiO_2) TID radiation and improved DC and AC performance at cryogenic temperatures [23], SiGe BiCMOS technology has emerged as a strong contender for extreme-environment applications such as space-based electronics, which must operate in radiation-rich conditions and at cryogenic temperatures (e.g., 43 K in the shadowed polar craters of the Moon). In bipolar transistors, unlike MOSFETs, the primary transistor action occurs away from any Si-SiO_2 interfaces; thus, the SiGe HBT is inherently hardened to TID damage without any process or layout modifications. Some TID-induced base leakage current does appear as a result of traps created along the EB spacer oxide, but this current is negligible up to multi-Mrad(SiO_2) and TID hardness has been demonstrated across all existing SiGe technology nodes [25, 54, 104]. SEE, however remain an area of concern for space-based SiGe circuits, and the inherent susceptibility of SiGe digital logic circuits to SEU [62, 84] is further compounded by the apparent increase in SEU (proton) sensitivity at cryogenic temperatures [102].

To mitigate SEE in SiGe ICs, a variety of circuit- [50, 62, 76] and device-level hardening techniques [82, 103] have been implemented with minimal impact on system

complexity. These radiation-hardening-by-design (RHBD) techniques can be supported considerably by modeling and simulation—at the device level through 3-D physical TCAD simulations of ionizing radiation effects, and at the circuit level either through traditional compact modeling [71, 78] or true mixed-mode simulations (compact models + 3-D TCAD) [106, 108]. Nevertheless, effective optimization of RHBD techniques can only be performed when there is sufficient fidelity between simulated and measured SET. The SEE response at the circuit level depends heavily on the circuit topology as a result of feedback effects, varying device biases, and for certain circuits, dynamic biases that evolve on the same time scale as that of measured device SET. Moreover, the importance of addressing this issue when modeling SET grows as circuit response times scale and become comparable to the duration of the individual transistor transients [106]. Clear guidelines must be established as to which approaches to modeling SET are valid for various conditions (circuit topology, technology node, device geometry, environment, etc.).

1.4 Long-Term Device and Circuit Reliability

In addition to radiation-related reliability concerns, device technologies must be ensured to be sufficiently immune to all types of degradation mechanisms associated with any extreme operating conditions. Circuits and systems impose a wide range of voltage and current conditions that can cause degradation of the device building blocks, passive elements such as inductors and capacitors, and back-end-of-line (BEOL) metal interconnects. Reliability of a given technology means that under typical circuit operating conditions, the circuits—and the systems ultimately constructed from those circuits—must not wear out or degrade to a level at which they fail over the entire functional life of the system. Reliability is ensured by extensive testing of each individual component of a given technology, each of which possesses unique degradation mechanisms.

Considering bipolar transistors, reliability has historically been ensured by first subjecting the devices to extreme operating conditions for extended periods of time, then quantifying the change in device figures-of-merit (FoM), and finally inferring the maximum stress conditions that ensure a tolerable change in those FoM over the lifetime of the device. Traditionally, the two extreme operating conditions imposed upon bipolar transistors have been: (1) a larger reverse bias stress applied to the emitter-base junction, causing hot-carrier damage (hot electrons, hot holes, or both), and (2) a high forward collector current density stress. Accelerated stress conditions are typically applied to minimize the stress time required to produce sufficient damage, and testing is performed at either elevated (for high J_C stress) or reduced (for reverse EB stress) temperatures to impose worst-case stress conditions. During a technology's qualification process, various process parameters will be tuned until all of the desired reliability metrics have been met. This methodology has been the standard practice for bipolar technologies for the past several decades, but more recent studies of the high-speed SiGe HBT prove that this methodology does not capture all possible degradation mechanisms [22]

The modern SiGe HBT has evolved significantly from the older Si bipolar technologies that were originally designed primarily for digital logic applications. The device speeds are much higher (with much higher peak doping concentrations), and in addition to the Ge grading of the SiGe base, many advancements have been made in the physical device structure that make it radically different from its Si BJT predecessor. Consequently, the full scope of possible degradation mechanisms must be reconsidered, including reverse EB and forward-J_C stress, but also addressing the impact of Ge film stability on process yield, radiation-induced degradation, breakdown voltage, bias instabilities, and impact-ionization-induced mixed-mode stress. The mixed-mode stress degradation mechanism, first reported in [118], poses a unique challenge for RF and mixed-signal applications because it arises when a high collector

current density and high collector-base voltage are applied simultaneously to a SiGe HBT; these of course are common bias conditions for many RF and mixed-signal circuits. It can be distinguished from conventional reverse EB and forward-J_C stress by the inverse-mode I_B degradation as well as its unique geometrical dependence. Fundamentally, the mixed-mode degradation mechanism results from hot carriers that originate in the collector-base junction, traverse to the shallow-trench isolation (STI) oxide and emitter-base (EB) spacer oxide interfaces, and depassivate silicon dangling bonds to cause a net increase in interface traps.

A variety of experimental [11, 16, 18, 118, 119, 120] and theoretical studies [15, 16, 110] of the mixed-mode degradation mechanism have been published. In the theoretical work, ideas have often been appropriated from studies of CMOS reliability and hot-carrier injection, such as the lucky-electron model [38, 41] and the reaction-diffusion model [42, 77, 83] to explain the dynamic nature of trap passivation and depassivation at the Si-SiO_2 interfaces. What has been made clear through all of this work is that the degradation mechanisms of the SiGe HBT are by nature highly dependent on the specific device technology, circuit type, and time-dependent circuit operating conditions. The reliability response of single device at fixed bias conditions does not provide sufficient information to assess overall circuit and system reliability, since the mixed-mode degradation mechanism exhibits a complex spectrum of damage and annealing regions [18, 19]. Thus, much more research is needed to establish methods by which the overall degradation response of modern mixed-signal circuits can be accurately predicted.

1.5 Research Objectives

Throughout the history of the semiconductor industry, the principal driving force behind IC technology innovation has been device scaling. The reduction of device dimensions has led to increasing levels of integration, a trend recognized by the

rise of labels such as small-scale integration (SSI), medium-scale integration (MSI), large-scale integration (LSI), and very-large-scale integration (VLSI), which refer to the total number of transistors on a given IC. At present, ICs with greater than one billion transistors are commonly available. However, simply scaling transistor dimensions to improve speed and transistor count is not sufficient to maintain the technological growth necessary to meet the challenges of emerging mixed-signal and RF applications. Replacements are needed for costly III-V IC components to drive down costs and facilitate design of complete SoC solutions. This requires continued scaling of silicon-based transistor performance, improving peak cutoff frequencies, gain, and other relevant performance metrics while minimizing mismatch, noise, and reliability concerns. Furthermore, as mixed-signal technologies become more widely accepted, they must overcome the additional hurdle of validation for reliable operation in extreme environments if they are to be employed within satellite or other space-based electronic systems. SiGe BiCMOS technology in particular is positioned as an excellent candidate to satisfy all of these requirements. The present research addresses performance-scaling and reliability of the SiGe HBT as well as operation of state-of-the-art CMOS and HBT devices in extreme environments. The purpose of this research is to develop enhanced predictive modeling tools that can be used to design new devices and subsequently predict their reliability on Earth and in extreme environments. Experimental measurements have been used to benchmark and validate these tools.

A significant amount of research has been completed to enhance predictive-modeling capabilities for advanced broadband technologies that have great promise for extreme-environment applications. This research has leveraged access to the IBM 45-nm RF-CMOS platform as well as multiple generations of commercially-available SiGe BiCMOS technology to conduct experimental measurements for support and validation of new modeling approaches. Contributions of this research include improved modeling

tools to facilitate continued performance scaling of the SiGe HBT, calibration of TCAD-compatible physical models for extreme environments, new measurements of device-level single-event transients, and improved modeling of circuit-level single-event transients. First, an improved 2-D transit time analysis is introduced in Chapter 2. This tool greatly enhances device optimization by enabling device designers to very quickly extract the f_T from a TCAD device model and identify the spatial distribution of the contributions to the total device delay. Subsequently, experimental measurements of recombination lifetime and resistance in a SiGe BiCMOS technology are reported in Chapter 3 and used to calibrate models of fundamental carrier-transport parameters that are key to accurate modeling of single-event effects. Chapter 4 presents and analyzes experimental measurements of SETs captured for the first time from a commercial 45-nm RF-CMOS technology. Chapter 5 builds upon the foundation of the material and device research from Chapters 2-4 to achieve a new understanding of underlying SET mechanisms and establish best practices for modeling circuit-level SET. In Section 5.2, existing SET data from a SiGe voltage reference is used to investigate the accuracy of various circuit-level SET modeling approaches, identifying the limitations of each approach and providing new insight into best practices for modeling circuit SET in different circuit topologies and device technologies. A similar approach is then applied in Section 5.3 to a building-block of digital circuits, the D flip flop, to reveal pitfalls in conventional circuit SET modeling approaches and highlight coupled mixed-mode TCAD simulations as an essential tool for understanding SEE in modern IC technologies. Chapter 6 extends this mixed-mode TCAD approach to address long-term device and circuit reliability due to hot carrier-induced degradation. In addressing circuit reliability associated with transient radiation effects, physics-based TCAD models of ion-strike charge collection are embedded in a compact model circuit environment to assess the impact on circuit operation. In the same way, physics-based TCAD models can be used to model device operation at or beyond classical "safe

operating area" (SOA) limits, where complexities arise such as pinch-in instabilities, thermal effects, current-dependent breakdown phenomena, complex 3-D effects, and stress-induced leakage. Finally, Chapter 7 will summarize the key contributions of this research and highlight the critical topics for which further research is needed to support the development of high-performance, reliable broadband IC technologies.

CHAPTER II

IMPROVED TRANSIT-TIME ANALYSIS FOR SIGE HBT PERFORMANCE OPTIMIZATION

Developing and enhancing effective TCAD optimization techniques will enable the design of more highly scaled device profiles. One such technique is the quasi-static transit-time analysis [26, 99, 109], which enables a detailed look at the regional contributions to the total transit time through the physical device. A key advantage of this technique is that it allows one to rapidly quantify and visualize the dynamic performance of a given device profile in greater detail than traditional frequency-domain simulations, and in only a fraction of the computational time. Moreover, since this technique is based on a DC bias-sweep simulation, the current gain β can be extracted simultaneously with the cutoff frequency, f_T. Variations of this technique have been successfully applied to npn SiGe HBTs at several different scaling nodes [2, 6, 21, 53, 96, 97] as well as complementary-SiGe (C-SiGe) HBTs [13].

In addition to estimating the cutoff frequency of a device, transit-time analysis can also provide insight into the regional contributions to its total delay time. For a 1-D transit-time analysis, the extraction of regional transit-time components is relatively straightforward [2, 53, 96, 109]. However, proper optimization of modern SiGe HBTs requires 2-D or even 3-D simulations to capture the intricacies of effects such as the high-injection Kirk effect and heterojunction-barrier effect (HBE). The result is an increase in the complexity of defining region boundaries and integrating over these to obtain the regional transit times. Previous 2-D transit-time studies have estimated the regional transit-time components, but these were limited to integrations over fixed regions chosen based on the doping profiles and the intrinsic and extrinsic

areas of the device [21, 97]. Although these provide rough approximations of the transit-time components, the region boundaries are more correctly delineated by the carrier dynamics throughout the active device. Moreover, these boundaries will change dynamically according to the device operating conditions, particularly at high-injection.

With this in mind, the traditional transit time analysis is enhanced here by introducing a 2-D streamline analysis that enables direct extraction of the regional transit times using bias-dependent region boundaries. This analysis tool is fully automated and integrated within commercially-available TCAD software [105] and is thus well-suited for TCAD-based optimized scaling of SiGe HBTs.

2.1 2-D Transit Time Analysis

The basis for the 2-D transit-time analysis is described in [99, 109]. In the quasi-static approximation, the transit time for a 2-D npn bipolar transistor from the emitter contact to the collector contact is given by

$$\tau_{ec} = \frac{1}{2\pi f_T} = \frac{\Delta Q_{tot}}{\Delta I_C}\Big|_{V_{CE}} = \frac{q}{\Delta I_C}\iint \Delta n(x,y)\Big|_{V_{CE}} dydx. \tag{1}$$

The small-signal quantities ΔJ_C and Δn are calculated from a pair of static device simulations modeling a small voltage perturbation, V_{BE}, around a DC bias point. For the results presented here, hydrodynamic simulations were performed in which the DC bias across the emitter-base junction was swept from low-injection to high-injection regimes, with a small-signal perturbation of $V_{BE} = 1$ mV, and fixed collector-base bias of $V_{CB} = 0.5$ V. Negligible change was observed in the estimated f_T for V_{BE} values from 1 mV to 10 mV. The simulation mesh was optimized and refined throughout the active device regions to ensure a precise representation of the 2-D charge distributions used in the transit-time calculations.

The transit-time streamline analysis resembles that of previous 1-D transit time analyses [2, 26, 96, 109] in that the cumulative transit time, τ_n, is defined along the

path of electron transport (or hole transport for pnp HBTs), which can then be divided into regional transit times by various methods. The fundamental difference, however, is that for a 1-D simulation there is a single carrier transport path, whereas for a 2-D simulation domain this path must be chosen carefully. Furthermore, in calculating τ_n, the local current density varies along the streamline path and must be used in place of J_C. Consequently, the cumulative transit time along the length of a streamline is given by

$$\tau_n(s) = q \int_0^s \frac{\Delta(x(s), y(s))}{|\Delta\vec{j}_n(x(s), y(s))|}\Big|_{V_{CE}} ds. \tag{2}$$

A desired streamline path can be generated for any point of interest from the small-signal electron current-density vector field, which is calculated as the difference in large-signal current-density vector fields of the two static solutions:

$$\Delta\vec{j}_n(x, y) = \vec{J}_{n2}(x, y) - \vec{J}_{n1}(x, y). \tag{3}$$

This is in contrast to the 3-D streamline analysis of [6], in which streamlines were computed from the large-signal current-density vector field. The small-signal current must be used since this is a small-signal analysis. This becomes critical when 2-D high-injection effects are present. It is important to note that this analysis can also be directly extended to 3-D TCAD models without any fundamental changes, and also applied to pnp SiGe HBT optimization.

2.2 Regional Transit Times

For highly-scaled bipolar transistors, the depletion approximation becomes invalid in the base and thus abrupt space-charge-region boundaries cannot be clearly identified [26]. Furthermore, the region in the base where $\Delta n \approx \Delta p$ disappears for sufficiently scaled devices [2, 96]. Although this indicates that diffusive delay in the base of such devices is negligible at low injection, there remains utility in subdividing space-charge and quasi-neutral-base transit times to capture the sharp increase in the

quasi-neutral-base transit time at high injection. For the analysis presented here, boundary definitions were chosen based on carrier and dopant densities to maximize the robustness of the boundary algorithm. The space-charge boundary in the emitter is defined to be the position at which the majority carrier concentration comes within 10% of the background doping concentration. At low injection, the same definition is used for the space-charge boundaries in the base. At high injection, this definition cannot be used, as the carrier densities increase well beyond the base doping concentration; instead, the base boundaries are chosen simply to be the position where $n = p$. The collector space-charge boundary is defined in the same way as the emitter boundary, with the exception that the majority carrier density is adjusted by subtracting the density of mobile carriers that comprise the collector current, J_C/qv_{sat}. Although velocity overshoot occurs within the collector-base space-charge region, the carriers are assumed to exit the space-charge region at approximately v_{sat}. The definition of these four boundaries gives rise to five regional transit-time components: the quasi-neutral transit times, τ_e, τ_b, and τ_c, and the space-charge transit times, τ_{be} and τ_{cb}.

A key advantage of this regional analysis over previous 2-D studies [21, 97] is that the region boundaries are bias dependent. To accurately reflect the regional transit times at high injection, it is critical that the base boundary follow the base push-out into the collector with the onset of the Kirk effect. Moreover, this analysis allows the space charge and quasi-neutral regions to be separately defined. Since the streamline computes the delay along a single path, multiple streamlines can be computed through various areas of the device to ascertain the regional transit times throughout the simulation domain.

2.3 Simulation Results

To verify that the quasistatic transit time analysis can be substituted for conventional frequency-domain simulations, the cutoff frequency of a calibrated model of a 50 GHz

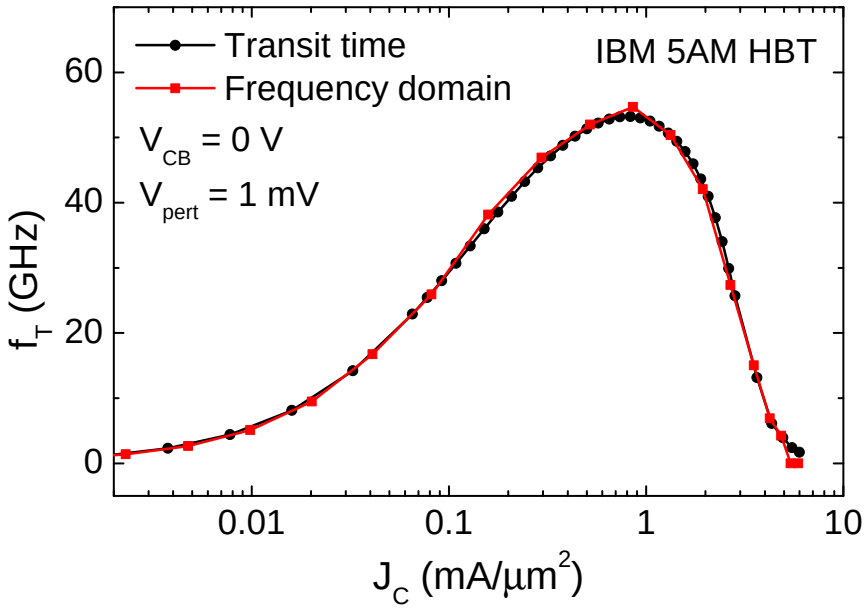

Figure 4: Comparison of f_T vs. J_C curves simulated for a calibrated model of a 50 GHz npn SiGe HBT, using both the 2-D transit time analysis and frequency-domain simulations.

SiGe HBT (IBM 5AM SiGe BiCMOS) was computed using both types of simulations. As shown in Figure 4, the two simulations produce remarkably similar results across the entire range of bias conditions.

2.3.1 Scaled npn SiGe HBTs

The 2-D transit-time analysis was performed on five prototype npn SiGe HBTs: GT0-GT4, with GT0 being a calibrated 200 GHz (f_T) SiGe HBT [114], GT1 a 375 GHz SiGe HBT, GT2 a 450 GHz SiGe HBT, GT3 a 550 GHz SiGe HBT, and GT4 a 700 GHz SiGe HBT [113]. A comparison of the transit-time f_T calculation with the f_T computed from hydrodynamic frequency-domain simulations is given in Figure 5 for GT0-GT2, demonstrating close agreement across the entire bias range and validating the suitability of substituting transit-time simulations in place of frequency-domain simulations even for highly scaled devices. The transit-time streamline analysis is explored in detail using the hypothetical GT1 transistor. Figure 6 shows a small-signal streamline at $V_{BE} = 0.8$ V originating near the center of the emitter (x = 0.01 μm)

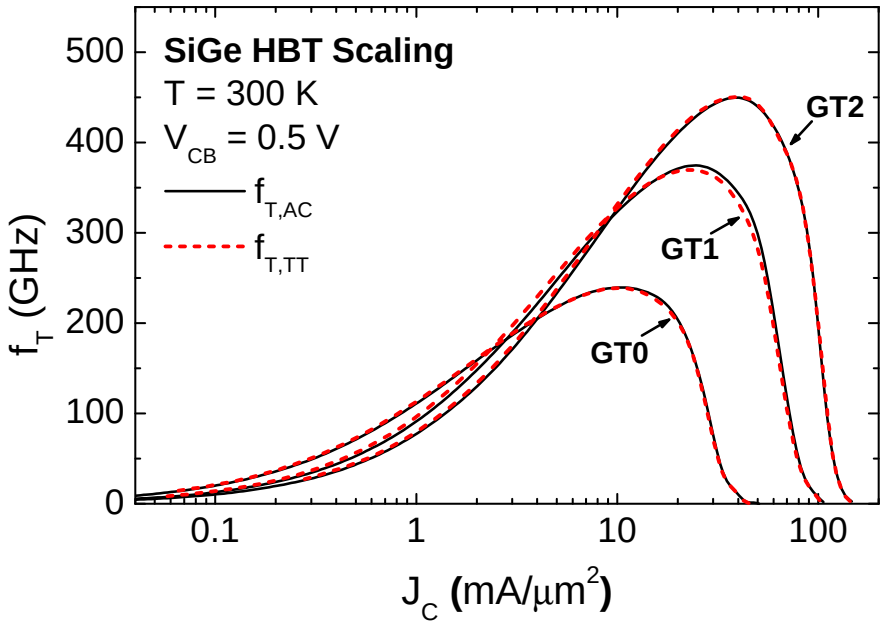

Figure 5: Comparison of f_T vs. J_C curves for scaled npn SiGe HBTs as computed by the 2-D transit time analysis and frequency-domain simulations.

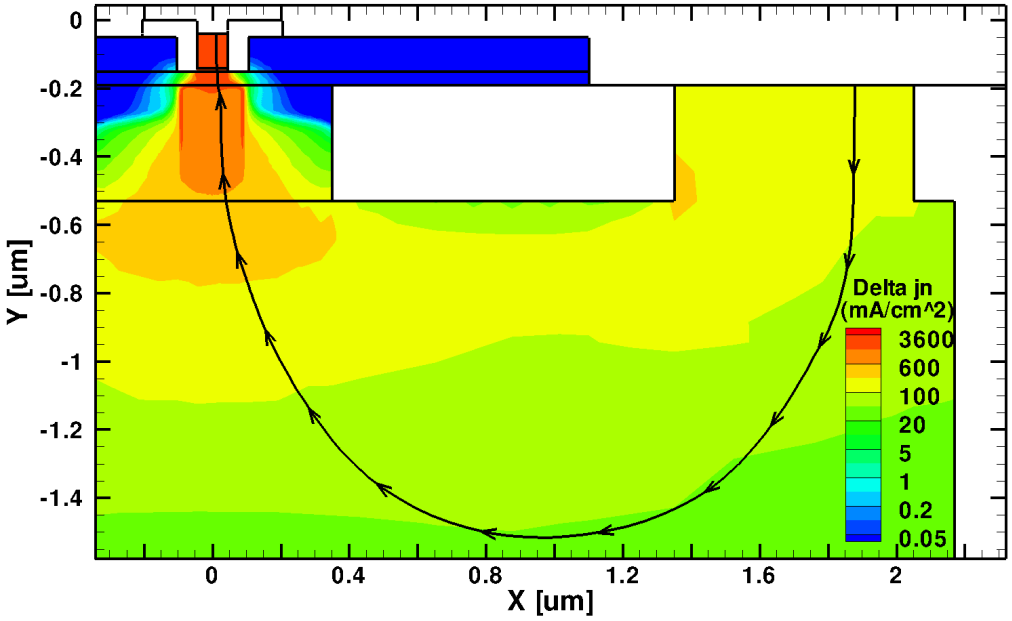

Figure 6: Example streamline computed for GT1 at $V_{BE} = 0.8\,V$ and $V_{CB} = 0.5\,V$.

and extending to the collector contact. The cumulative small-signal transit time was computed at DC bias points ranging from $V_{BE} = 0.7\,V$ to $V_{BE} = 1.05\,V$, illustrated in Figure 7 as a function of distance traversed along the streamline. It is important to note that the transit time is not a measure of the physical time it takes for a single carrier to transit from emitter to collector, but rather is a function of the change in the distributions of the carrier concentration and the current density. As such, the decrease in transit time near the collector space-charge boundary is not unphysical; instead, its physical origin is the extension of the collector-base space-charge region toward the sub-collector, which causes a decrease in the electron concentration at the edge the space-charge region, thus reducing the total transit time [26, 96]. At low V_{BE} (0.7 V and 0.8 V), the base-emitter space-charge transit time dominates the total transit time. As the bias is increased, the higher transconductance leads to a reduction in the space-charge transit times, τ_{be} and τ_{cb}. The relative contribution of τ_{be} is reduced while that of τ_{cb} is increased, since the reduction in τ_{cb} is limited by the mobile electron storage in the collector-base space-charge region. The peak cutoff frequency occurs just before the suppressed τ_{be} and τ_{cb} are overcome by the sharp increase in τ_b at the onset of the Kirk effect and HBE ($V_{BE} = 0.9\,V$). Early onset of the Kirk effect is evident at $V_{BE} = 0.95\,V$, with an increased base width and a higher contribution of τ_b to the total transit time. HBE is in full effect at $V_{BE} = 1.05\,V$, causing τ_b to dominate the total transit time. These trends are clearly exhibited in Figures 8-10, which plot the regional transit times as a function of collector current density for transistors GT0-GT2. One intriguing scaling trend for these devices is the increasing contribution of τ_{cb} to the total transit time at peak f_T, highlighting the importance of collector profile optimization.

The scaling trends are more clearly shown in Figure 11, which gives the relative magnitude of each regional transit time component when the GT0-GT4 are biased at peak f_T. The largest component of the total transit time is consistently the collector

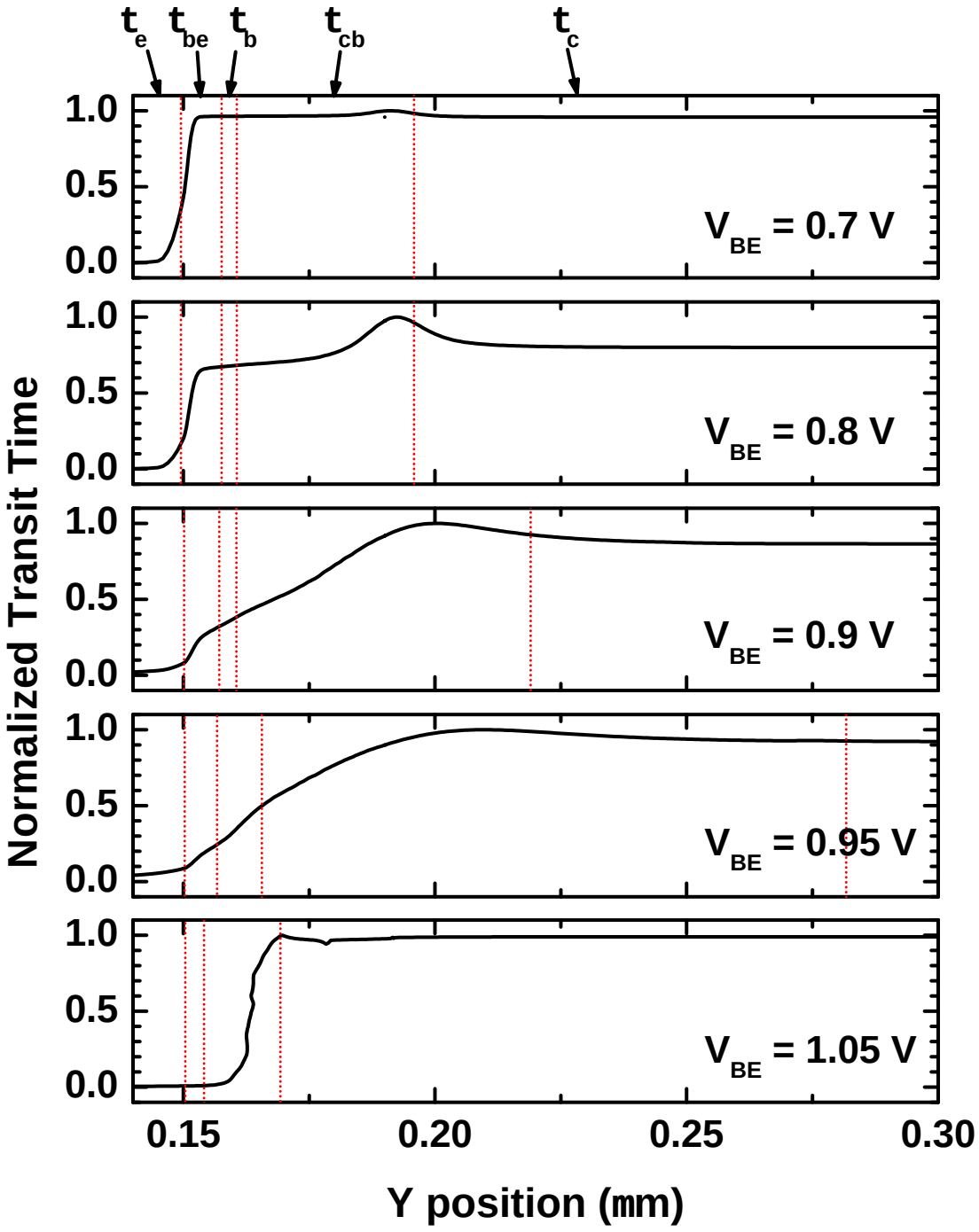

Figure 7: Normalized cumulative transit time for GT1. Computed region boundaries are indicated by vertical red lines, with emitter on the left.

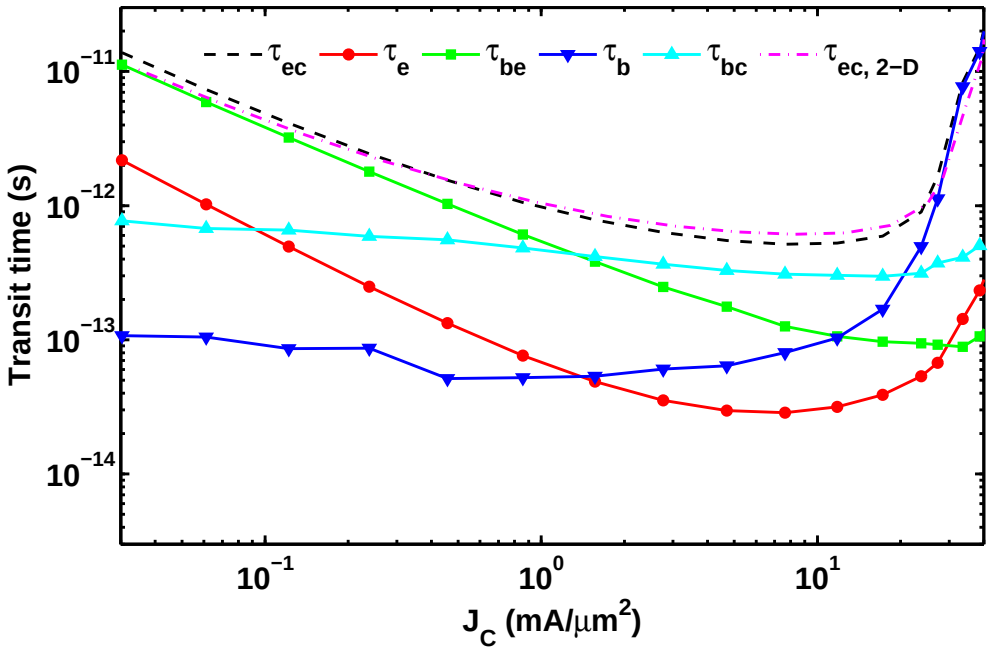

Figure 8: Regional transit times vs. bias for GT0 with $V_{CB} = 0.5\,V$.

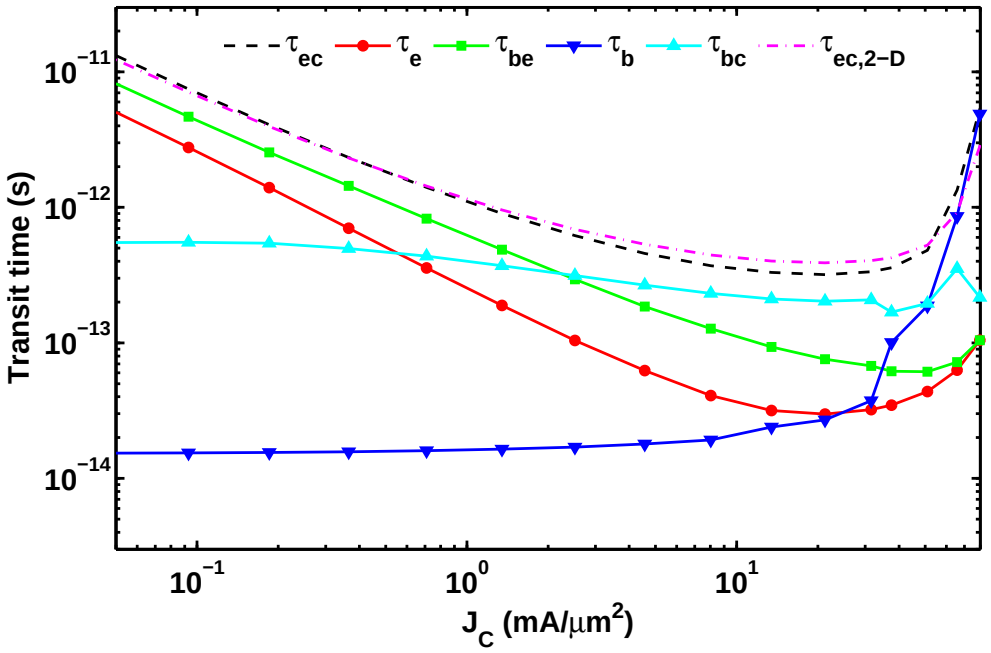

Figure 9: Regional transit times vs. bias for GT1 with $V_{CB} = 0.5\,V$

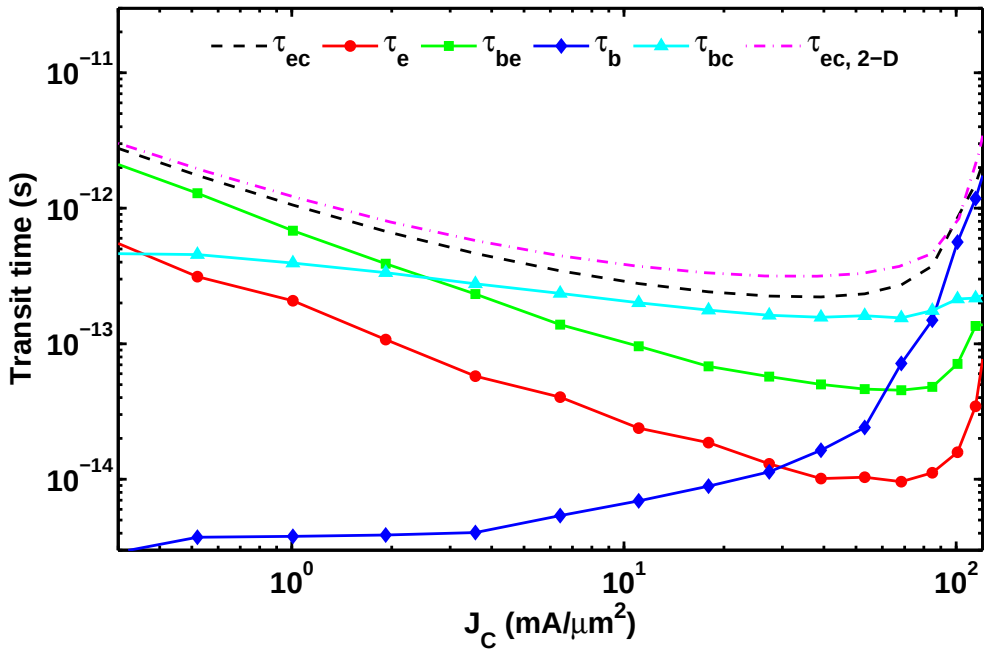

Figure 10: Regional transit times vs. bias for GT2 with $V_{CB} = 0.5\,V$

transit time. With increasing vertical scaling, the relative contribution of the neutral base transit time is reduced, and the collector transit time becomes increasingly dominant. This trend is unavoidably linked to vertical scaling of the HBT. A smaller base width and higher base doping will strongly reduce the base and base-emitter transit times. The collector transit time will also be reduced as the collector doping is increased, but to a lesser degree, since it can only be improved by increasing the collector doping, and that increase is limited by breakdown constraints.

2.3.2 Application Towards a C-SiGe Scaling Roadmap

In [12], a complementary SiGe HBT scaling roadmap is presented, including candidate profiles for matched npn and pnp SiGe HBTs at the 100 GHz and 200 GHz peak-f_T performance nodes. This roadmap was developed using a sophisticated TCAD framework that enables realistic estimates of the key device performance metrics based on highly parametrized device models, which in turn enable predictive optimization toward targeted performance nodes. Integrated within this TCAD framework is the

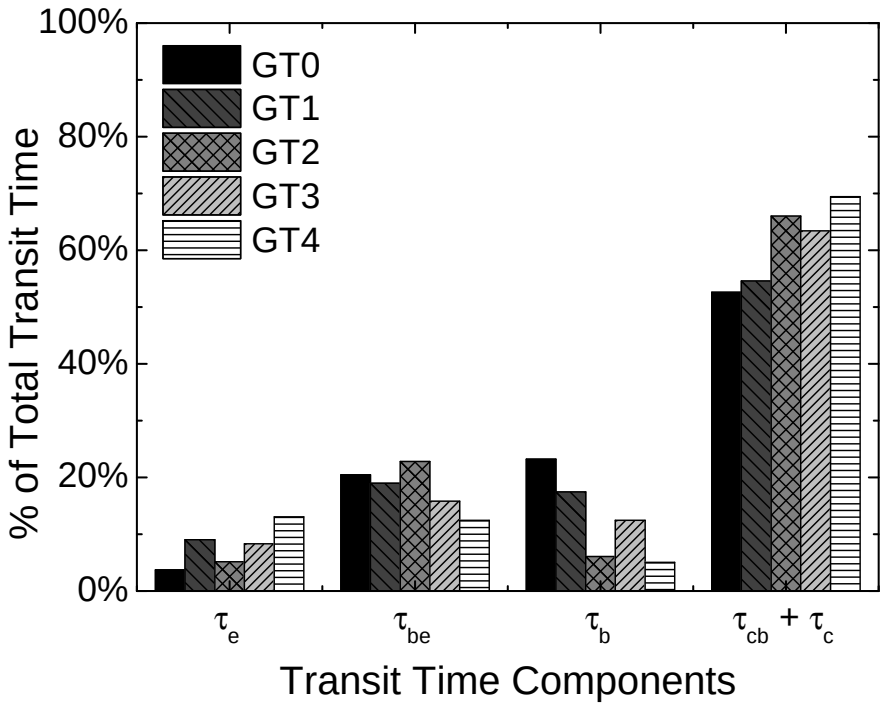

Figure 11: Regional transit time components for scaled npn SiGe HBTs biased at peak-f_T, normalized to total transit time.

quasistatic transit time analysis presented here, including the streamline-based regional transit time analysis. The utility of the transit time analysis was demonstrated through its use in the simultaneous optimization of the vertical profiles of both npn and pnp SiGe HBTs to achieve matched dc and ac performance across a wide range of bias conditions, which is no small challenge because of the valence band offset and lower minority carrier mobility associated with pnp SiGe HBT design [13]. The regional transit time analysis was very helpful to fine tune the doping and Ge profiles based on the limiting regions across bias and the onset of the Kirk effect and HBE.

Figure 12 shows the close correlation of the ac performance (f_T, f_{max}) of the npn and pnp devices at the highest performance node optimized in [12] (200 GHz). From the regional transit time analysis, the transit time components are shown in Figure 13, in which similar trends are shown for both npn and pnp devices. The total transit time, τ_{ec}, is closely matched by design. Moderate differences exist between several

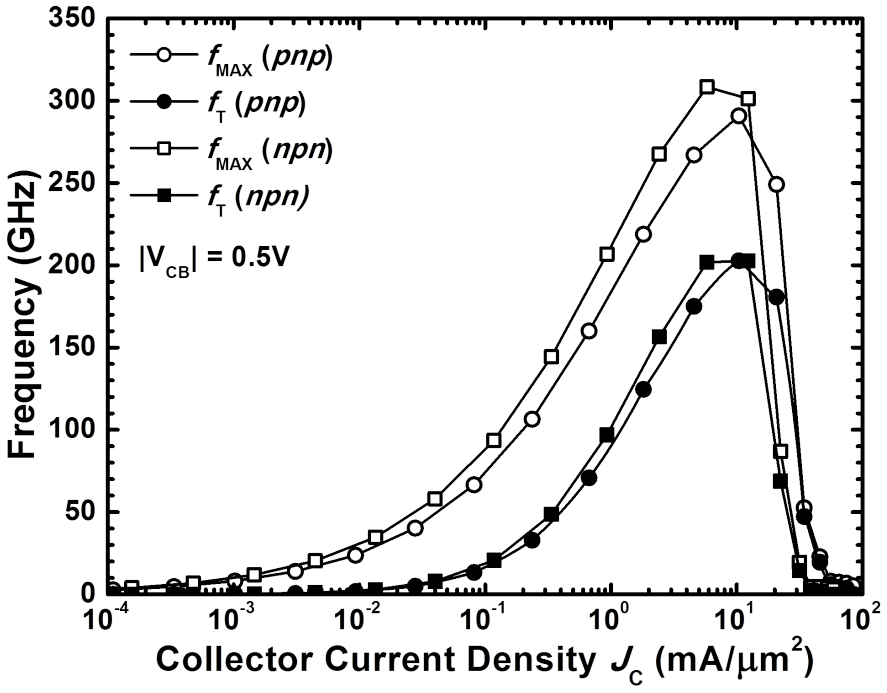

Figure 12: Comparison of the ac performance (f_T, f_{max}) for the calibrated 200 GHz C-SiGe HBT models.

of the individual components; this can be attributed to differences in the vertical profiles as well as to differences in hole vs. electron mobilities. For example, the emitter transit time of the pnp device is relatively higher at low-injection compared to the npn device because of a slightly lower current gain at those bias conditions, although it becomes comparable to that of the npn at peak-f_T, where the current gains are well matched. At higher injection, the base transit time of the npn device is larger than that of the pnp device because of the higher collector doping that delays the onset of the Kirk effect. In both devices, however, the same general trends are observed as were for the scaled npn SiGe HBT models of the previous section, where the base-emitter transit time dominates at low injection, the collector transit time dominates near peak-f_T, and the base transit time increases sharply following the onset of high-injection effects. In Figure 14, the regional transit times at peak-f_T are shown for the 100 GHz and 200 GHz C-SiGe device models as a fraction of the total transit time, with the results from the calibrated 200 GHz npn SiGe HBT model

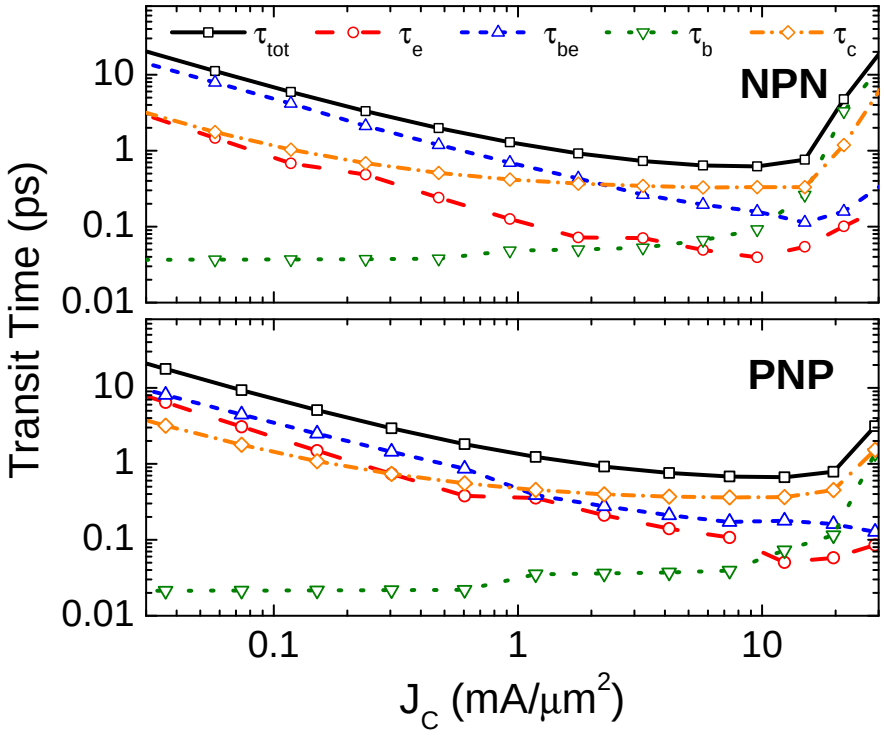

Figure 13: Regional transit time components for the 200 GHz C-SiGe HBTs vs. collector current density.

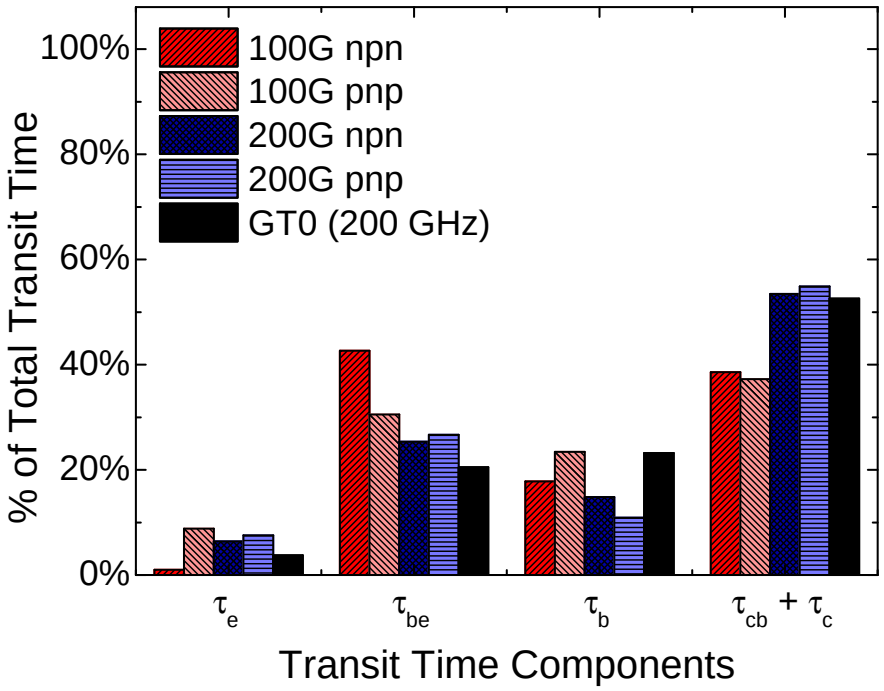

Figure 14: Regional transit time components for 100 GHz and 200 GHz C-SiGe HBTs biased at peak-f_T, normalized to total transit time.

also shown for reference. Comparing with the results of Figure 11, there is a close correlation between the 200 GHz devices due to their similar doping levels and base widths. Furthermore, the transit time components of the 100 GHz devices fall as expected on the scaling continuum, with the contribution of the collector transit time relatively smaller as compared to the higher performance nodes.

2.4 High-Injection Effects

At the onset of high injection, the small-signal streamline analysis reveals a gradual two-dimensional onset of the Kirk effect and HBE across the width of the collector-base junction. Base push-out first occurs at the center of the device where the current density is greatest, exposing the retrograded Ge profile [90]. The accumulation of holes at the valence band barrier induces a conduction-band barrier that inhibits electron transport. Since the barrier first appears at the center of the device, the small-signal electron transport path is diverted laterally within the base to flow around the barrier, as illustrated in Figure 15. As the electron current density in the collector-base junction increases with bias, the onset of the Kirk effect and associated HBE expands laterally, driving the small-signal electron transport path further out along the width of the collector, as shown in Figure 16. At sufficient bias, HBE is in effect across the width of the collector-base junction; thus, the small-signal path relaxes toward the center of the device as it is forced to flow through the conduction-band barrier. From this point onward, current spreading is primarily resistive.

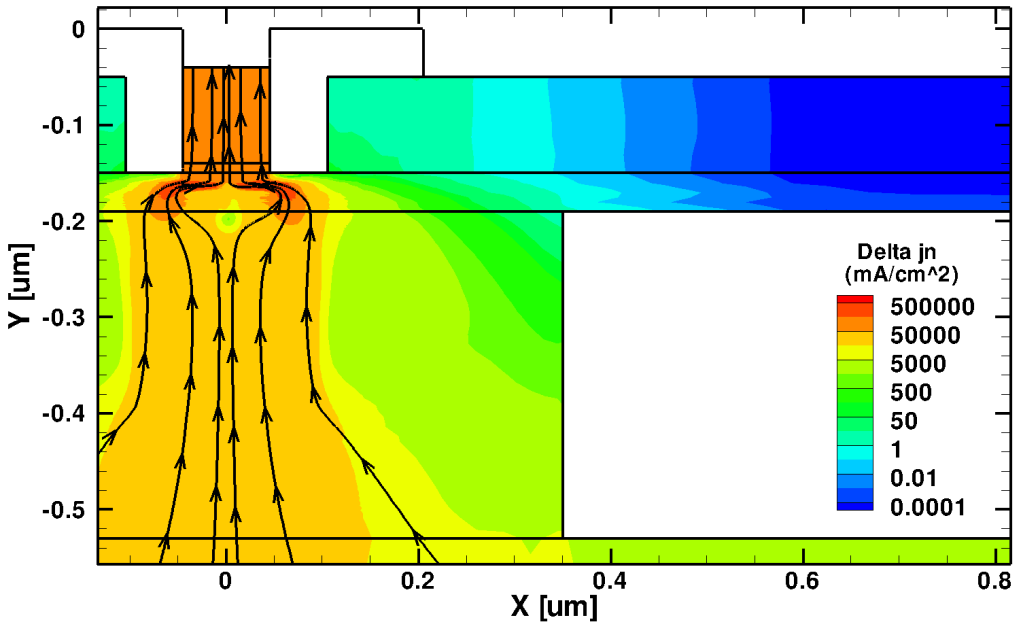

Figure 15: Streamlines computed for GT1 at early onset of HBE, with $V_{BE} = 1.01\,V$ and $V_{CB} = 0.5\,V$.

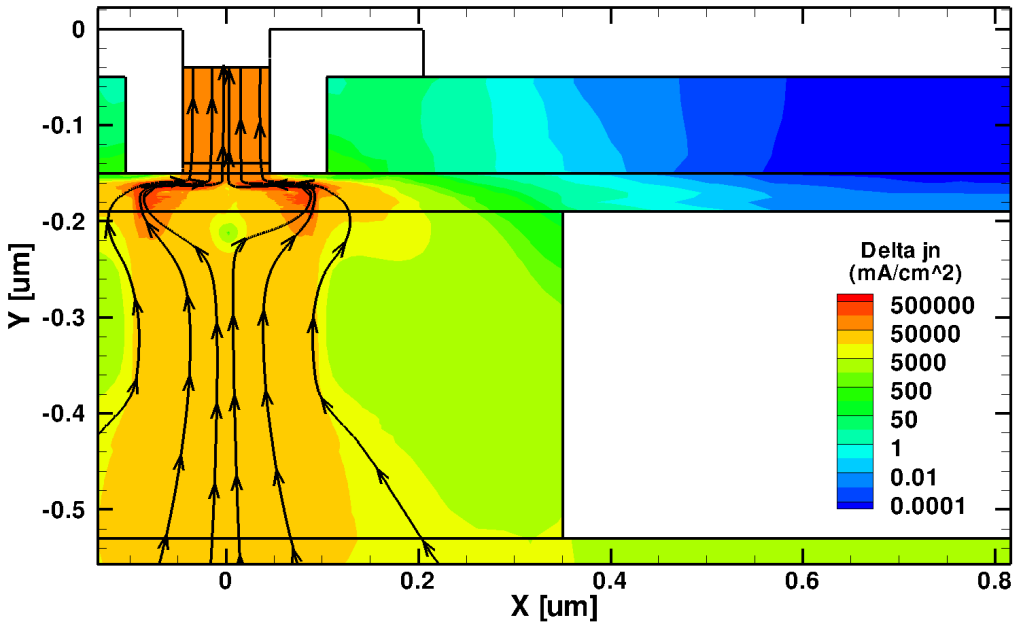

Figure 16: Streamlines computed for GT1 with $V_{BE} = 1.03\,V$ and $V_{CB} = 0.5\,V$.

CHAPTER III

MEASUREMENT AND MODELING OF CARRIER TRANSPORT PARAMETERS FOR EXTREME ENVIRONMENTS

SEU hardening techniques can be further developed and optimized by analyzing the transient dynamics of charge generated by a heavy-ion strike. This is most easily accomplished with TCAD simulations of charge collection at the terminals of an HBT following a heavy-ion strike, which can be used to study the influence of variables such as ion species, ion energy, strike angle, and strike location [70, 75, 81, 111]. Such simulations can directly evaluate device-level radiation-hardening techniques or, by introducing the simulated strike-induced current transient into circuit simulations, they can also be used to assess circuit implications and evaluate circuit-level hardening techniques. However, the validity of these simulations is entirely contingent upon the integrity of the underlying physical parameter models.

Calibrated recombination, mobility, and ionization models must be developed to accurately predict the behavior of SiGe HBTs and the circuits in which they are employed. Consequently, experimental measurements of these parameters must be collected from within the actual technology that will be utilized, a fact that is particularly evident for parameters such as recombination lifetime, whose magnitude and temperature dependence vary greatly depending on the composition of defects and impurities present in a given technology. This research addresses the theoretical temperature dependence of the Shockley-Read-Hall (SRH) recombination lifetime [37, 98]. Experimental measurements of recombination lifetime from the substrate of IBM's 8HP SiGe BiCMOS technology are presented, including the lifetime response to 63.3 MeV proton

irradiation. This data are then used to develop calibrated temperature-dependent parameter models for TCAD. Similarly, experimental measurements of resistance from different regions within IBM's 5HP/5AM SiGe BiCMOS technology are presented, and then used to calibrate TCAD-compatible parameter models for mobility and incomplete ionization.

3.1 Minority Carrier Recombination

Few published temperature-dependent bulk lifetime measurements exist for commercial silicon-based technologies. Furthermore, published data such as in [34] are limited to a temperature range of 300 K to 400 K. Despite the solid foundations of SRH theory and many recent contributions by the photovoltaic community, carrier lifetime studies have been either restricted to theory or confined to laboratory test cases designed to characterize the behavior of a particular impurity or defect. For example, these studies characterize the recombination behavior of individual impurities such as copper [56], iron [7, 87], aluminum [89], and nickel [55], along with intrinsic defects such as the boron-related defect [57], oxygen-related defect [92], metastable defect in boron-doped Czochralski-grown silicon [86], and radiation-induced defects [8, 44]. However, despite all of these prior contributions, accurate modeling of "real-world" high-speed devices necessitates an empirical knowledge of carrier lifetimes within the actual technology that will be used.

3.1.1 Experimental Details

The minority electron lifetime in the 8–10 $\Omega \cdot$ cm p– substrate of IBM's 8HP SiGe BiCMOS technology was measured using custom diode structures constructed from the substrate-subcollector junction. Similarly, the minority hole lifetime in the n-well region was measured using diodes constructed from a p+ diffusion layer (extrinsic base) within the n-well. A representative cross-section of a SiGe HBT that depicts these regions is given in Figure 1. Multiple samples of both $100 \times 100\,\mu m^2$ and

$200 \times 200\,\mu m^2$ diodes were measured across temperature and injection level.

The compensated open circuit voltage decay (OCVD) technique for lifetime measurement [35] was chosen because it is purely electrical, can measure lifetimes in the bulk of the semiconductor, can be used to measure injection dependence by varying the diode current, and is compatible with packaged measurements within cryogenic and high-temperature test fixtures. Abrupt current pulses were applied to the test diodes by passing a 1 kHz square wave through a 1N4148 fast-switching diode while an oscilloscope captured the transient voltage decay. The introduction of a properly tuned shunt resistance can effectively cancel the effects of excess parasitic capacitance, restoring the ideal linear decay region necessary for lifetime extraction. Accordingly, a tunable shunt resistance was provided by a Keithley switching matrix along with a custom PCB containing a range of resistors. Fig. 17 shows an example of how the required linear voltage decay was recovered by varying the shunt resistance with all other conditions fixed. Low temperature measurements (50–325 K) were carried out using a closed-cycle liquid-helium cryogenic test system, whereas high temperature measurements (300–500 K) were carried out using a Delta Design temperature chamber.

At many operating conditions, the voltage decay does not follow the ideal result of Fig. 17. This can result from excessive noise, very short lifetimes, non-ideal temperature effects, or inadequate compensation of parasitic capacitance. Furthermore, any non-idealities are amplified when the transient voltage decay is differentiated and inverted in order to obtain the carrier lifetime. To mitigate these issues, the measured decay curves were passed through a low-pass Butterworth filter, followed by a Gaussian filter. This was the most effective approach to remove noise without distorting the magnitude of the extracted lifetime.

Furthermore, due to the dynamic nature of the diode parasitics, the optimal compensation resistance varies for each set of measurement conditions (i.e., bias,

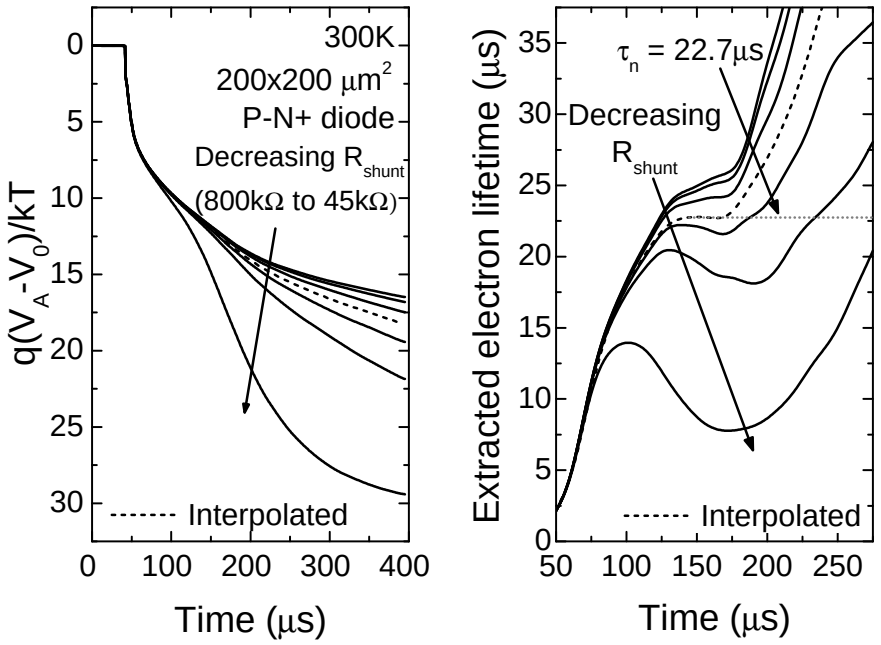

Figure 17: Example set of OCVD and corresponding lifetime extraction curves. Dashed lines represent the interpolated ideal curves used to determine the carrier lifetime.

temperature, device size and type). Without adequate compensation, the extracted lifetime can deviate significantly from the true carrier lifetime. To address this issue, the transient voltage decay was measured across a wide range of shunt resistances for each set of measurement conditions. The transient decay corresponding to the optimal compensation was then interpolated from these measured curves as shown in Fig. 17, increasing the accuracy and precision of the extracted carrier lifetime.

3.1.2 Measurement Results

In Figure 18, the injection dependence of the substrate minority electron lifetime is shown across a range of temperatures, demonstrating a decrease in lifetime with decreasing temperature, along with an increase in lifetime with increasing current density. The corresponding temperature dependence is given in Figure 19, in which the bias current is fixed to decouple the lifetime injection dependence. The consistency

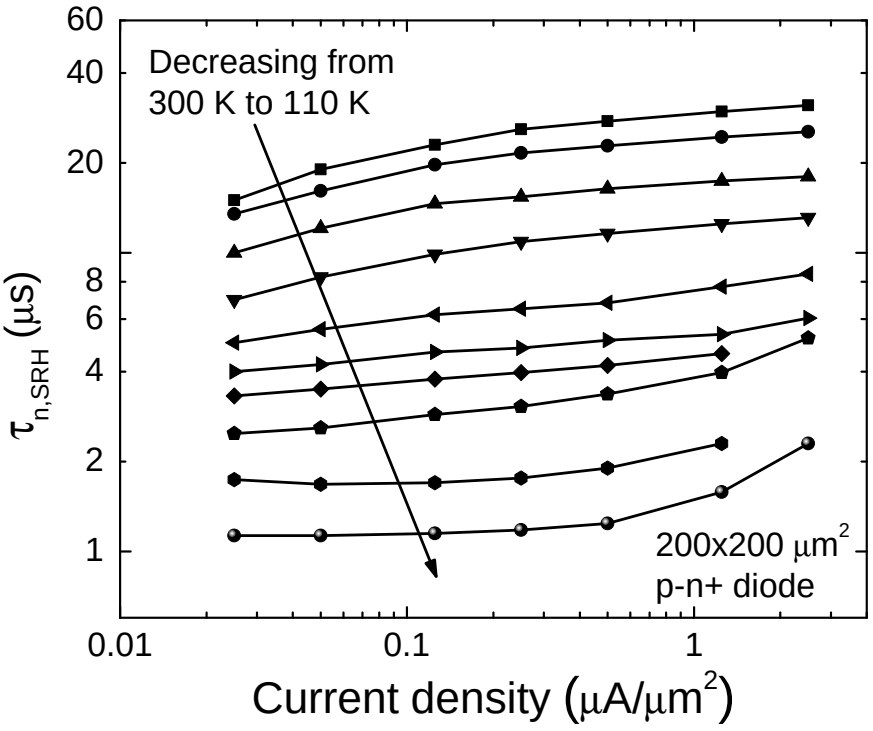

Figure 18: Minority electron and hole lifetimes as a function of diode current density.

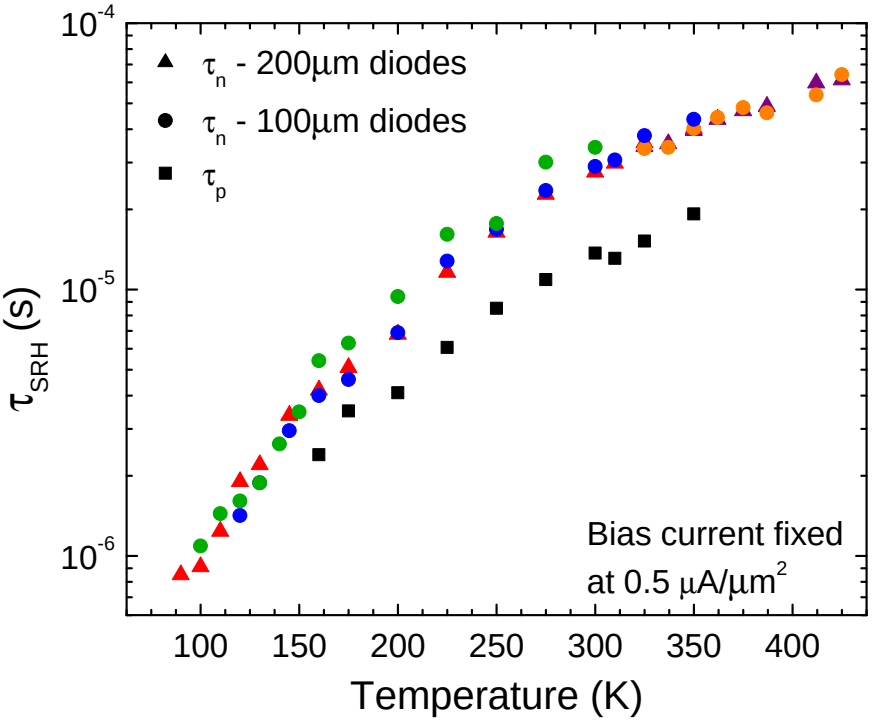

Figure 19: Minority electron and hole lifetimes across temperature for fixed current density of $0.5\,\mu A/\mu m^2$.

of the measured electron lifetime across multiple samples and test structure sizes supports the validity of my experimental approach. Also shown in Figure 19 is the n-well hole lifetime, which exhibits a similar temperature dependence to that of the electron lifetime. Lifetime extraction was limited to a temperature range from 90 K to 425 K because of excessive parasitics at the temperature extremes that caused the measured voltage decay curves to depart from the expected behavior. The fact that both the electron and hole lifetimes increase with increasing temperature indicates that their respective capture cross-sections have an inverse dependence on temperature. The lack of a rapid increase in lifetime up to a temperature of 425 K indicates that the dominant trap energy level is fairly deep or near the middle of the bandgap.

Proton irradiation experiments were conducted to assess the effects of displacement and ionization damage of the minority carrier lifetimes. The diode test structures were subjected at room temperature to 63.3 MeV proton irradiation up to a total accumulated dose of 1 Mrad(Si). The samples were irradiated at the Crocker Nuclear Laboratory at the University of California at Davis. A five-foil secondary emission monitor calibrated against a Faraday cup was used for dosimetry measurements. The radiation source (Ta scattering foils) was located several meters upstream of the target, and this established a beam spatial uniformity of about 15 % over a 2.0 cm radius circular area. Prior to irradiation, the diode I-V curve exhibits a nearly ideal slope, as expected for a properly fabricated diode. As shown in Figure 20, with increasing proton dose, the slope of I-V curve increasingly departs from the ideal slope of 60 mV/decade as a result of increased recombination. Temperature dependent lifetime measurements were carried out, with the results given for both electron and hole lifetimes in Figure 21. As expected intuitively and from the I-V characteristics, both hole and electron lifetimes decrease substantially because of the increased recombination associated with displacement and ionization damage. Furthermore, the slope of the temperature dependence flattens, indicating the presence of additional defect types.

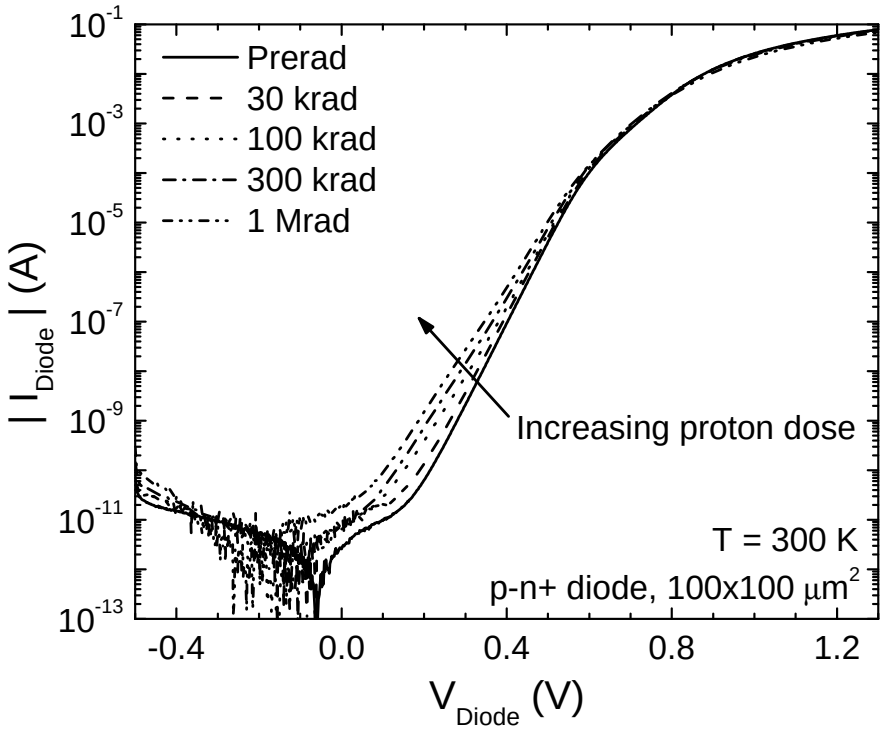

Figure 20: Radiation response of p-n+ diode I-V characteristic.

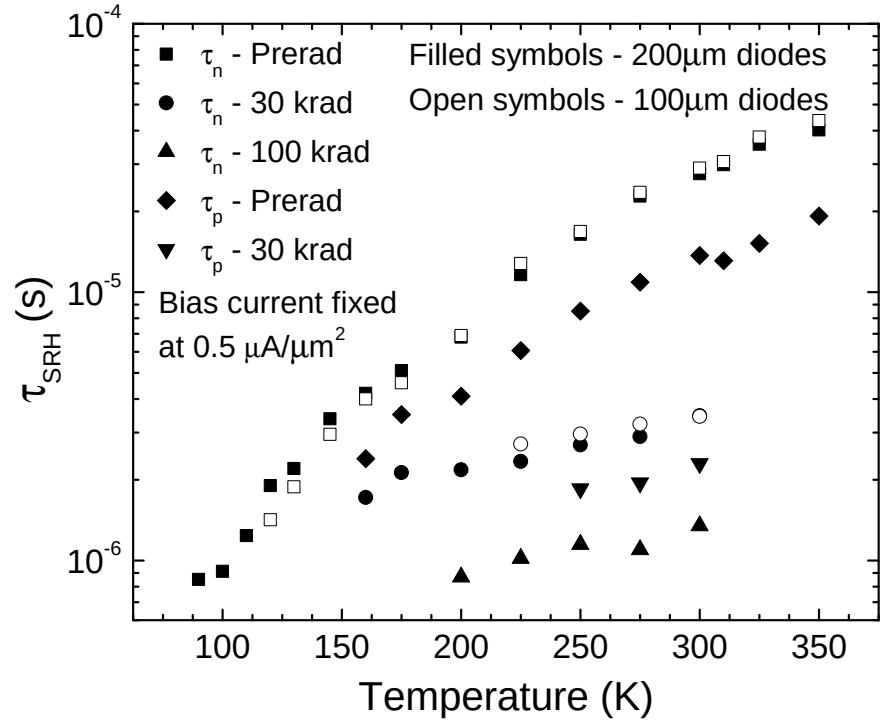

Figure 21: Radiation response of minority electron and hole lifetimes across temperature. Filled symbols correspond to 200 µm diodes at various levels of irradiation and open symbols correspond to 100 µm diodes.

3.1.3 Calibrated Trap Modeling

Based on the analysis of Rein [85], a convenient expression of the SRH recombination lifetime for a defect of energy level E_T is given by

$$\tau_{SRH} = \tau_{n0}\left[\frac{p_0 + p_1 + \Delta n}{p_0 + n_0 + \Delta n} + k\frac{n_0 + n_1 + \Delta n}{p_0 + n_0 + \Delta n}\right] \tag{4}$$

where n_0 and p_0 are the equilibrium densities of electrons and holes, n_1 and p_1 are the SRH densities, Δn is the excess carrier density, and τ_{n0} and τ_{p0} are the respective capture time constants of electrons and holes, defined as

$$\tau_{n0} \equiv (N_T \sigma_n v_{th})^{-1} \quad \text{and} \quad \tau_{p0} \equiv (N_T \sigma_p v_{th})^{-1} \tag{5}$$

with inverse dependencies on the thermal velocity, v_{th}, the defect concentration, N_T, and the capture cross-sections σ_n and σ_p [85]. In (4), a symmetry factor k has been defined that depends only on the defect structure rather than on the absolute quantities of N_T and $\sigma_{n,p}$,

$$k \equiv \frac{\sigma_n}{\sigma_p} = \frac{\tau_{p0}}{\tau_{n0}} \tag{6}$$

Having introduced the symmetry factor k, the absolute defect parameters N_T and σ_n only appear in the electron capture time constant τ_{n0}, which is a common factor of both terms of (4). Consequently, τ_{n0} acts solely as a scaling factor for τ_{SRH}, whereas the relative defect parameters E_T and k form the basis for the interrelated injection and temperature dependencies of the SRH lifetime.

Considering p-type material, the temperature dependent terms are n_1, p_1, and τ_{n0}. The majority carrier concentration p_0 is also temperature dependent due to carrier freeze-out. This is clearly critical at cryogenic temperatures, but must also be considered across all temperatures for doping levels near the Mott transition [91]. Assuming a trap center above mid-bandgap, the low-level injection SRH lifetime reduces to

$$\tau_{SRH}^{LLI,p} = \tau_{n0}(T)\left[1 + k\frac{n_1(T)}{p_0}\right] \tag{7}$$

There are two contributions to the overall temperature dependence of the SRH lifetime: τ_{n0}, which merely reflects the temperature dependencies of the capture cross-section σ_n and the thermal velocity v_{th}, and the SRH density, n_1, which increases exponentially with increasing temperature. The thermal velocity has a power-law dependence on temperature, whereas the SRH density derives its temperature dependence from the conduction band density of states N_C. The temperature dependence of σ_n typically follows a power law, but depends entirely on the nature of the trap in question (consider, for example, the E1 and E4 defects, which have opposite dependencies [44]).

At moderate temperatures, the contribution of n_1 to the overall temperature dependence can be neglected. Therefore, the temperature dependence of the SRH lifetime is given directly by $\tau_{n0}(T)$, and is proportional to the inverse product $(\sigma_n(T)v_{th}(T))^{-1}$. From this dependence, the superimposed dependence of the thermal velocity can be removed, revealing the capture cross-section temperature dependence. As the temperature increases, n_1/p_0 cannot be neglected and eventually begins to dominate (7), resulting in a steep increase in the SRH lifetime. The critical temperature for the onset of this steep increase is largely driven by the trap energy level. A shallow trap will manifest this increase at a much lower temperature than a deep level trap, due to its higher SRH density. Consequently, the trap energy level can be determined from either the onset temperature itself or the slope of the lifetime for temperatures above the onset temperature.

Since the minority electron lifetimes reported here do not exhibit a steep increase at high temperatures, the dominant trap in the substrate must be a deep-energy-level trap with an onset temperature above 425 K. The temperature and injection dependent lifetime data were simultaneously fit by using the SRH model and varying trap energy, trap density, $\sigma_n(T)$, and $\sigma_p(T)$. The resulting parameters are listed in Table 1, where the modeled trap is labeled as $\mathrm{T_n}$. With this data, it is important to note that only the combined magnitude of the $N_T \cdot \sigma_n$ product can be extracted, since the magnitude of

Table 1: Trap parameters for calibrated carrier lifetime models.

Trap	Energy [eV]	$\sigma_p(T)$ [cm^2]	k	N_T [cm^{-3}]
T_n	E_C - 0.5	$1\times10^{-15}\times(T/300)^{-3.49}$	3.5	4.5×10^{12}
T_p	E_V + 0.32	$3.1\times10^{-15}\times(T/300)^{-3.05}$	1	4.3×10^{12}

the lifetime is dependent on the product of these parameters. Additional information is needed to decouple these two parameters. However, since their relative magnitudes do not affect the shape of the lifetime injection and temperature dependence, the trap densities for T_n and subsequent defect models can be arbitrarily chosen. Figure 22 demonstrates the accuracy of the fit across the entire temperature range. A trap energy level of 0.5 eV below E_C was used for this calculation. It was found that when the trap energy level was within 0.4 eV of the conduction or valence band edges, the high-temperature fit began to diverge substantially from the data because of the onset of the steep lifetime increase.

The quality of the injection-dependent lifetime fit using the T_n model is shown in Fig. 23 for both 225 K and 300 K. The injection level is defined as the ratio of the injected minority carrier density to the majority carrier concentration (neglecting freeze-out, $\Delta n/N_A$). In order to estimate the injected carrier density, the diode voltage was measured immediately after open-circuiting the diode in order to remove excess voltage drop due to series resistance. Using this actual applied voltage and known doping concentration, the injected carrier concentration was determined from the following boundary condition: $\Delta n_p = n_i^2/N_A(\exp(qV_A/kT) - 1)$. For a current density of 1 µA/µm^2, the actual diode voltage is 0.665 V, corresponding to an injected carrier density of approximately 1.8×10^{16} cm^{-3}, which translates to an injection level of 20.13 ·

Similar to the electron lifetime, the dominant trap indicated by the minority hole lifetime in the n-well is a mid to deep level trap. The extracted trap parameters are given in Table 1, where the trap is labeled as T_p. In Figure 22, the resulting

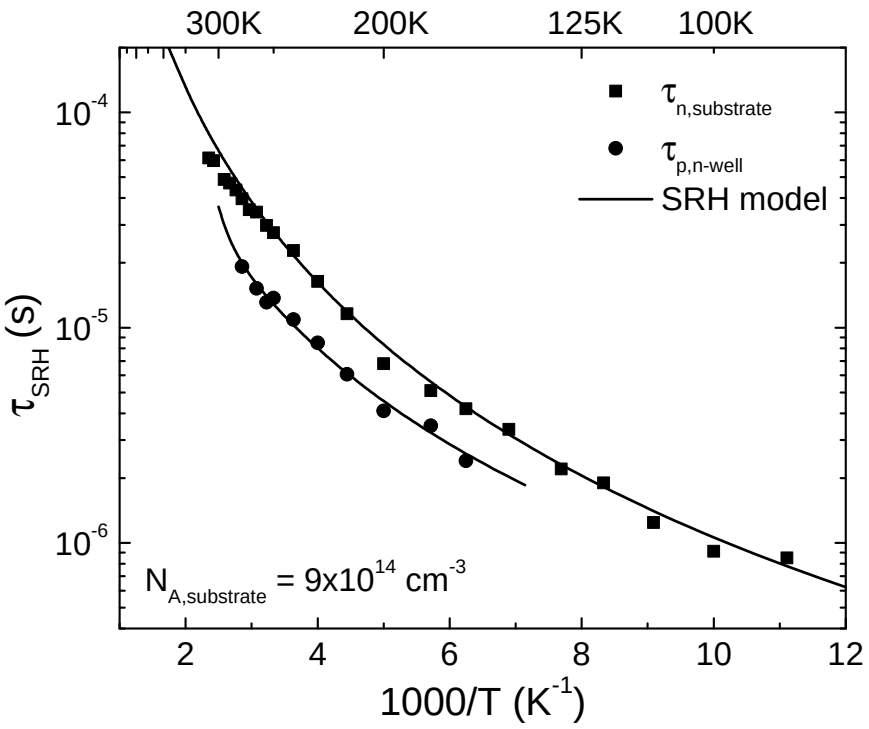

Figure 22: Comparison of simulated and experimental temperature dependence of minority electron and hole lifetimes.

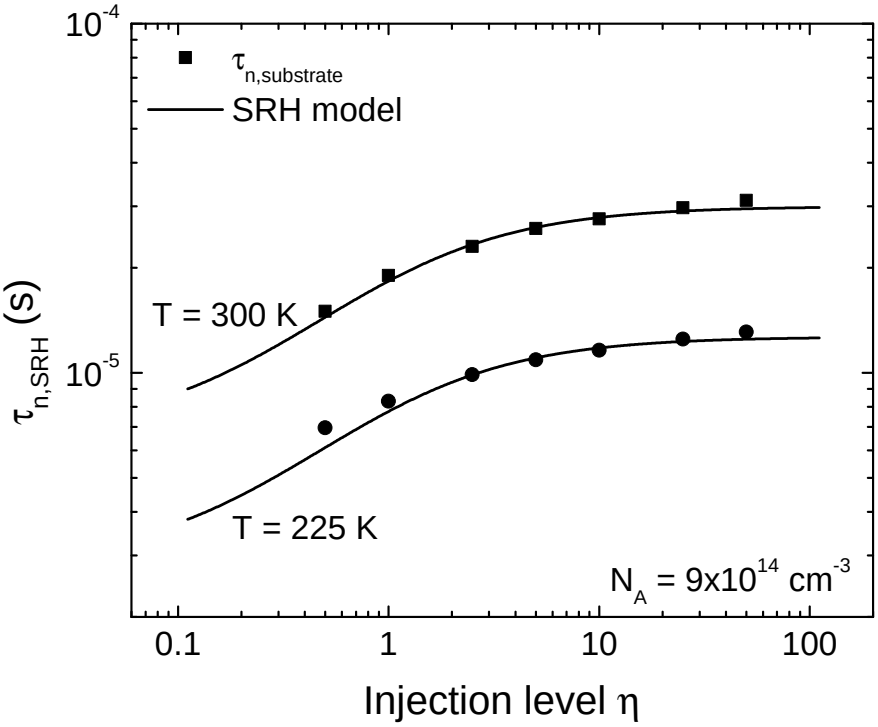

Figure 23: Comparison of simulated and experimental injection dependence of minority electron lifetimes.

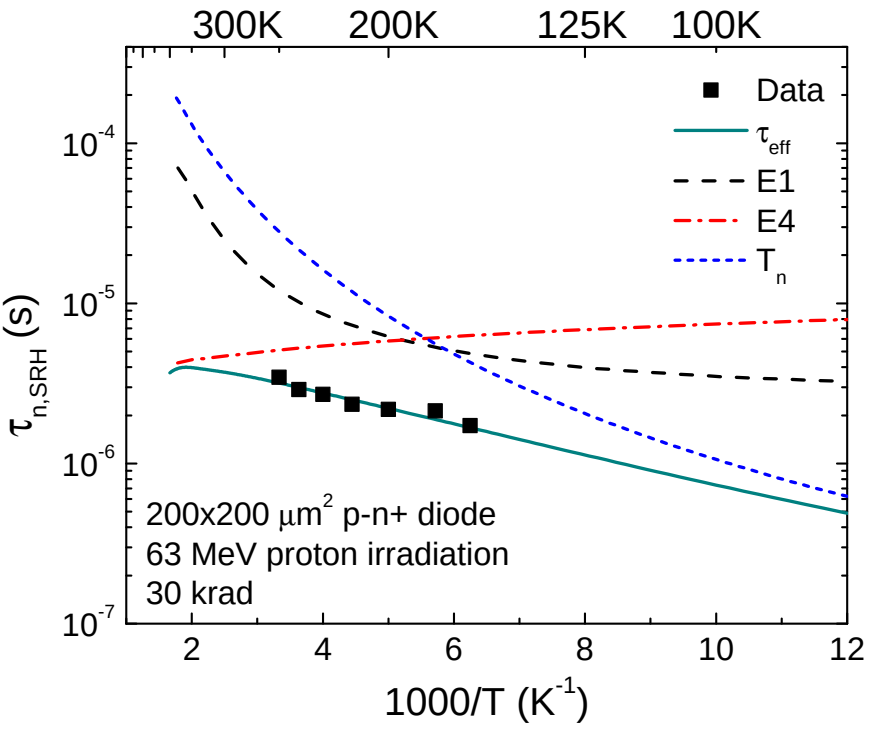

Figure 24: SRH model fit of minority electron lifetime vs. temperature after 30 krad proton irradiation. The effective lifetime is determined by the inverse sum of the reciprocal lifetimes of all contributing traps.

temperature-dependent model is plotted against the measured lifetimes. A trap energy level of 0.32 eV above E_V was used for this calculation.

The irradiated electron lifetime data can be modeled by introducing additional trap energy levels. As discussed in [44], the primary energy levels produced by electron irradiation in p-type silicon are the E1 (vacancy-oxygen complex) and E4 (divacancy) defects. These defects can be introduced as a starting point for post-irradiation lifetime calculations. By maintaining the same T_n trap density and increasing the trap densities of the E1 and E4 levels, it was possible to closely fit the measured temperature dependence of the 30 krad sample, shown in Figure 24. The resulting trap densities used in this model are $N_{T,E1} = 1 \times 10^{12}\,\mathrm{cm}^{-3}$ and $N_{T,E4} = 4.2 \times 10^{12}\,\mathrm{cm}^{-3}$.

3.2 Mobility and Incomplete Ionization

The Philips unified mobility model [45, 46] is a physics-based analytical model that unifies the descriptions of the majority and minority carrier mobilities. Besides lattice, donor, and acceptor scattering, this model also incorporates the effects of impurity screening by charge carriers, electron-hole scattering, clustering of impurities [45], and a full temperature dependence for both majority and minority carrier mobility [46]. Moreover, since the model gives the carrier mobility as an analytical function of the donor, acceptor, electron, and hole concentrations, it is a natural fit for implementation within a TCAD device simulator.

The strong temperature dependent nature of the lattice scattering mobility is explicitly shown in its definition,

$$\mu_{i,L} = \mu_{max} \left(\frac{300}{T} \right)^{\theta_i} \tag{8}$$

Similarly, the majority impurity scattering mobilities, $\mu_{e,D}$ and $\mu_{h,A}$, directly depend on temperature and are expressed as

$$\mu_{i,I}(N_I, c) = \mu_{i,N} \left(\frac{N_{ref,1}}{N_I} \right)^{\alpha_1} + \mu_{i,c} \left(\frac{c}{N_I} \right) \tag{9}$$

where (i, I) stands for (e, D) or (h, A). For the majority impurity scattering mobility at low temperatures, $\mu_{i,N}$ will dominate, since it has a direct power law dependence on temperature and $\mu_{i,c}$ has a inverse power law dependence. The minority impurity and electron-hole scattering mobilities derive their temperature dependence both from their direct dependence on the majority impurity expression and from the parameter P_i within their respective mobility ratio functions $G(P_i)$ and $F(P_i)$ [45]. Assessing which of these scattering components drives the overall temperature dependence of the mobility at extremely low temperatures is an important step in evaluating and calibrating an accurate mobility model for use down to cryogenic temperatures. From Fig. 25, it is clear that the lattice scattering mobility dominates the temperature dependence of

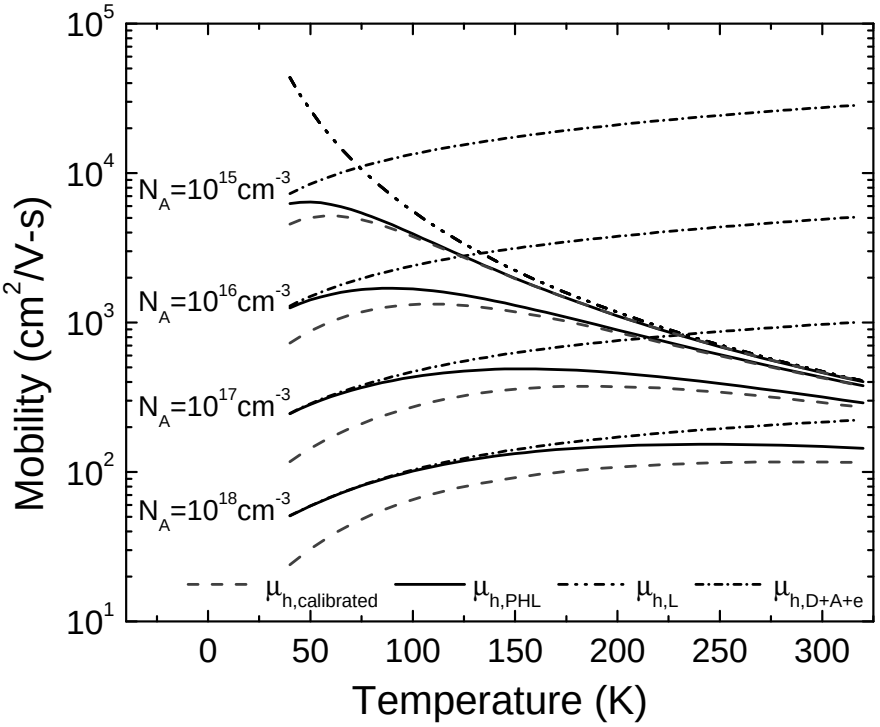

Figure 25: Comparison of the temperature dependence of scattering mechanisms used in Philips unified mobility model for various doping concentrations. The effective hole mobility from the calibrated model is shown by the dashed lines.

the carrier mobility at lower doping concentrations and higher temperatures, whereas the combined majority/minority impurity and carrier scattering mobility increasingly dominates the temperature dependence for higher doping concentrations and lower temperatures. This provides a reasonable starting point for evaluating the mobility model against experimental resistivity measurements across temperature and doping concentration.

Similarly, an accurate model for the incomplete ionization of dopants is necessary not only to meaningfully link experimental resistivity data to theoretical mobility values, but is in its own right a critical component of accurate low-temperature device models. Recently in two companion papers by Altermatt *et al.* [3, 4] a parameterization of the density-of-states near the band edge of doped silicon was derived and subsequently applied to calculate dopant ionization level. In the model

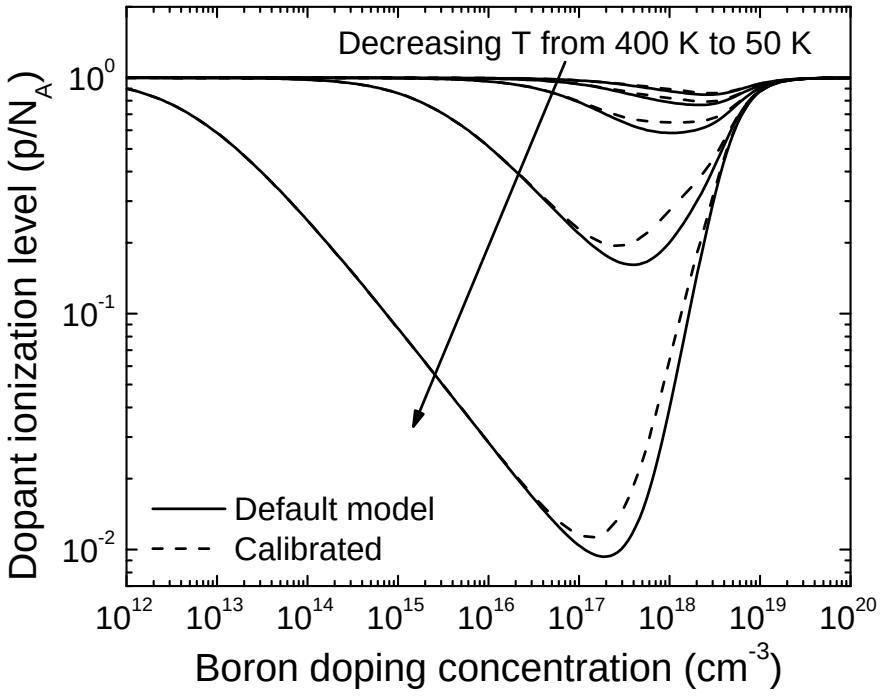

Figure 26: Ionization level as a function of boron doping concentration across a wide temperature range. These curves were calculated using the modified ionization model presented in [4, 91]. The dashed lines reflect the ionization level from the calibrated ionization model.

derivation, the dopant band was shown to only touch the conduction band at the Mott transition and to merge with the conduction band at considerably higher doping levels, agreeing with the experimental data that at these high doping levels the dopants are completely ionized. Marked occupation of dopant states occurs when the Fermi level is located near the dopant level, leading to incomplete ionization of dopant atoms and a diminished free carrier density. Up to 25 % of dopant atoms may be non-ionized for certain doping concentrations [91]. Consequently, incomplete ionization at moderate temperatures is an important concern for doping levels from roughly $1 \times 10^{17}\,\mathrm{cm}^{-3}$ to $1 \times 10^{19}\,\mathrm{cm}^{-3}$, clearly shown in Figure 26.

3.2.1 Resistance Measurements

All resistance measurements over temperature were conducted within IBM's 5HP/5AM SiGe BiCMOS technology, using an Agilent 4156 Semiconductor Parameter Analyzer to

perform Kelvin measurements on a variety of test structures. Temperature-dependent measurements of packaged test structures were carried out with a closed-cycle liquid-helium cryogenic test system capable of DC to 100 MHz operation from 4 K to 400 K. The measurements of p-type resistance reported here include the resistivity of the lightly-doped substrate (constant doping of around $1 \times 10^{15}\,\mathrm{cm}^{-3}$), the intrinsic-base sheet resistance of the SiGe HBT (peak doping near $1 \times 10^{18}\,\mathrm{cm}^{-3}$), and the sheet resistance of the p+ diffusion layer (extrinsic base—peak doping greater than $1 \times 10^{20}\,\mathrm{cm}^{-3}$). Measurements of n-type resistance include the sheet resistances of the lightly-doped epilayer in which the SiGe HBT intrinsic collector is defined (with self-aligned collector implant in place—constant doping near $5 \times 10^{15}\,\mathrm{cm}^{-3}$), the collector of the high-breakdown variant of the SiGe HBT (with a different collector implant—peak doping near $5 \times 10^{16}\,\mathrm{cm}^{-3}$), and the heavily-doped SiGe HBT subcollector (peak doping near $1 \times 10^{20}\,\mathrm{cm}^{-3}$).

The substrate resistivity was measured using the standard four-point-probe technique. A custom test structure was designed with four collinear 1.6 µm substrate contacts that were equally spaced by 150 µm. The base sheet resistance was measured using a ring-dot structure with the high-fT collector doping profile; this structure consists of an emitter ring bounded by two inner and two outer base contacts. Since the data was used calibrate mobility and incomplete ionization models that are generalized models for the Si material system, it was necessary to avoid any Ge related effects. Consequently, the particular structure that was measured did not include Ge grading in the base; with this exception, the test structure was fabricated according to the standard SiGe BiCMOS process. The remaining sheet resistances were measured using rectangular Kelvin structures of varying geometry.

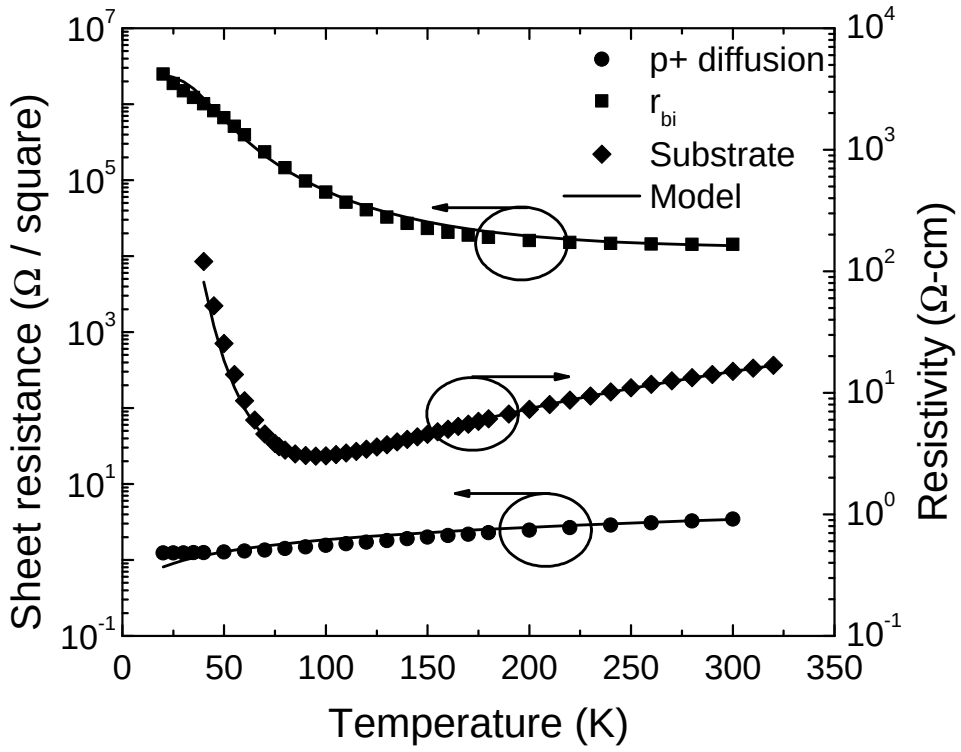

Figure 27: P-type resistivity and sheet resistance measurements in IBM's 5HP/5AM SiGe BiCMOS technology: (a) p-type, (b) n-type

3.2.1.1 P-type data

In Figure 27, the p-type temperature-dependent resistance data are shown, including the substrate resistivity, intrinsic-base sheet resistance, and p+ diffusion sheet resistance. For the lightly-doped substrate, a significant increase in resistivity is seen as the temperature decreases below 100 K. This can be attributed to the significant degree of incomplete ionization that is expected for a boron density of $9 \times 10^{14}\,\mathrm{cm}^{-3}$. The decrease in substrate resistivity from room temperature down to 100 K can be attributed solely to the expected increase in mobility, since the dopants are completely ionized in this temperature range. In contrast, the intrinsic-base sheet resistance exhibits clear signs of incomplete ionization even at temperatures above 200 K, as a result of its higher doping concentration. Reexamining Figure 26 for doping levels near $1 \times 10^{18}\,\mathrm{cm}^{-3}$, incomplete ionization is already in effect at room temperature and the

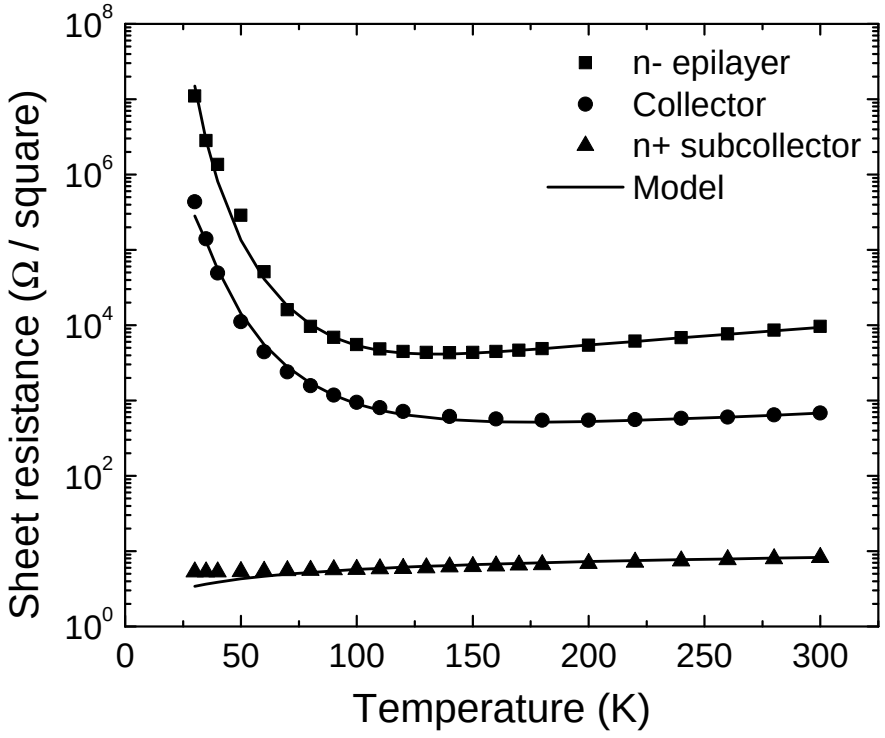

Figure 28: N-type resistivity and sheet resistance measurements in IBM's 5HP/5AM SiGe BiCMOS technology

level of ionization steadily decreases with decreasing temperature, albeit at a slower rate than for lower doping concentrations. This overrides the mobility-related decrease in resistance and leads to an increase in r_{bi} with decreasing temperature, accelerating as the temperature decreases below 200 K. Finally, the p+ diffusion sheet resistance exhibits very little temperature dependence because of its extremely high doping concentration. Complete ionization of dopants holds across the entire temperature range, and thus the slight decrease (less than 3x) in sheet resistance from 300 K to 20 K can be attributed to a corresponding increase in mobility with cooling.

3.2.1.2 N-type data

Figure 28 shows the n-type temperature dependent resistance data, including the sheet resistances of the n– epilayer, HBT collector, and n+ HBT subcollector. The n– epilayer, which has a relatively higher doping concentration than that of the

substrate, displays a similar temperature dependence. The weaker dependence at higher temperatures reflects the fact that the mobility dependence is also weaker because of the higher doping concentration. The collector layer is merely the n– epilayer after ion implantation. Thus, the collector sheet resistance demonstrates a similar overall temperature dependence compared to the n– epilayer, with several key differences: the overall magnitude is lowered because of a higher carrier density, the moderate temperature region is suppressed, reflecting a reduction in mobility because of increased impurity scattering, and the onset of incomplete ionization occurs at a higher temperature because of the higher doping concentration. Finally, the n+ subcollector sheet resistance data mirrors that of the the p+ diffusion layer, indicating complete ionization across all temperatures.

3.2.1.3 Substrate resistivity radiation response

Proton irradiation experiments were conducted in order to assess the effects of displacement and ionization damage on the substrate resistivity. Resistivity test structures were subjected at room temperature to 63.3 MeV proton irradiation up to a total accumulated dose of 1 Mrad (Si). Fig. 29 shows the changes induced in the temperature dependent resistivity at accumulated doses of 100 krad, 300 krad, and 1 Mrad. By normalizing the irradiated resistivities to the pre-radiation data, the specific nature of the radiation-induced changes to resistivity is more easily seen. At moderate temperatures, the increase in resistivity indicates that radiation-induced displacement damage results in higher lattice scattering. Below 100 K, however, where impurity scattering dominates at this particular doping concentration, the resistivity decreases. This decrease could be caused by boron dopant deactivation from radiation damage. Dopant deactivation would cause a temperature-independent increase in resistivity due to a lower carrier concentration, along with decreased ionized impurity scattering that would be manifested as decreased resistance at low temperatures. Moreover,

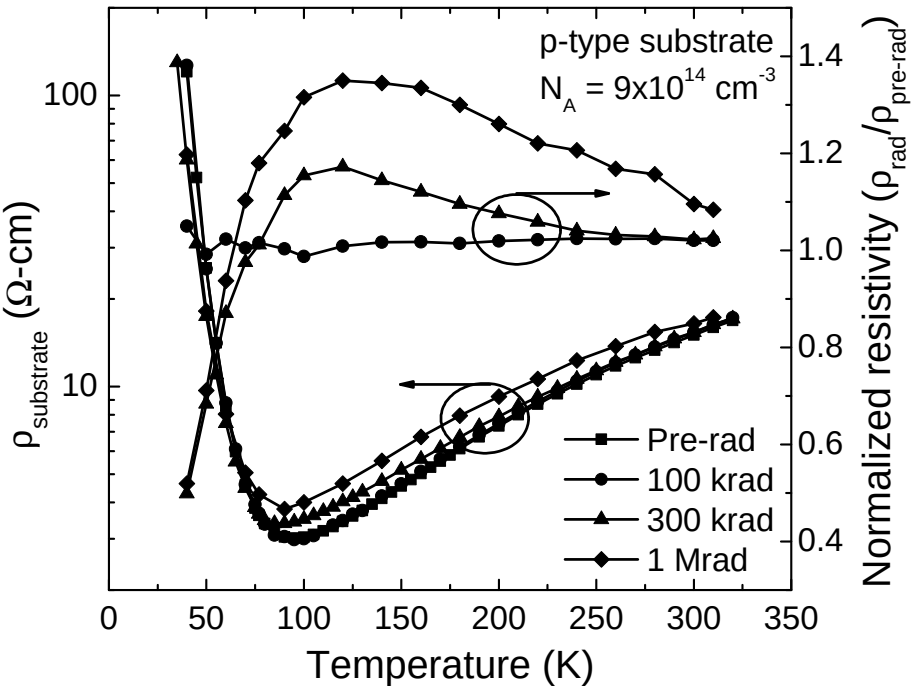

Figure 29: Proton radiation response of the substrate resistivity across temperature.

the lower carrier concentration would lead to a relatively lower degree of incomplete ionization, resulting in lower resistivity in the deep cryogenic temperature regime.

3.2.2 Calibrated Modeling

To develop mobility and ionization models that together produce accurate models of resistivity for the doping-dependent p-type and n-type resistance data presented here, the most reliable approach is to retain models that are physics-based and focused on material systems rather than particular technologies. This approach minimizes the reliance on assumptions that could potentially break down under conditions for which the models have not been experimentally tested. Models that are purely empirical or have been developed specifically for a particular technology often do not extend well to other technologies or physical conditions. Consequently, the approach in this paper has been to carefully calibrate the parameters of the Philips mobility model and the Altermatt ionization model, since both of these models were developed out of fundamental theory and aimed for silicon-based systems in general. All of the

Table 2: Calibrated parameters used in mobility model for arsenic-, phosphorus-, and boron-doped silicon.

Parameter	As	P	B
$N_{ref,1}$ (cm^{-3})	1.45×10^{17}	1.1×10^{17}	1.5×10^{17}
$N_{ref,I}$ (cm^{-3})	1×10^{22}	1×10^{22}	1×10^{22}
α_1	0.85	0.65	0.8
θ_i	1.72	1.72	1.82

Table 3: Calibrated parameters used in incomplete ionization model for arsenic-, phosphorus-, and boron-doped silicon.

Parameter	As	P	B
N_{ref} (cm^{-3})	3×10^{18}	7×10^{17}	8.5×10^{17}
c	1.5	0.8	1.4
N_b (cm^{-3})	9×10^{18}	6×10^{18}	4.5×10^{18}
d	1.8	1.3	2.4

experimental data were fit using a single set of model parameters. The modified parameters used for the calibrated Philips model and the calibrated Altermatt model are given in Tables 2 and 3.

Figure 27 shows the calibrated model fits for the p-type resistance data. The substrate resistivity was used to calibrate the lattice-scattering coefficient θ_i, since at moderate temperatures lattice scattering dominates and incomplete ionization is negligible. Properly accounting for the low-temperature increase in resistivity resulting from carrier freeze-out, combined with the increasing influence of impurity scattering, required simultaneous tuning of the impurity-scattering temperature coefficient α_1 and the ionization model parameters. The resulting model produces close fits to the three data sets across the entire temperature range.

Calculation of the intrinsic-base sheet resistance also required the highly-variable doping concentration across the base to be properly taken into account. Substituting an effective base doping is insufficient, since mobility and ionization level are strong functions of doping concentration. For that reason, the variable base profile was

discretized into very thin layers of constant doping using data from SIMS measurements. The resistivity and corresponding sheet resistance for each layer were calculated, then all of the individual sheet resistances were added together as parallel resistances to determine the total effective base sheet resistance.

Following the same procedure, the n-type resistance data were used to calibrate the n-type mobility and incomplete ionization parameters. The resulting model fits are given in Figure 28. For the calculations of n– epilayer and collector sheet resistance, a constant effective doping level was used. For the n+ subcollector, however, the known doping profile was discretized in the same manner as the base profile to account for the varying doping concentration through the subcollector. Since both the modeled p+ diffusion and n+ subcollector sheet resistances begin to diverge from the data below 70 K, the error can be attributed to inadequate modeling of ultra-high doping effects in the mobility model.

3.3 Applications

Due to their accuracy at cryogenic temperatures, the parameter models presented here are especially well-suited for TCAD simulations of heavy-ion strikes. One future application is to address the increased SEU sensitivity of SiGe digital bipolar logic circuits at cryogenic temperatures. Cryogenic TCAD simulations of a heavy-ion strike on a single HBT can be used to investigate device mechanisms that drive SEU sensitivity. In addition, "mixed-mode" simulations can be carried out by inserting a TCAD HBT model into the compact model of a SiGe digital circuit. The mixed-mode simulation a heavy-ion strike at cryogenic temperatures and its resulting circuit-level effects would represent the most complete simulation to date of SEE as it actually occurs within space-based electronics.

CHAPTER IV

SINGLE-EVENT TRANSIENTS IN 45-NM RF-CMOS ON SOI

As SOI CMOS technology is scaled to more advanced nodes below 90 nm, its radiation response must be carefully characterized, particularly as advances are made in device materials and structure. The total-ionizing-dose (TID) tolerance of SOI CMOS has previously been reported for several advanced technology nodes [5, 58, 59]. In [59], the shallow-trench isolation (STI) oxide was shown to be the driving factor of the TID sensitivity of partially-depleted 65 nm RF-CMOS transistors. Body-contacting schemes that eliminate oxide sidewalls were demonstrated to be a potential solution for this sensitivity, since they remove the STI leakage path between the source and drain wells. However, the use of body contacts for improved TID tolerance imposes a significant reduction in RF performance as a result of increased parasitics, and thus establishes an important trade-off between performance and reliability.

In addition to TID tolerance, the sensitivity of advanced CMOS nodes to single-event effects (SEE) must be evaluated. In this chapter I investigate the SEE sensitivity of an advanced 45 nm SOI CMOS technology by capturing heavy-ion microbeam and pulsed-laser-induced current transients at the device level, and for the first time address the SEE impact of various body contacting schemes. Although SOI CMOS is generally less sensitive to SEE than bulk CMOS because of its smaller sensitive volume, it still suffers from effects such as parasitic bipolar amplification stemming from floating-body effects [32, 93]. Body ties are commonly used to lessen floating-body effects and increase SEE hardness [30, 40, 93]. The observed dependencies of single-event transients on bias conditions and strike location provide insight into

additional trade-offs that must necessarily be considered in conjunction with RF performance and TID tolerance in operating SOI CMOS in a space environment, particularly with respect to the optimal choice of the various body-contacting schemes that are available.

4.1 Experimental Details

Strain-engineered, partially-depleted 45 nm RF-CMOS on SOI devices are investigated here for SEE sensitivity. This high-performance technology features a 1.16 nm gate oxide, thin-film SOI, and employs the following FET-specific performance elements to overcome the inherent penalties associated with aggressive gate pitch scaling: enhanced dual stress liner (DSL), advanced embedded SiGe (eSiGe) integration, optimized stress memorization (SMT) processing, and advanced activation annealing to enhance nFET gain [74]. In addition to the standard T-body body-contacting scheme, FETs are available with a layout optimized for high-frequency RF applications, and are referred to as the "notched-body" contacts. Floating body nMOS and pMOS transistors within this technology can achieve peak f_T's of 485 GHz and 345 GHz, respectively [52].

The layout differences between the standard T-body and notched-body devices are illustrated in Figure 30. The key alteration in the notched-body layout is the reduction in the area of the heavily-doped body contact. This limits the body junction area and confines it to the region just below the polysilicon gate. For RF devices, this greatly reduces parasitic capacitance and gate leakage current, thus improving RF performance. For a longer than minimum channel length device with an optimized body-contact scheme, a peak f_T of 245 GHz was reported with no degradation in critical RF figures-of-merit [52].

Laser-induced transients were measured at the Naval Research Laboratory in Washington, DC, using a two-photon absorption (TPA) backside pulsed laser system capable of supplying a 1.2 µm diameter charge distribution profile [64]. This system

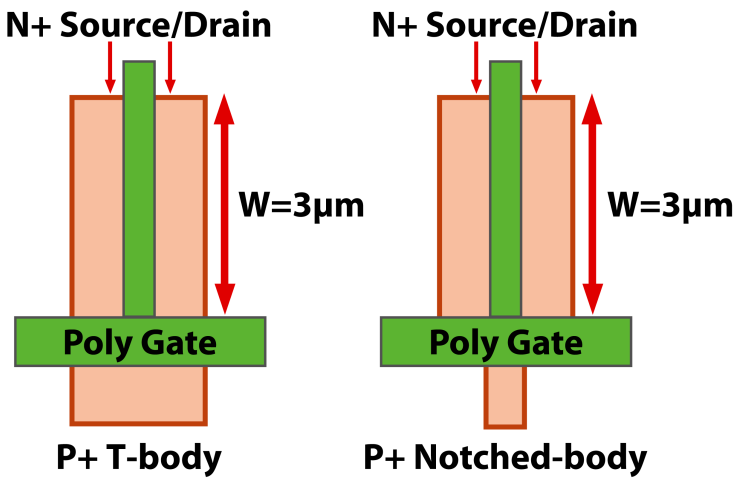

Figure 30: Layout geometries for T-body and notched-body contacting schemes (not to scale).

was employed because it enables 3-D position-dependent time-resolved measurements of single-event transients. In this system, device-level current transients are induced by injecting carriers using TPA from a sub-bandgap pulsed laser and are then recorded using high-bandwidth measurement equipment, including a Tektronix TDS71254 50 GHz real-time oscilloscope. The system is configured to produce optical pulses at 800 nm, a repetition rate of 1 kHz, and a pulsewidth of approximately 120 fs. The xyz translation platform has a position resolution of 0.1 µm; all data were collected in a rectangular xy grid at a fixed "z" with a step size of 0.25 µm. Upon inserting each DUT, the "z" position was optimized to place the sensitive volume at the peak focus of the laser beam. A similar test setup is described in [79]. Ion-strike-induced transients were measured at Sandia National Laboratories using a focused-ion microbeam with 36 MeV oxygen ions to provide a basis of comparison for the TPA results. A similar test setup is described in [80].

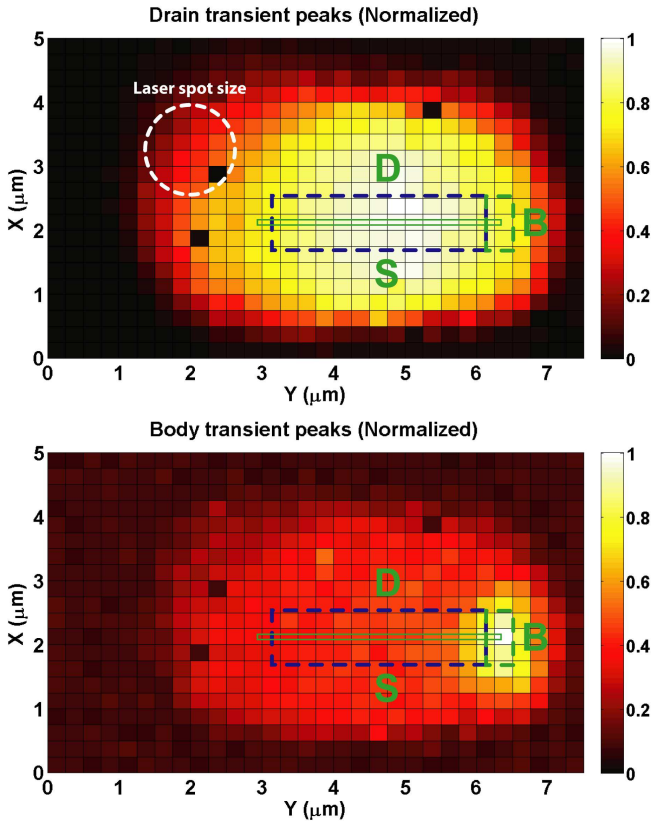

Figure 31: Normalized peak drain and body currents for laser-induced transients as a function of laser position for a 3.0 µm x 0.056 µm SOI T-body nFET with an incident-laser pulse energy of 10.7 nJ. The device is biased at V_{DS} = 1.0 V and V_{GS} = 0 V.

4.2 Experimental Testing of Different Body-Tie Geometries

4.2.1 Preliminary Results

T-body and notched-body nFETs with fixed dimensions (3.0 µm x 0.056 µm) were chosen to allow a quantitative comparison of the SET response between devices. A range of incident laser pulse energies ranging from 1.0 nJ to 10.7 nJ were used. In Figures 31-32, the peak currents of the drain and body transients from T-body and notched-body nFETs with a laser pulse energy of 10.7 nJ are plotted as a function of incident-laser position, with the device geometry and approximate laser spot size marked for reference. Due to the fact that the notched-body layout has a reduced silicon volume compared to that of the T-body layout, the transient peak magnitudes

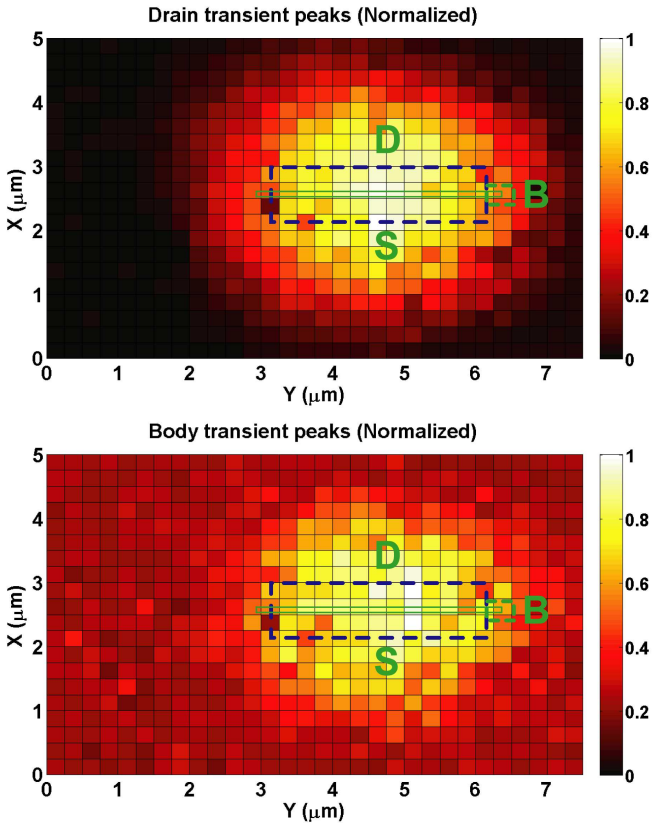

Figure 32: Normalized peak drain and body currents for laser-induced transients as a function of laser position for a 3.0 µm x 0.056 µm SOI notched-body nFET with an incident-laser pulse energy of 10.7 nJ. The device is biased at V_{DS} = 1.0 V and V_{GS} = 0 V.

are normalized to the overall peak magnitude for each channel of each device to provide a better basis for comparison. The difference in silicon volume comes entirely from the difference in the body tie area; the devices are otherwise identical in geometry and doping. For a 3 µm width device, the silicon volume of the T-body device is 13 % larger than that of the notched-body device.

The body transient peak of the T-body device in Figure 31 highlights the location of the body contact, as the maximum body current occurs when striking near the body contact. The body transient magnitude reduces as the laser spot moves away from the body contact; for these positions, the laser spot does not overlap the body-tie and no charge is deposited directly at the doped body-tie region. In addition, the sensitive area appears larger than the actual device geometry; this is due to the finite spot

size of the laser pulse coupled with the very large pulse irradiance used for this set of experiments. The measured sensitive area can thus be interpreted as the convolution of the laser spot size with the device geometry. The apparent size of the sensitive area for these data sets is a consequence of the overlap of the tails of the Gaussian laser pulse profile with the sensitive region of the device. Figure 32 illustrates the peak drain and body currents for laser-induced transients from a notched-body nFET. Two differences emerge when the notched-body results are compared to the T-body transient peaks: (1) in the notched-body device, no region exists with a prominent body transient similar to the T-body results, and (2) the T-body device exhibits a larger sensitive area even when normalized to account for increased charge deposition in the larger body-tie region. This indicates that for these body-contacting schemes there is a fundamental difference in the underlying physics of their transient responses. Since the data are normalized for each terminal of each device, the notched-body device appears to have a large body transient centered along the gate width; however, the normalized data simply highlights the largest transients relative to the rest of the 2-D scan area, which in this case are much smaller and closer to the background noise than the T-body transients.

The differences in the transient data are further illustrated in Figure 33, which plots the peak transient currents as a function of laser position along the width of the two devices. The notched-body device exhibits uniformly smaller source and drain transients compared to the T-body device (20 % less at its peak), with a reduced peak body transient across the width of the device, particularly at the location of the body-tie, where it is 3x smaller. It is expected that the body transient of the T-body device would in part be larger because of the larger active area at the body-tie end of the device; however, the two devices are identical in geometry and doping levels away from the body contact, and the fact that the transient magnitudes are reduced across the entire width of the device precludes the possibility that the laser spot size has any

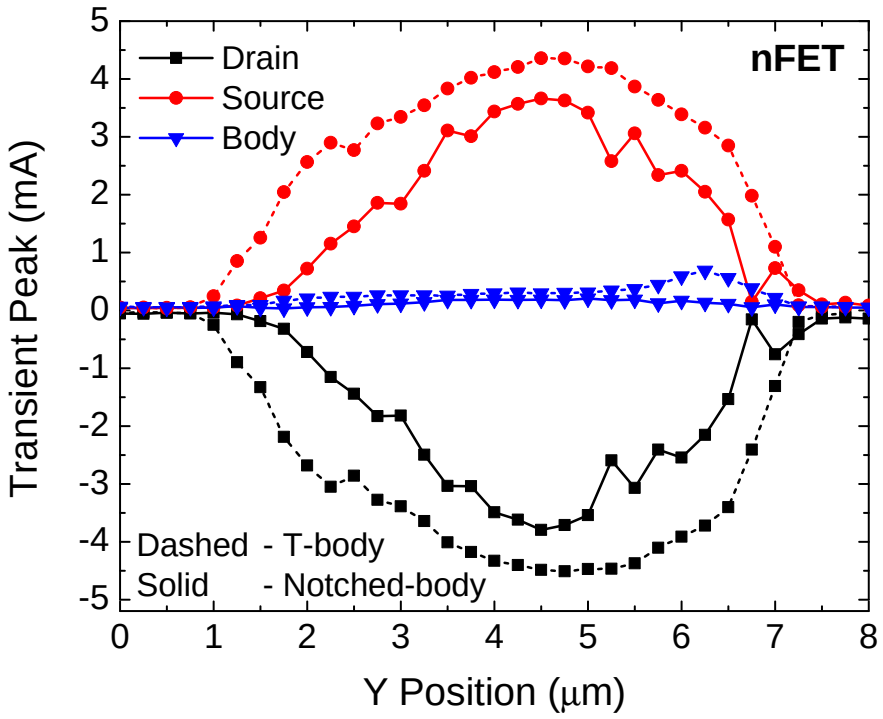

Figure 33: Peak laser-induced transient currents as a function of laser position along the width of 3.0 µm x 0.056 µm T-body and notched-body SOI nFETs with an incident-laser pulse energy of 10.7 nJ. The devices are biased at $|V_{DS}| = 1.0$ V and $V_{GS} = 0$ V.

significant impact on the key observations and conclusions of this work. Rather, the differences indicate an increased sensitivity of the T-body device that should also be exhibited in heavy-ion-induced transients.

In Figure 34, the peak transients are plotted as a function of laser position along the width of T-body and notched-body pFETs of the same dimensions (3.0 µm x 0.056 µm). Similar trends to the nFET data are observed, with the T-body device exhibiting much stronger body transients for laser positions near the body contact. The T-body pFET exhibits only marginally larger drain-transient peaks compared to the notched-body pFET, but additional measurements need to be collected to confirm whether this is an inherent difference between body-contacted nFETs and pFETs.

Figures 35-36 shows the time-resolved transients at the same position for each of the body-contacted nFETs, at the channel center near the body-tie end of the

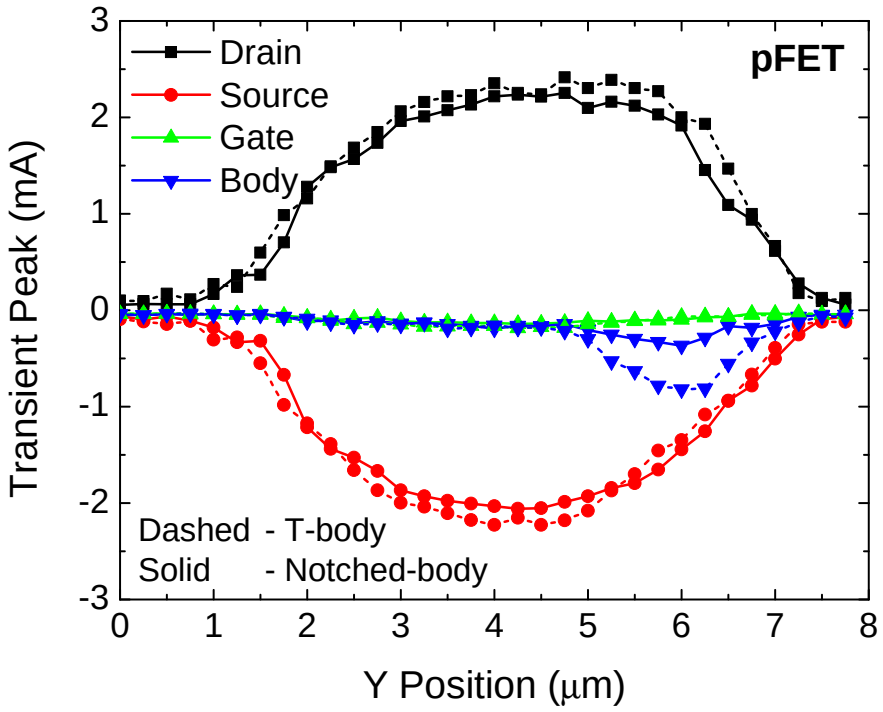

Figure 34: Peak laser-induced transient currents as a function of laser position along the width of 3.0 µm x 0.056 µm T-body and notched-body SOI pFETs with an incident-laser pulse energy of 10.7 nJ. The devices are biased at $|V_{DS}| = 1.0$ V and $V_{GS} = 0$ V.

device. The notched-body nFET exhibits smaller transients (2x reduction in peak drain current), demonstrating the increased sensitivity of the T-body nFET to strikes near the heavily-doped body contact. The larger T-body transients are in part due to the fact that at this position, the laser spot is generating carriers within the larger volume of the its body contact, leading to higher-magnitude transients and increased total charge collection. However, the degree of difference between the two devices is much greater than the increase in volume. One possible cause of the increased sensitivity is that the heavily-doped well used for the body contact has a much greater cross-sectional area, since the T-body layout has a larger area adjoining to the source and drain wells. Not only does this impact the RF performance of the device through increased parasitic capacitance, but it also introduces a *pn* junction between the drain and doped body tie that is reverse-biased when the device is biased at a worst case of

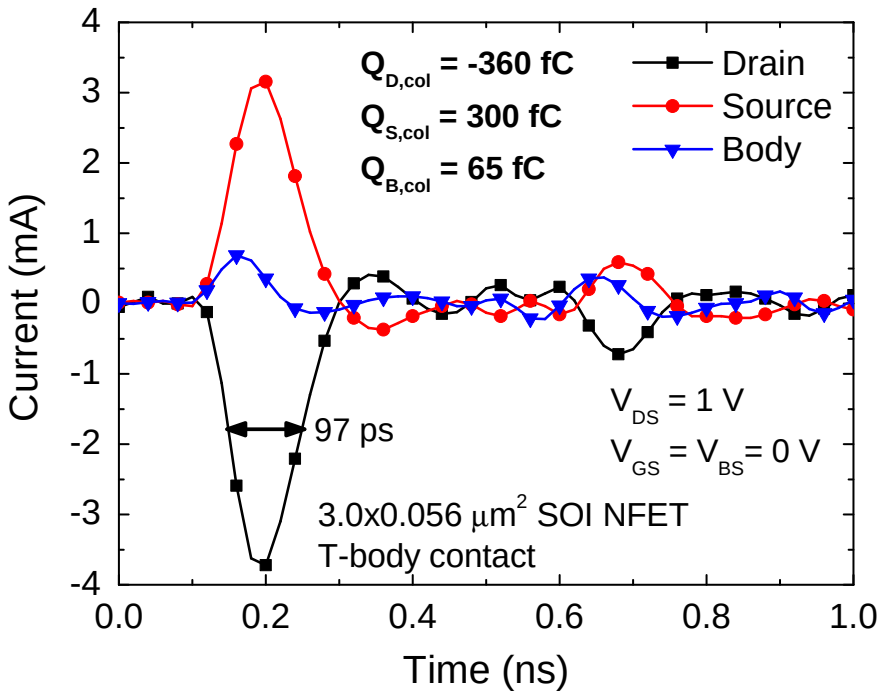

Figure 35: Transient currents for incident laser near body contact of a 3.0 µm x 0.056 µm T-body nFET with a laser pulse energy of 10.7 nJ. The device is biased at $V_{DS} = 1.0\,V$ and $V_{GS} = 0\,V$.

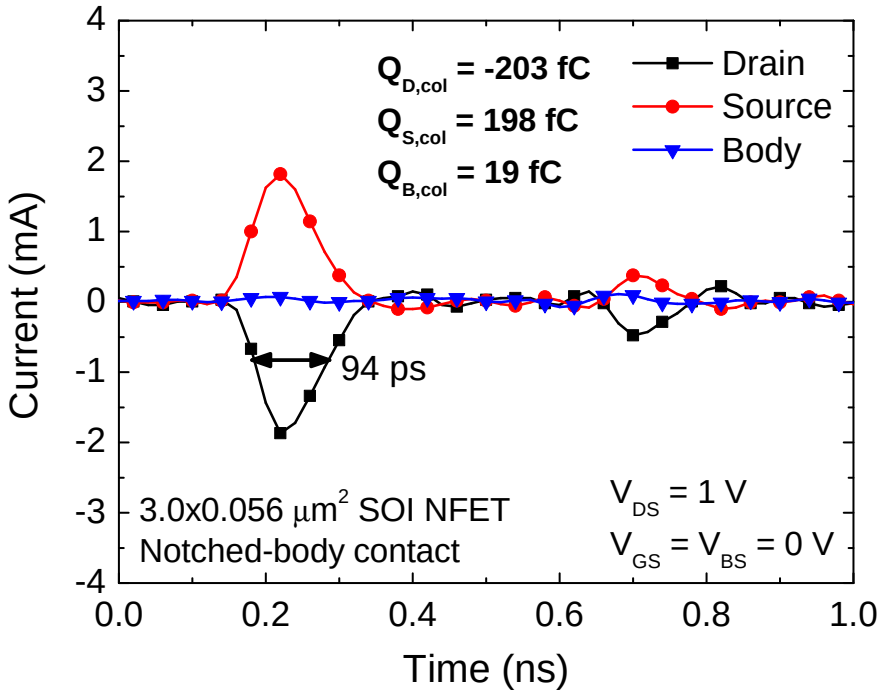

Figure 36: Transient currents for incident laser near body contact of a 3.0 µm x 0.056 µm notched-body nFET with a laser pulse energy of 10.7 nJ. The device is biased at $V_{DS} = 1.0\,V$ and $V_{GS} = 0\,V$.

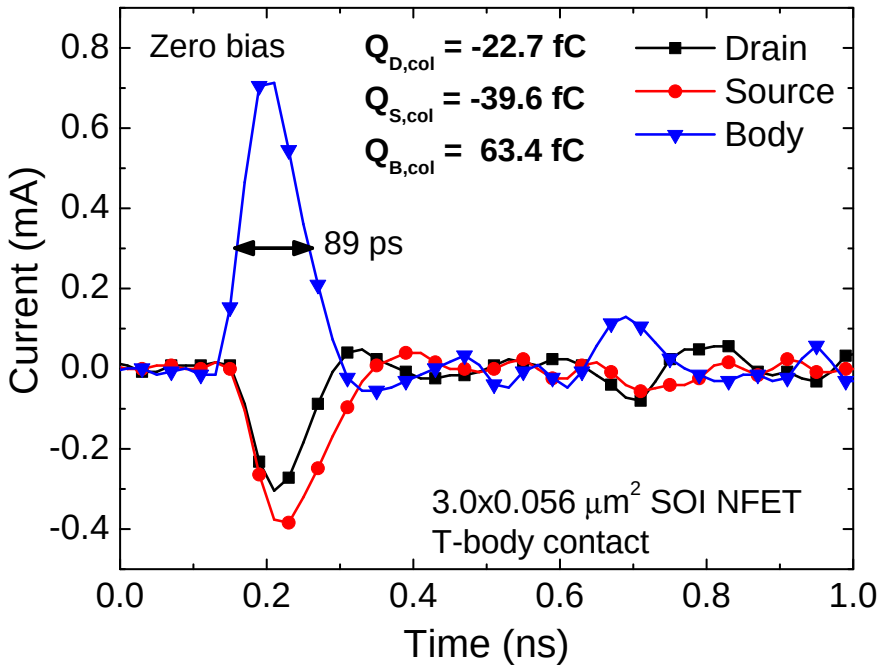

Figure 37: Transient currents at zero bias for incident laser pulse near body contact of a 3.0 µm x 0.056 µm T-body nFET with a laser pulse energy of 10.7 nJ.

V_{DS} = 1.0 V. When charge carriers are introduced into a T-body nFET as a result of a laser or heavy-ion strike, this reverse-biased junction will sweep excess holes into the body terminal and excess electrons into the drain terminal. Excess electrons will be swept into the drain by both the reverse-biased drain-body junction and the electric field across the channel that is caused by the applied V_{DS}, whereas excess holes will be collected by both the source and body terminals. In a notched-body device, the doped body-tie is confined to the area below the polysilicon gate, limiting the degree to which the heavily-doped body tie is exposed to the drain well. Accordingly, the advantage of the notched-body nFET is not simply that the total volume of silicon is reduced, but rather that the area of the drain to doped body-tie junction is significantly reduced.

To further probe the differences between the transient response of the two body-tie schemes, transients were also captured for the case in which all terminals were grounded, as shown in Figures 37-38. As described in [31], these transients can be used to calibrate the deposited charge within the device to a specific laser energy,

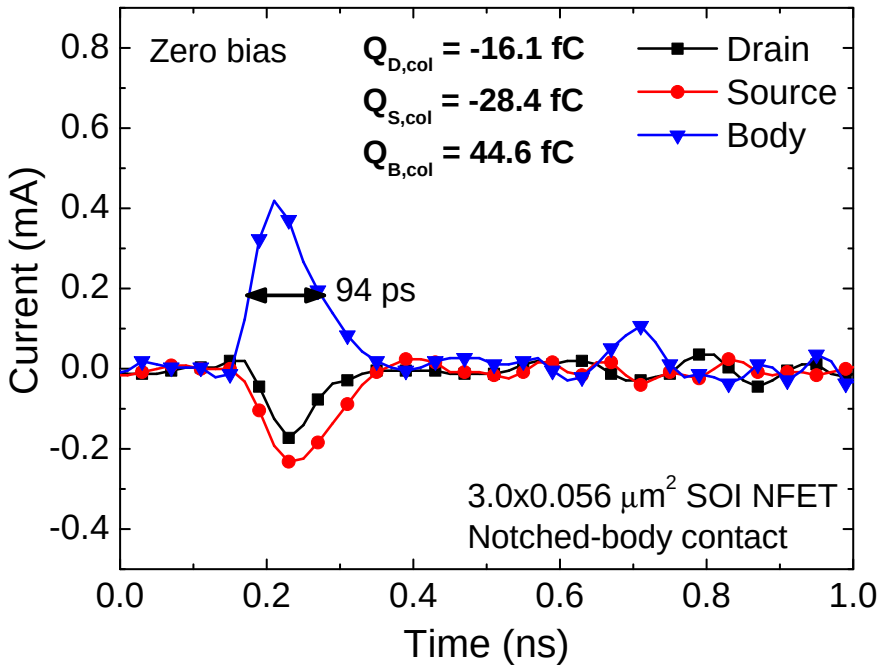

Figure 38: Transient currents at zero bias for incident laser pulse near body contact of a 3.0 µm x 0.056 µm notched-body nFET with a laser pulse energy of 10.7 nJ.

since at zero bias it is assumed that all electrons will be captured by the source and drain terminals and all holes will be captured by the body terminal. This can then be used to calibrate the laser energy for this specific device. The values for total charge collection are given in Figures 37-38, with the notched-body nFET collecting 30 % less charge than the T-body nFET. Again, this difference is greater than the difference in silicon volume, indicating an fundamentally different transient response between the body-contact schemes. These results can be used with those of Figures 35-36 to calculate the parasitic bipolar gain coefficient for these devices. A value of 5.7 is obtained for the T-body device and 4.5 for the notched-body device.

4.2.2 Additional Testing at Lower Laser Pulse Energies

Additional experiments were performed for a range of laser pulse energies using an modified experimental approach. In this case, multiple transients were captured at each position to account for fluctuations in the laser pulse energy as well as any other variations in the measurement setup, with an average of four transients per data point

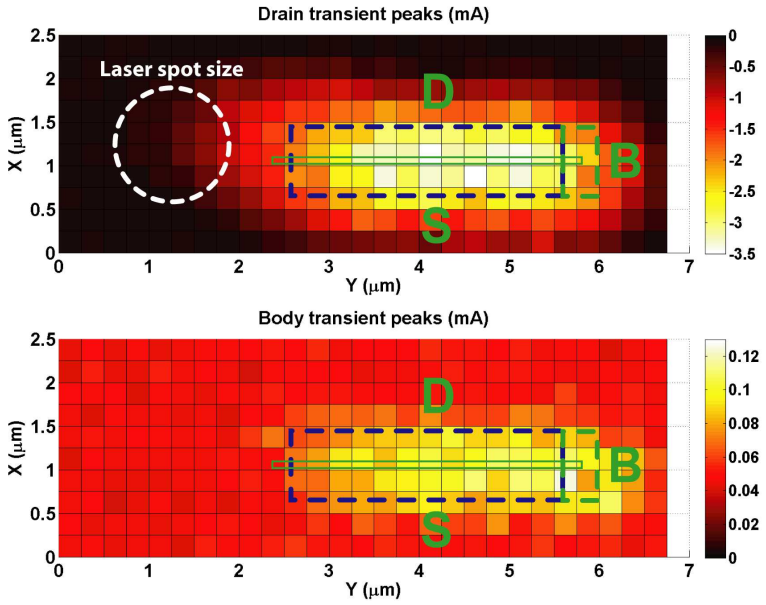

Figure 39: Peak drain and body currents for laser-induced transients as a function of laser position for a 3.0 µm x 0.056 µm T-body nFET with an incident laser pulse energy of 1.0 nJ. Each data point corresponds to an average of four transient measurements. The device is biased at $V_{DS} = 1.0\,V$ and $V_{GS} = 0\,V$.

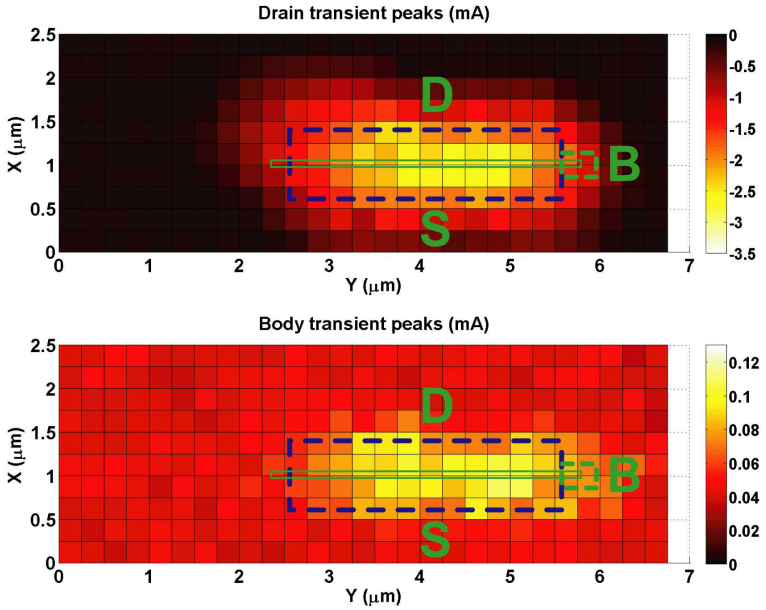

Figure 40: Peak drain and body currents for laser-induced transients as a function of laser position for a 3.0 µm x 0.056 µm notched-body nFET with an incident laser pulse energy of 1.0 nJ. Each data point corresponds to an average of four transient measurements. The device is biased at $V_{DS} = 1.0\,V$ and $V_{GS} = 0\,V$.

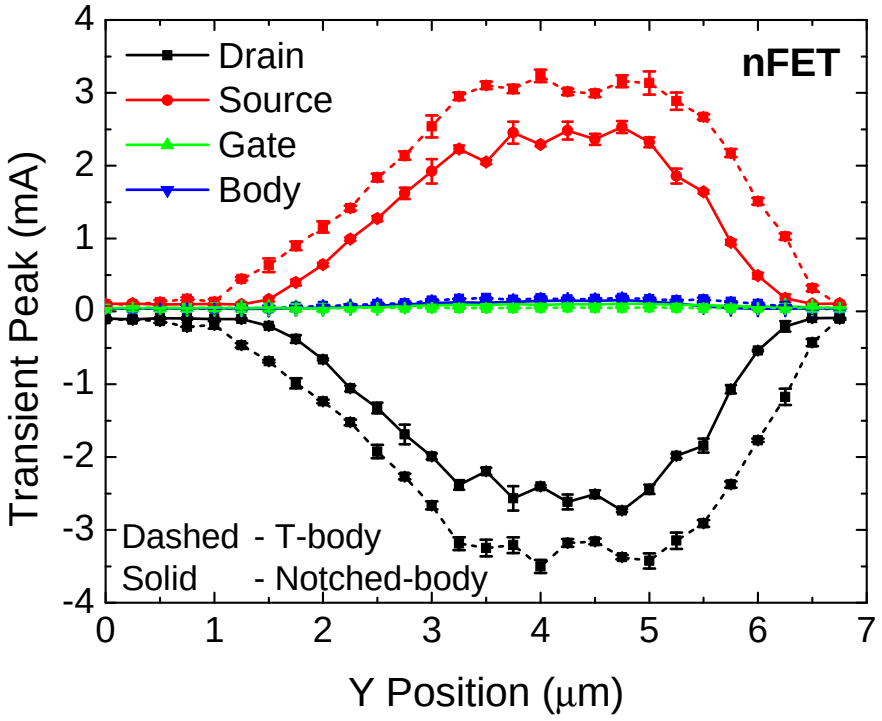

Figure 41: Peak transient currents for laser-induced transients as a function of laser position along the width of 3.0 µm x 0.056 µm T-body and notched-body SOI nFETs with an incident-laser pulse energy of 1.0 nJ. The devices are biased at $V_{DS} = 1.0$ V and $V_{GS} = 0$ V.

used to generate the data of Figures 39-41. 2-D plots of peak transient currents for a T-body nFET and a notched-body nFET are shown in Figures 39-40. Measurements were performed at laser pulse energies of 0.5 nJ, 1.0 nJ, and 3.7 nJ. Similar results were obtained at each laser energy; thus, for brevity, the measurements shown here were captured at a laser pulse energy of 1.0 nJ. According to the empirical expression developed in [94], a 1.0 nJ energy corresponds to a factor of 115 less deposited charge than in the case of Figures 31-38. The absolute current values are shown, in contrast to the normalized plots of Figures 31-32. The extent of the highly-sensitive area corresponds closely to the geometry of the each device, with the amplitudes fading off as the laser position moves away from the device in all directions. Furthermore, the data support the results of the previous experiment: compared to the notched-body device, the T-body device exhibits much larger drain and body transients for laser

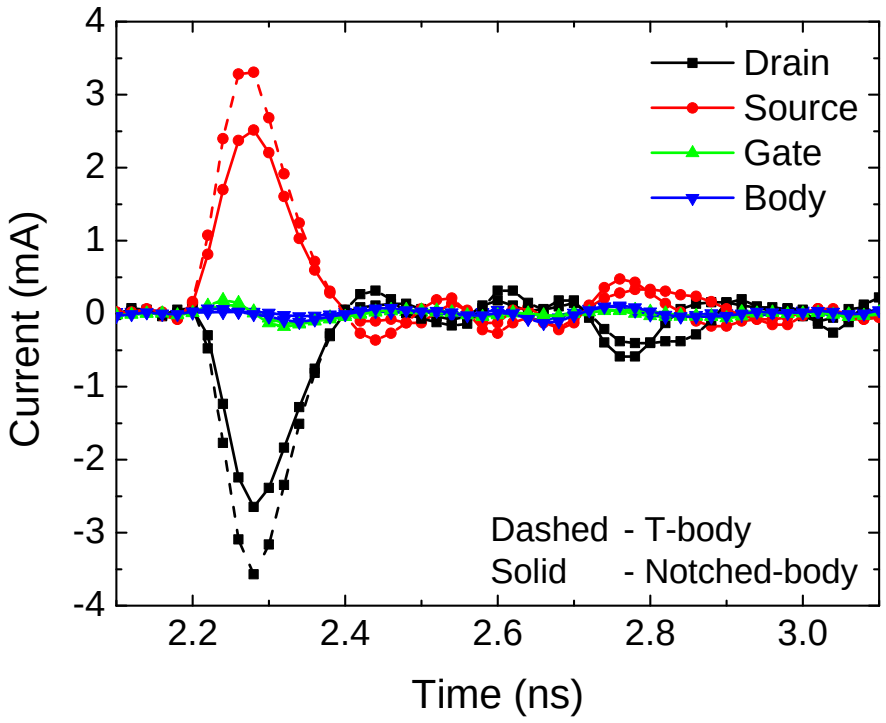

Figure 42: Transient currents for a laser position at the center of the device width of 3.0 µm x 0.056 µm T-body and notched-body SOI nFETs with a laser pulse energy of 1.0 nJ. The devices are biased at $V_{DS} = 1.0$ V and $V_{GS} = 0$ V.

positions near the body contact as well as uniformly larger transient magnitudes across the width of the device. This is more clearly illustrated in Figure 41, which plots the peak transient amplitudes for each terminal across the width of the devices, along with the corresponding error bars calculated from the multiple captured transients. The body transient is not as prominent as in Figure 33, because of the lower laser energy used in this set of measurements. A comparison of the T-body and notched-body transients taken from the center of the device width is shown in Figure 42. At this laser position, the center of the laser spot lies 1.5 µm away from the edge of the body-tie; as such, the difference between the transient responses is a result of fundamental differences between the charge-collection mechanisms of each device. There is no indication that the laser spot size has any effect on this conclusion. Furthermore, additional 2-D scans were performed to capture the transient response at a variety of laser energies, with comparable results to those plotted here. Transistor gate width

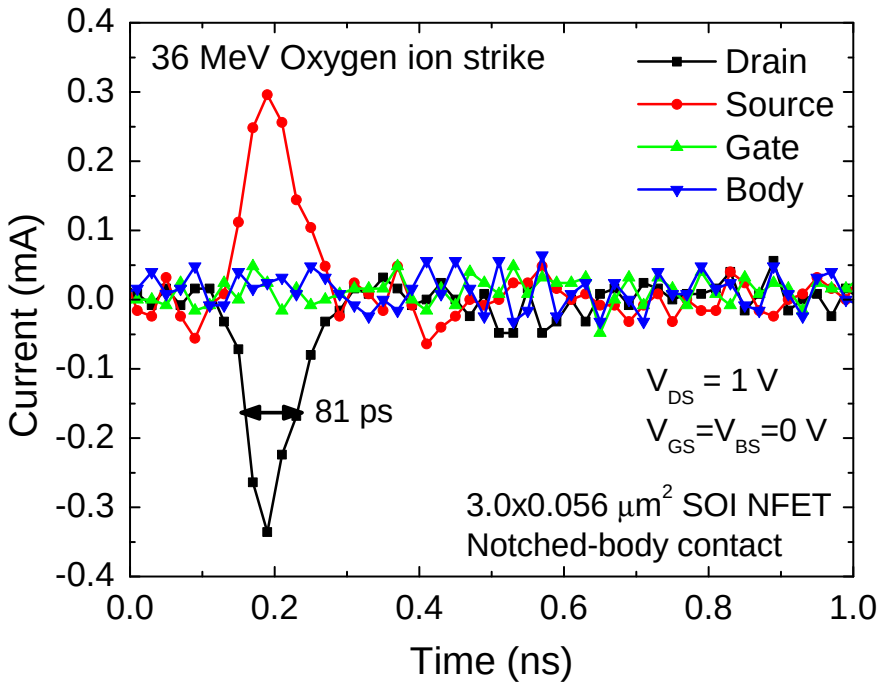

Figure 43: Single event transients in a 3.0 µm x 0.056 µm notched-body nFET biased at $V_{DS} = 1.0$ V for a 36 MeV oxygen ion strike.

was not found to have an effect on the difference in SET magnitudes between body tie layouts. Scans performed on 0.654 µm x 0.056 µm devices showed the T-body transients to be larger; however, since the gate width is smaller than the laser spot size, only the peak magnitudes could be compared.

Figure 43 shows the peak (drain current) heavy-ion-induced transient for a 2-D scan across a notched-body nFET. As expected, since the heavy-ion-induced transient results from a single oxygen ion with an LET of roughly 5.4 MeV $\cdot$ cm^2/mg, there is significantly less charge deposition and therefore much smaller current transients as compared to the laser-induced transients, in which higher energies can be achieved and charge is deposited throughout the larger laser spot size. Nevertheless, the transient waveform duration (FWHM 80–100 ps) and relative magnitudes between terminals are consistent, further validating the data collected from the TPA laser experiments. Microbeam-induced transients were measured from both notched-body and T-body nFETs, but no comparisons could be made between the two sets of data because

of the low transient amplitudes, with the body transient well below the noise floor; moreover, the drain and source transients varied too greatly in amplitude as a result of the 0.5 µm uncertainty in strike location. Ions with higher LET are necessary to resolve a large enough signal to meaningfully compare the various body-contacting schemes.

4.2.3 Trade-offs of Body-Tie Variants

In summary, position-dependent, laser-induced transients from 45 nm CMOS devices indicate that the SEE hardness of the notched-body MOSFET is enhanced with respect to the standard T-body MOSFET. The magnitudes of the drain transient are uniformly larger across the width of the T-body nFET as compared to the notched-body nFET, a conclusion that is supported across multiple pairs of devices. This coincides conveniently with the fact that the notched-body scheme carries the same benefits as the T-body scheme with respect to TID hardness. Moreover, the notched-body scheme minimizes the additional parasitics inherent to a body-contacted device; nevertheless, RF performance degradation still occurs in any body-contacted device compared to floating-body performance. However, for circuit and system applications intended for a radiation environment, body-contacts are necessary to suppress the parasitic bipolar amplification that enhances floating-body SEE sensitivity. These results point toward the notched-body device layout as the best compromise between RF performance, TID hardness, and SEE hardness.

CHAPTER V

ESTABLISHING BEST PRACTICES FOR MODELING CIRCUIT-LEVEL SINGLE-EVENT TRANSIENTS

As performance requirements increase and new applications emerge for space-based ICs, there is a greater need for simulation tools that can reliably predict their behavior when irradiated. Since the cost of carrying out a space mission is so great, its success is too critical to allow any risk. However, the cost of exhaustively testing and qualifying an IC technology is itself a significant hurdle. Consequently, many space-based systems rely on older IC technologies that already have a proven history of reliable operation in space environments. Highlighting this point is the fact that while 45-nm and 32-nm CMOS technologies are readily available in consumer electronics today, the most advanced radiation-hardened ICs on the market based on custom processes are at the 150-nm lithographic node [29]. Consequently, the computing power available for space-based electronics lags significantly behind that available for terrestrial electronics. Nevertheless, demanding new applications necessitate a move to more advanced IC technologies, particularly for broadband satellite communications. With this in mind, predictive modeling tools play a critical role in developing space-qualified IC technologies, since they can be applied to quickly assess the dominant radiation sensitivities and test radiation-hardening techniques, minimizing the time and expense of experimental testing. A key requirement, however, is that the device and circuit modeling tools are truly predictive; they must be tested and validated against experimental results of radiation effects in modern IC technologies. This is particularly challenging in the realm of single-event effects, because of the diverse nature of single-event effects and their variance based on the underlying transistor

type, circuit class, circuit topology, and system-level configuration. A wide variety of physical mechanisms can operate within a semiconductor device during a transient radiation event, and the resulting current and voltage transients can impact the circuit in a multitude of ways.

5.1 Motivation for Using Mixed-Mode TCAD to Model Single-Event Effects

Modeling techniques for single-event effects at the circuit level can be divided into two general categories: decoupled and coupled. Decoupled modeling techniques integrate device-level current transients within a circuit without accounting for feedback from the circuit. The circuit-level SET simulation is typically performed by first choosing a set of device-level transient waveforms to represent the ion strike, then employing independent current sources to inject these currents into the circuit nodes of interest (e.g., the terminals of the device that is struck by a heavy ion) [1, 50, 76]. The SET currents can be injected in a variety of ways, but the most common attribute is that a current source is employed, either in the circuit netlist or directly within the compact model. An example of the basic current-injection configuration surrounding a SiGe HBT is given in Figure 44. The device-level transient waveforms are often analytical formulations, such as double-exponential or trapezoidal current pulses, or more recently, piecewise-linear current waveforms obtained from 3-D TCAD simulations of the device of interest [65, 78]. The key shortcoming of these decoupled approaches is that the current sources inject a predefined current waveform into the circuit nodes without regard for how the circuit response might influence the device-level transient current. In older technologies where the circuit response times are orders of magnitude longer than the device-level transient durations, this is not an issue, but in modern high-speed circuits (e.g., SiGe), the effect of circuit feedback on device transients becomes increasingly important. Circuit feedback can transform the device-level transient waveforms and affect the total charge collection, leading to greatly different results

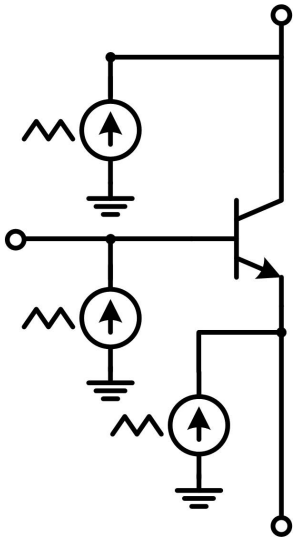

Figure 44: Basic configuration of device-level current transient modeling for decoupled current-injection approaches.

at the circuit output. Moreover, if ideal current sources are used, then the circuit is forced to conform to the injected currents and can exhibit unphysical behavior [43, 63, 106], which not only can lead to the wrong SEE predictions, but can also obfuscate the underlying physical SEE mechanisms that drive circuit- and system-level errors.

In contrast, coupled mixed-mode TCAD possesses the advantage of capturing the dynamic interaction of the physical charge collection process with the surrounding circuit. Mixed-mode TCAD capabilities are provided by a number of different TCAD software tools, including Sentaurus Device from Synopsys, NanoTCAD from CFD Research Corporation, ATLAS from Silvaco, and MINIMOS-NT. In these tools, the ion-strike and ensuing charge transport are modeled within a multidimensional model of the semiconductor device, and the charge transport equations are solved simultaneously with the operating conditions of the compact modeled circuit at each time step of the transient event. Thus, as the nodal biases and impedances looking out from the irradiated TCAD device evolve over the course of the SET, they impose new boundary conditions and can result in significantly different device-level charge

transport. The impact of circuit feedback is highly dependent on the physical processes within the device as well as the circuit topology and loading conditions on the TCAD device terminals. Although compact models have been developed that address the bias dependence of device-level transients and the influence of the external circuit [33, 43], these cannot be applied generally to all circuits, since the nature of circuit feedback is so greatly dependent on the circuit topology, bias conditions, and device type, doping, and geometry. Moreover, the models in [33, 43] were themselves derived using mixed-mode TCAD results, highlighting the unique ability of fully coupled mixed-mode TCAD to accurately predict circuit-level SET and reveal the dominant physical mechanisms.

Beyond its ability to capture circuit feedback, mixed-mode TCAD provides another unique ability to designers, in that it can be used to test and develop device-level radiation-hardening-by-design (RHBD) techniques. Since device-level RHBD techniques involve changes to the physical device, they cannot be readily tested using compact models. During research and development of a new device structure, it would be prohibitively slow for a device designer to build a custom compact model for each variation. However, it is relatively straightforward to develop a physical TCAD device model; once done, mixed-mode TCAD simulations of test circuits could be performed to evaluate the effectiveness of different designs, enabling a simple iterative process to optimize an RHBD device to achieve the most effective SEE mitigation.

5.2 Accurate Modeling of Single-Event Transients in a SiGe BiCMOS Voltage Reference

Single-event effects (SEE) are also a key area of concern for space-based SiGe circuits. Clear guidelines must be established as to which approaches to modeling SET are valid for various conditions (circuit topology, technology node, device geometry, environment, etc.). With this in mind, four different approaches to modeling circuit-level SET are investigated here, using as a test case a SiGe bandgap reference (BGR) circuit [73] for which measured data have been published [72]. The goal is to first assess how well the simulation approaches correlate to the measured data, then illuminate for the first time the underlying circuit SET mechanisms, and finally use these conclusions to explain the differences between the simulation results and illuminate possible pitfalls and best practices for circuit SET modeling. The first three simulation approaches utilize a strictly compact-model-based circuit in which injected current transients model the effects of a heavy ion strike. This technique has been used to successfully match simulated transients to measured data in an analog circuit [1], but in this older-generation circuit the SEE response was several orders of magnitude longer than the single-device response. The differences between the compact-model-based approaches lie in the origin of the injected current transients: (1) analytical double exponentials; vs. (2) 3-D TCAD computed transients for a negative substrate bias; vs. (3) 3-D TCAD computed transients at the corresponding circuit nodal biases. The fourth approach is to use a full mixed-mode simulation, in which a 3-D physical TCAD model is substituted in place of a compact model and solved simultaneously with the full compact-modeled circuit [108].

First, the BGR circuit and experimental results will be described, followed by the circuit-level simulation approaches and the bias dependence of the HBT transient response as illuminated by 3-D TCAD simulations. The simulated SET response of the voltage reference circuit using each approach is then presented and compared to

the measured data. The BGR output transient is deconstructed using the mixed-mode simulation results to determine its driving components; having established these mechanisms, the reasons underlying the similarities and differences of the simulation approaches are revealed, providing insight for future work on circuit SET in different circuit topologies and device technologies.

5.2.1 Background

Experimental data have been reported on laser- and microbeam-induced transients in single devices from a first-generation SiGe BiCMOS technology [80]. Recently, measured microbeam-induced transients have also been published for a SiGe precision voltage reference circuit fabricated in this same SiGe technology [72]. Results of these studies show the strong bias dependence of individual HBT transients and the much longer transients at the output of the voltage reference compared to that of the single transistor (~300 ns and ~10 ns, respectively). The SiGe voltage reference is described in [72, 73]; its schematic is reproduced in Figure 45. An exponential, curvature-compensated BGR [73] was used to provide the reference voltage to the positive input of an operational amplifier (opamp) [72]. The opamp is a two-stage amplifier followed by an emitter-follower buffer from which the experimental SETs were measured, and it is biased with an on-chip current source. The SiGe BGR circuit is designed to generate an output voltage of 1.17 V at room temperature, and an output voltage of 1.65 V is expected from the regulator. During normal circuit operation, the substrate is grounded, and a power supply of 3.3 V is used to bias the circuit. The output voltage of the BGR is defined by the sum of two components: (1) the V_{BE} of transistor Q3 and (2) the voltage drop across resistor R2, which is determined by the sum of the base current flowing through Q3 and the collector current flowing through Q5.

CFDRC's NanoTCAD tool was used to perform simulations of normally incident

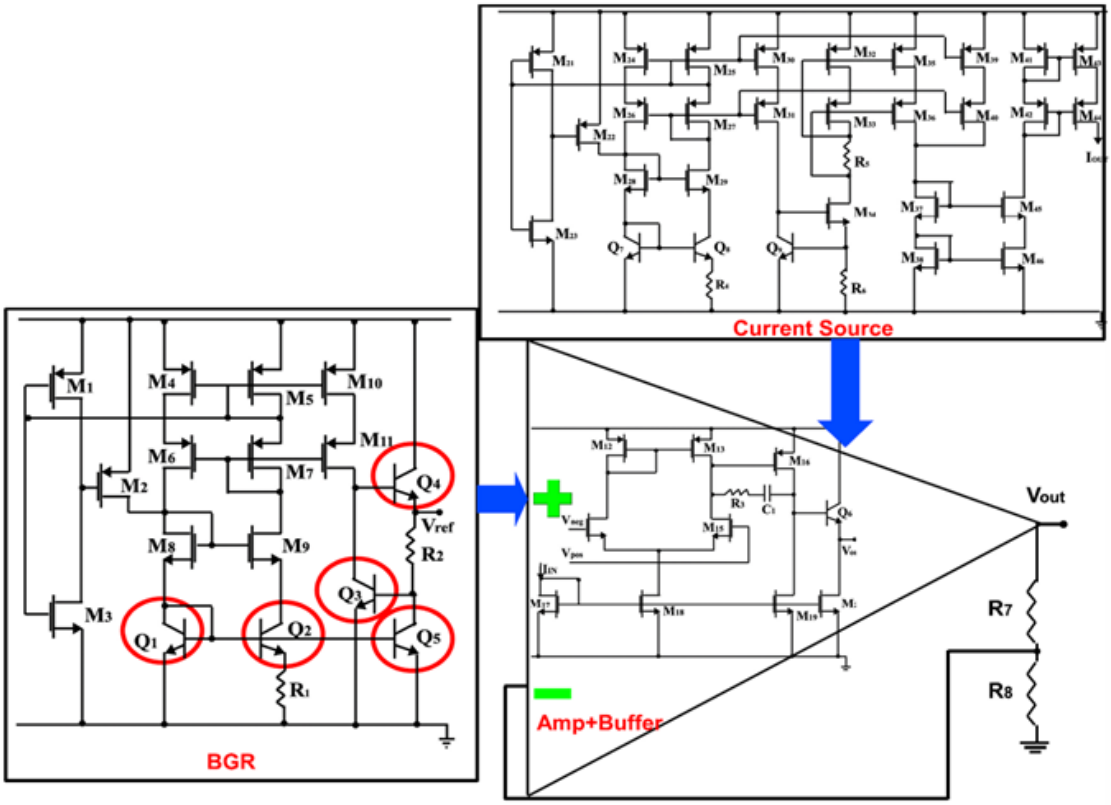

Figure 45: SiGe voltage reference circuit schematic. The circuit exhibits the greatest sensitivity to transients induced in transistor Q2 [72].

emitter-center ion strikes on transistor Q2 (0.5 µm x 2.5 µm) SiGe HBT), since experimental ion strikes on this transistor produced the largest circuit output transients. To model the complex ion track, the SRIM software tool [100] was used to compute an energy-deposition vs. depth profile for the 36 MeV oxygen ion used in the microbeam tests, taking into account the back-end-of-line (BEOL) layers present above Q2. This variable linear energy transfer (LET) profile was then imported into NanoTCAD using its automated ion track meshing capability. In the voltage reference circuit, transistor Q2 is biased at $V_B = 0.74\,V$, $V_E = 0.05\,V$, $V_C = 0.74\,V$, with the substrate grounded. The peak of all TCAD ion strikes presented here occurs at 2 ps. The measurement setup and experimental conditions are detailed in [72, 80].

5.2.2 Single-Event Transient Simulation Approaches

5.2.2.1 Compact Model Approaches

The first simulation approach is to inject an analytical double-exponential current transient within the Spectre model of the BGR at the terminals of Q2. Analytical current sources have been used extensively for both analog [78] and digital circuits [71] as well as to investigate novel device architectures [82]. This approach is easily implemented and enables a straightforward analysis of the critical LET or collected charge associated with SEE. In addition, the use of an analytical transient avoids convergence issues that can arise with piecewise-linear transient sources based on TCAD or measured transients [43]. The double-exponential source used in this work is calibrated to 3-D TCAD simulations and is illustrated in Figure 46. In the second approach, a piecewise-linear current source is used to inject the 3-D TCAD computed transients at the terminals of Q2. In this case, the device terminals are grounded with a substrate bias of $-4\,\mathrm{V}$, since bipolar logic is primarily sensitive to strikes on the off transistor as a result of charge collection at the collector terminal through the collector to substrate junction. This approach was chosen to provide a point of comparison both for the bias dependence of the device-level transients as well as a measure of how this bias dependence is reflected at the circuit output. The third simulation approach uses piecewise-linear transients based on TCAD transients, biased in this case at the nodal potentials of Q2. Both sets of TCAD transients are shown in Figure 46. The primary drawback of these approaches is that they are based on single-transistor transients in which there is no loading or feedback from the external circuit [43, 71].

5.2.2.2 Bias Dependence of SiGe HBT SET

Figure 46 demonstrates that the SET of a SiGe HBT possesses a strong bias dependence. The SET simulated at the nodal biases of Q2 demonstrates greatly increased collector and emitter transients compared to that of the HBT with a negative substrate bias

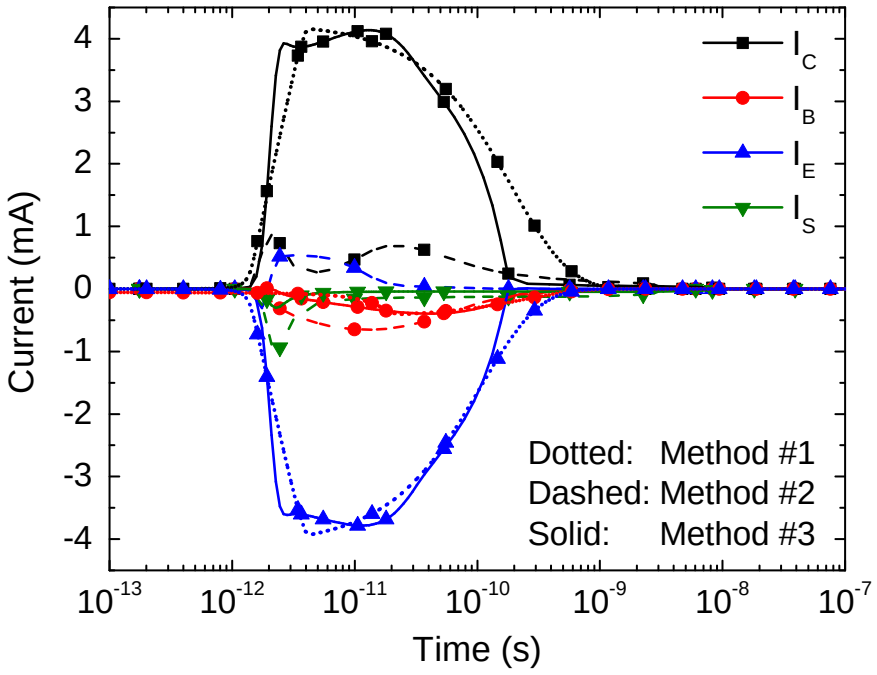

Figure 46: Comparison of injected device transients from each simulation approach: (1) double exponential, (2) 3-D TCAD SiGe HBT transients at $V_{SX} = -4\,V$, and (3) 3-D TCAD transients at Q2 biases. Transients are for a single device not within the circuit.

and grounded emitter, base and collector. Two mechanisms have been proposed to explain the large emitter to collector transient: (1) bipolar action, in which modulation of the base potential induces an increase in forward bipolar current, and (2) the ion shunt effect, in which the carrier densities are sufficiently high along the ion track such that the emitter is shorted to the collector by an electron-hole plasma wire, leading to large transient currents for nonzero V_{CE} [39, 47, 48]. The energy bands in Figure 47, taken from a line probe through the center of the HBT, show a nearly linear slope from the emitter to collector for the duration of the large emitter to collector current, with slight deviations in the base region resulting from the presence of the SiGe layer. Furthermore, the electron-density profile during this time period is approximately flat except for minor variations in the region with bandgap grading, consistent with resistor-like shunt current from emitter to collector. Although these facts suggest that the ion shunt effect dominates the collector and emitter transients of transistor Q2,

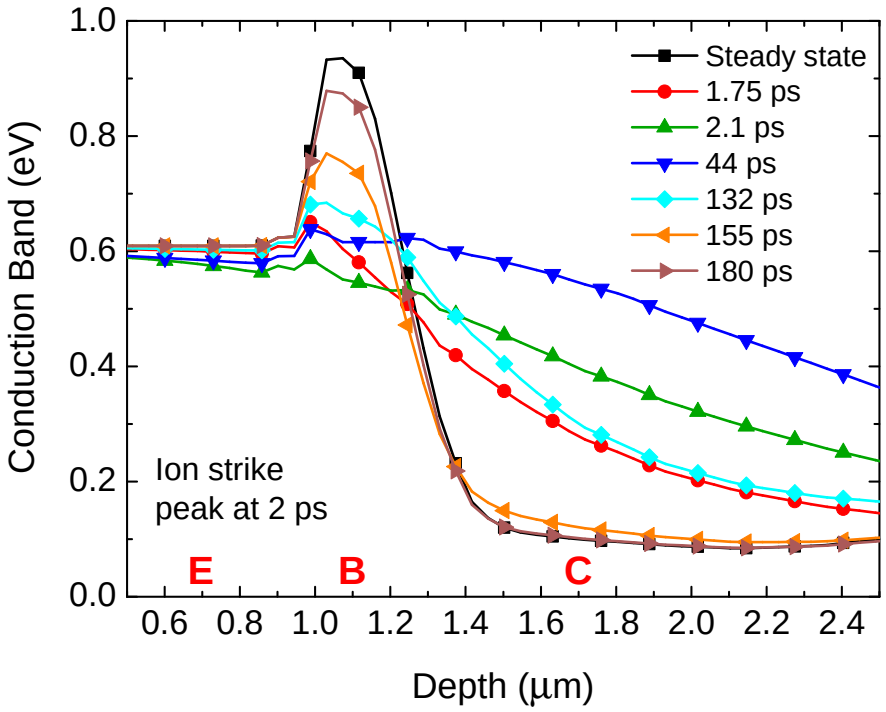

Figure 47: Time evolution of the conduction band energy within a SiGe HBT during a simulated ion strike, taken from a line probe through the center of the emitter. The ion strike peak occurs at 2 ps, with a Gaussian decay of 250 fs.

more research should be done to confirm the details of the SET mechanism. However, regardless of which specific mechanism is at work, the net result of an ion strike on a SiGe HBT biased in the forward-active region is an amplification of the SET and a corresponding increase in the total collected charge, highlighting the importance of addressing the impact of transistor bias when investigating SEE in analog/RF circuits.

5.2.2.3 Coupled Mixed-Mode TCAD

In contrast to strictly compact-model-based approaches, true mixed-mode simulations, as described in [106, 108], possess the advantage of the 3-D TCAD device being exposed to the dynamic biases present in the circuit throughout the SET, at the cost of increased computational complexity. The compact-model-based approaches require an initial TCAD transient simulation for each transistor and bias case of interest (one of which typically completes within four hours), followed by circuit transient simulations

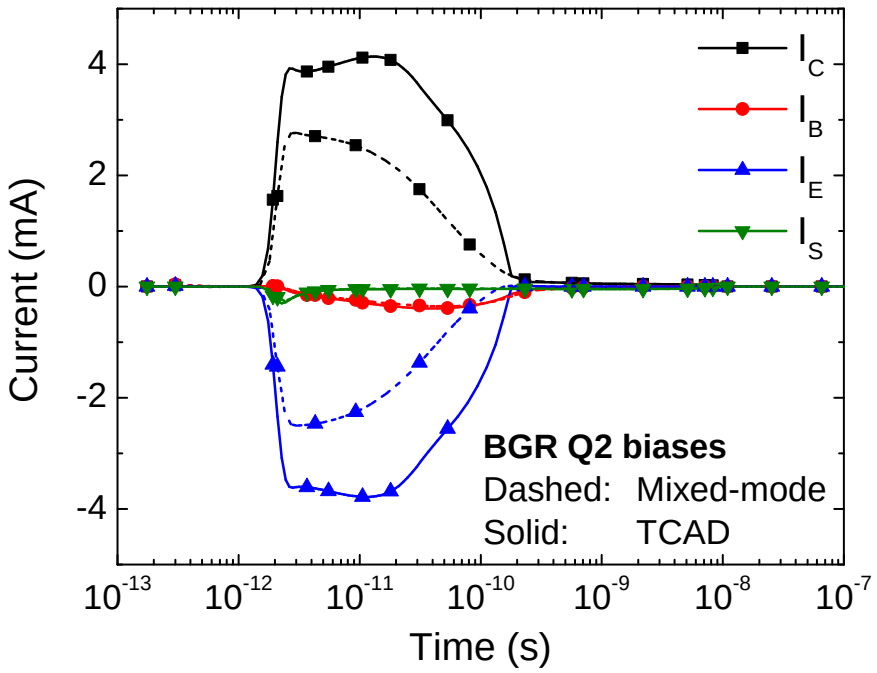

Figure 48: Comparison of device terminal transients computed from a 3D TCAD simulation and from within the full 3-D mixed-mode circuit simulation.

of minimal duration. In contrast, a single mixed-mode transient simulation does not require an initial TCAD transient simulation and typically completes within 10-12 hours. Computing the TCAD solution within a circuit enables mixed-mode simulations to account for feedback and loading from the external circuit that can in principle alter the device transient currents as they evolve over time. The mixed-mode simulations in this paper were performed using CFDRC's MixCad tool (3-D NanoTCAD interface to Cadence Spectre) [108], using a calibrated 3-D TCAD model of a first-generation SiGe HBT. Whereas mixed-mode studies such as [112] are limited to basic SPICE passives and compact models, the unique interface between NanoTCAD and Spectre allows this mixed-mode tool to be used directly with the compact models included in commercial process design kits. Thus, the mixed-mode simulations presented here contain a 3-D TCAD device operating within the final circuit design as submitted for fabrication, including extracted layout parasitics.

Figure 48 illustrates the difference in the transient currents at the HBT terminals

for a 3-D TCAD simulation at steady-state circuit biases (Method #3 in Figure 46) and for a TCAD device within a mixed-mode circuit simulation. The magnitude of the mixed-mode emitter and collector transients is significantly reduced, leading to a much lower collected charge (0.49 pC vs. 0.74 pC at the collector); thus, the current-injection approach overestimates the SET at the device level within this particular circuit. Given the strong bias dependence of SET in the SiGe HBT, along with the widely varying impedances to which devices in analog/RF circuits are exposed, this issue will be heavily dependent on the circuit topology in question. In this case, the emitter transient current causes the emitter voltage to rise because of the voltage drop across R1, whereas the collector transient current causes the collector voltage to drop, thus reducing V_{CE} and limiting the shunt current between the emitter and collector.

5.2.3 SiGe Voltage Reference Simulations

5.2.3.1 Circuit-Level Transients

To provide a basis for comparing the simulated SET to that reported in [72], the entire measurement path was modeled. Starting from the circuit output, the modeled elements include the bond-pad capacitance, bond-wire inductance, distributed coaxial transmission line, bias tee, and oscilloscope resistive and capacitive loads; a representative diagram of these circuit elements is given in [107].

Figures 49-50 shows the simulated transients at the outputs of the BGR and the regulator circuits. The duration of the BGR output transient increases by an order of magnitude relative to that of the single transistor (from 10 ns to 100 ns), agreeing with the lengthening of transients in circuits versus individual devices that has been observed experimentally [72]. Likewise, the transient at the output of the voltage regulator increases in duration by nearly a factor of two, approaching the duration of the measured voltage-reference transient. This transient lengthening is consistent across all simulation approaches, indicating that the transient lengthening is driven by the parasitics of the circuit as the signal propagates to the output.

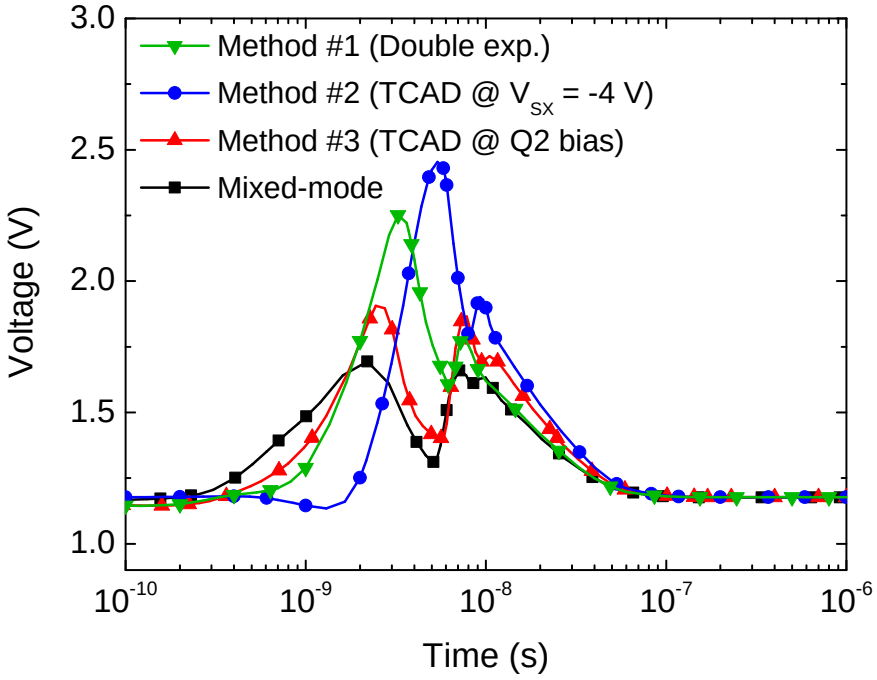

Figure 49: Comparison of circuit transients at the BGR output as simulated according to the different approaches.

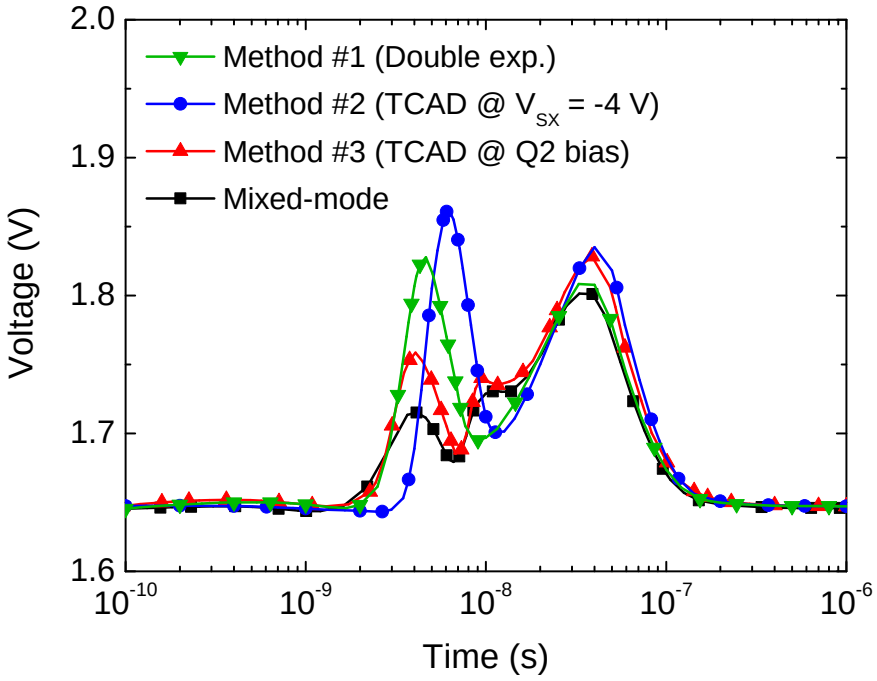

Figure 50: Comparison of circuit transients at the regulator output as simulated according to the different approaches.

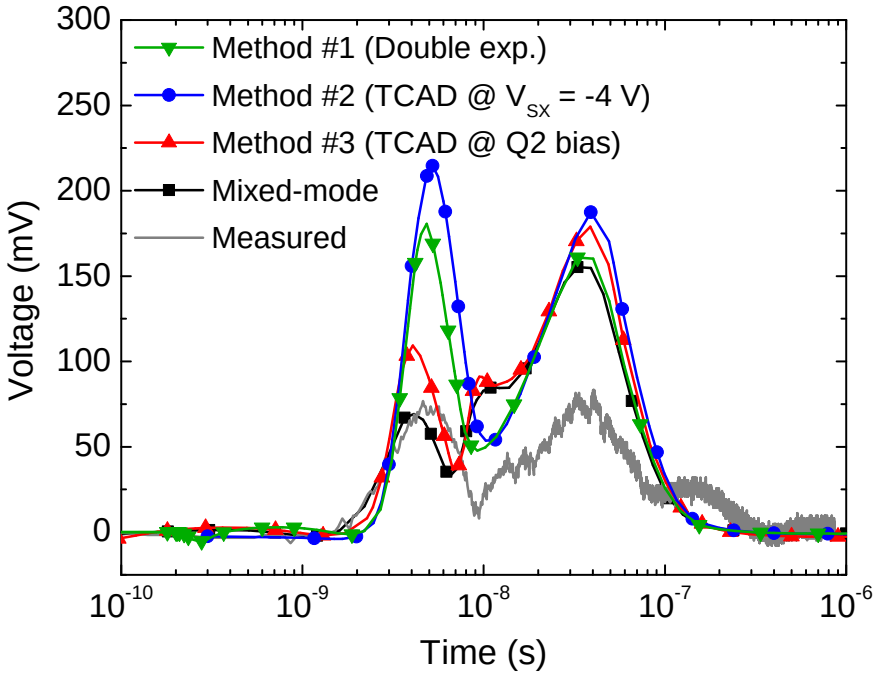

Figure 51: Comparison of simulated and measured transients at the oscilloscope input. The time scales of the simulated transients have been shifted to align with the measured transient.

Figure 51 shows the simulated transients as they would be measured by the oscilloscope, overlaid upon an actual measured transient from a 36 MeV oxygen ion strike, taken from Location 1 of [72], a representative emitter-center strike at Q2. The bias tee present at the oscilloscope terminal removes the DC bias of Figure 50. The first peak of the mixed-mode result correlates well with the data, and its second peak qualitatively follows the shape of the measured transient. As expected from the overestimated transients in the single-transistor simulations of Figure 48, Method #3 (fixed circuit biases) overshoots the mixed-mode transient at the oscilloscope, further exceeding the measured data. Method #1 produces a similar result, since its double exponential input transients are calibrated to the 3-D TCAD transients of Method #3. Despite the large difference in its device-level transients, Method #2 produces a comparable transient at the oscilloscope. Although the mixed-mode simulation demonstrates the closest agreement to the magnitude and temporal structure of the

measured transient, all four approaches give similar qualitative insights into the circuit transient response.

5.2.3.2 Analysis of BGR Output Transient

The factors that drive the temporal shape of the BGR output transient can be identified by carefully tracing the current transients from their origin at Q2 to the circuit output. Since the four approaches give similar qualitative results, I analyze in detail the mixed-mode simulation. The first important fact is that Q2 is composed of an array of 32 HBTs wired in parallel; in both experiment and simulation, a given SET event is recorded when only one of these 32 HBTs is subject to a heavy-ion strike. Charge sharing between adjacent HBTs is negligible, indicated by the fact that the sensitive areas of each HBT in Q2 as shown in [72] do not overlap. Moreover, 3-D TCAD simulations encompassing two HBTs spaced apart as in the physical layout demonstrate that for an ion strike to the emitter-center of one HBT, the peak transient collector current on the adjacent HBT is three orders of magnitude less than that of the irradiated device; this is also confirmed in [75]. The mixed-mode circuit simulation shows that a large portion of the collector and emitter current originating in the irradiated HBT is absorbed by the 31 parallel compact model HBTs. Figure 52 shows the ion-induced transients from the 3-D TCAD HBT along with the total current transients at the 31 parallel HBTs. The resulting net transients from the Q2 array are significantly different from the irradiated device transients, with reduced peak collector and emitter currents, and increased negative base current. Referring to the schematic in Figure 45, the base of Q2 is shared with transistors Q1 and Q5. Consequently, the net base current flowing from Q2 is divided between Q1 and Q5.

Figure 53 shows the terminal current transients for Q5. The initial rising base transient is capacitively coupled to the emitter, but bipolar conduction begins to

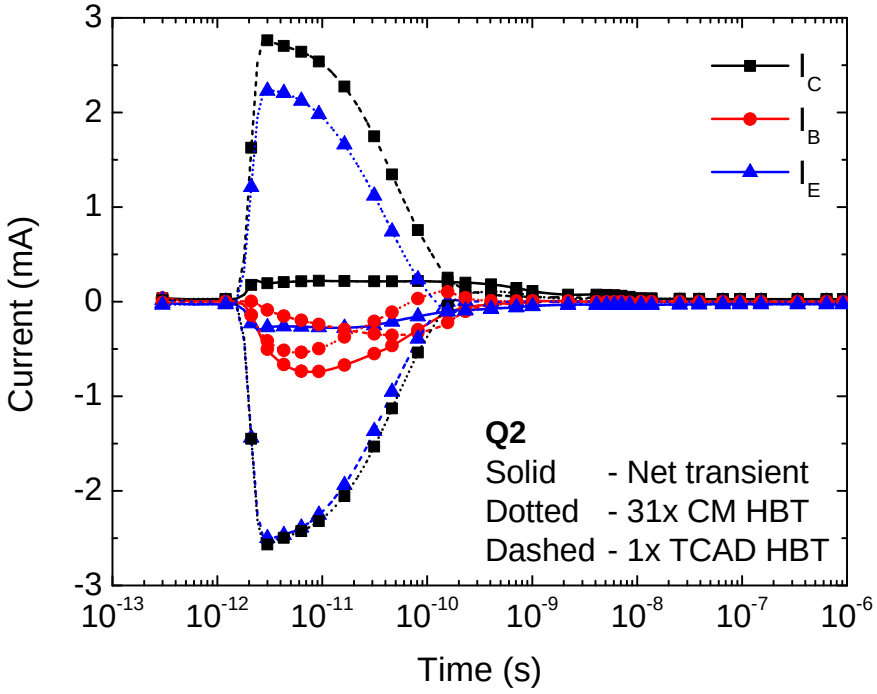

Figure 52: Mixed-mode transient currents at Q2: currents at the single irradiated TCAD transistor (dashed), currents at the 31 parallel compact model HBTs, and net transient currents to the remainder of the circuit from the 32x array (solid).

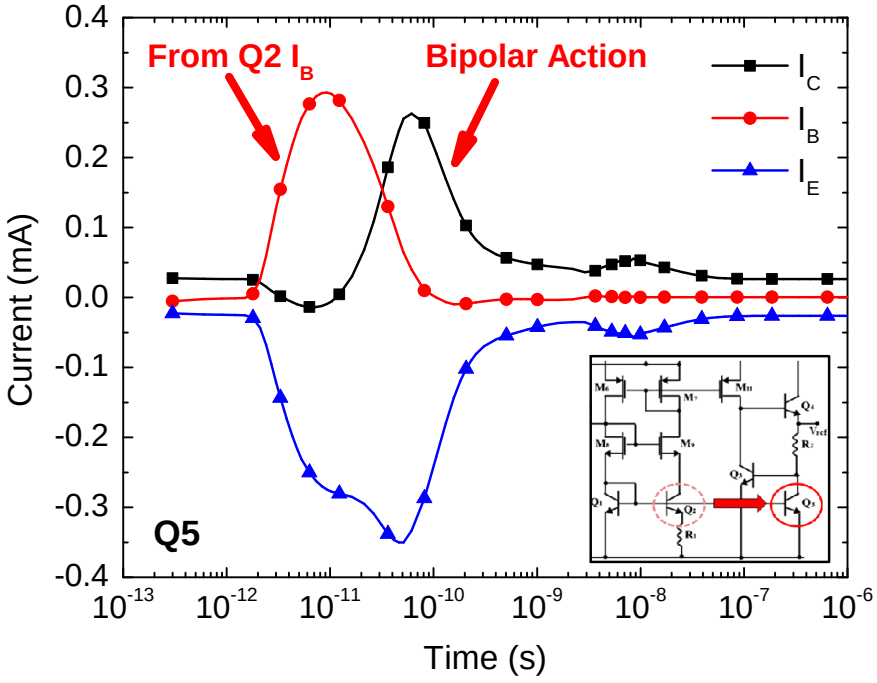

Figure 53: Mixed-mode current transients at terminals of Q5.

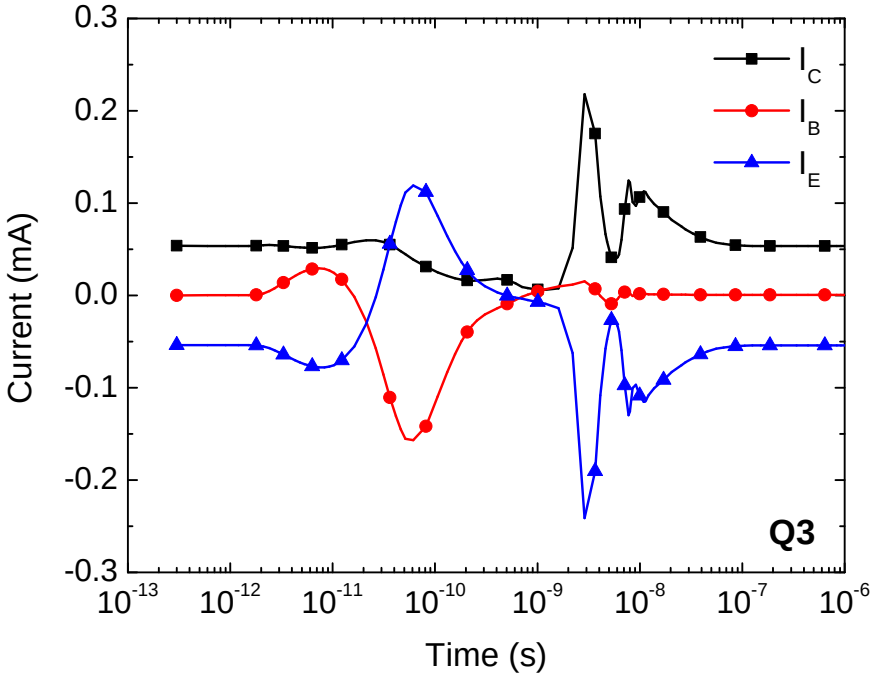

Figure 54: Mixed-mode current transients at terminals of Q3.

dominate as a result of modulation of the base potential for times greater than approximately 30 ps. Although the Q5 collector current transient will lead to modulation of the BGR output voltage through the voltage drop across R2, the influence of Q3 must be considered, since the base of Q3 is tied to the collector of Q5. Figure 54 plots the terminal current transients for Q3, in which the initial negative base and positive emitter currents show that capacitive coupling of the base and emitter of Q5 supplies a portion of the collector current of Q3. The influence of the resistive drop resulting from current in the output branch can be estimated by summing the Q3 base and Q5 collector transients, as plotted in Figure 55. Three peaks emerge in the sum of the two currents: near 100 ps, near 2 ns, and near 10 ns. The first and third peaks are driven by the collector current of Q5, with the second peak driven by the base current of Q3. The output voltage then increases according to the total transient current through R2. Due to the distributed resistance and parasitic capacitance of R2, the high-frequency components of the current transient are transformed to lower

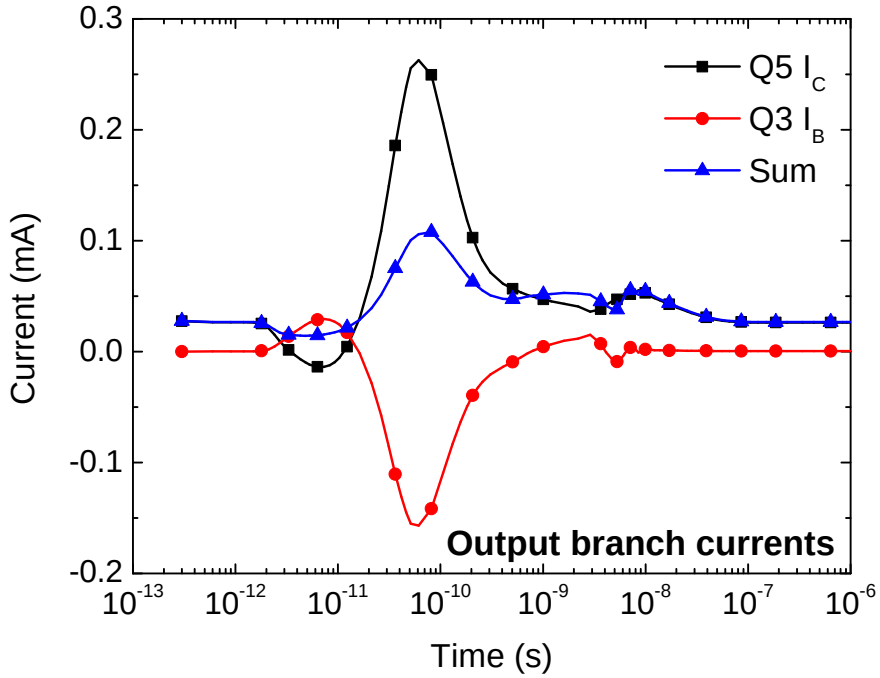

Figure 55: Mixed-mode current transients that contribute to the BGR output voltage through the resistive drop across R2.

frequencies as it propagates through the resistance network, as illustrated in Figure 56. Consequently, the first and second peaks of the summed current in Figure 55 are conflated together in the output transient.

Whereas the initial rise in the output voltage originates with the increasing Q5 collector current, the subsequent output response is determined by a damped feedback loop involving Q3, Q4, Q5, R2 and M11. As the output voltage rises because of the current transient through R2, the voltage at the drain of M11 also increases because of the emitter follower Q4 until M11 is ultimately driven into the linear regime near 1 ns. This process is illustrated in Figure 57, which plots key transistor biases for the duration of the SET. The simulation results show that the source voltage of M11 rises slightly to compensate for the rising drain voltage, but is limited by the cascoded M10. The V_{SG} of M11 also increases slightly as M7 conducts more current to support the collector transient of Q2 as it propagates through M9. As M11 is pushed out of saturation, its drain current decreases significantly, reducing the base current that

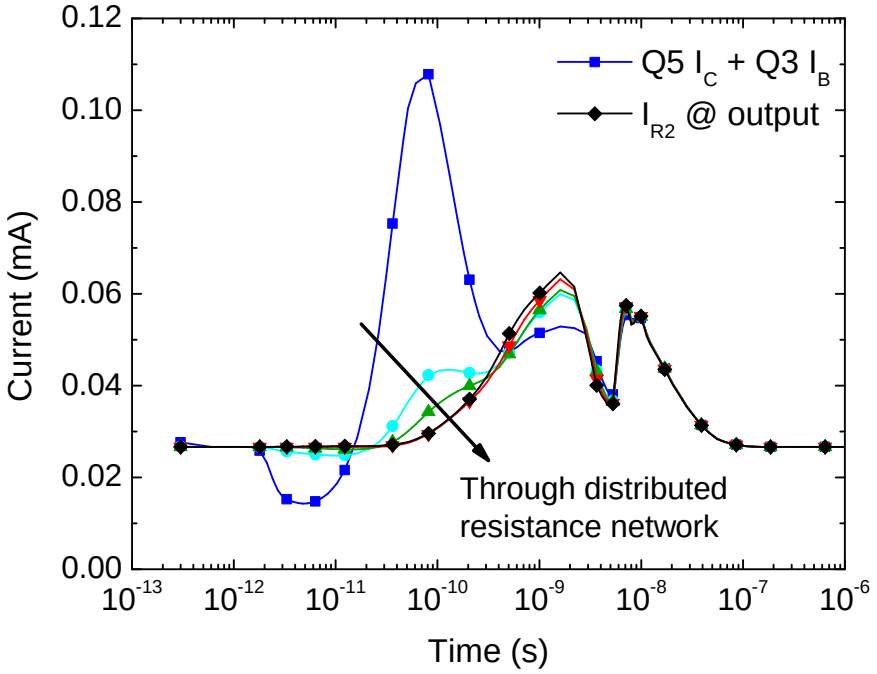

Figure 56: Transformation of current transient as it flows through resistor R2, a polysilicon resistor modeled as a distributed network of resistors and parasitic capacitors.

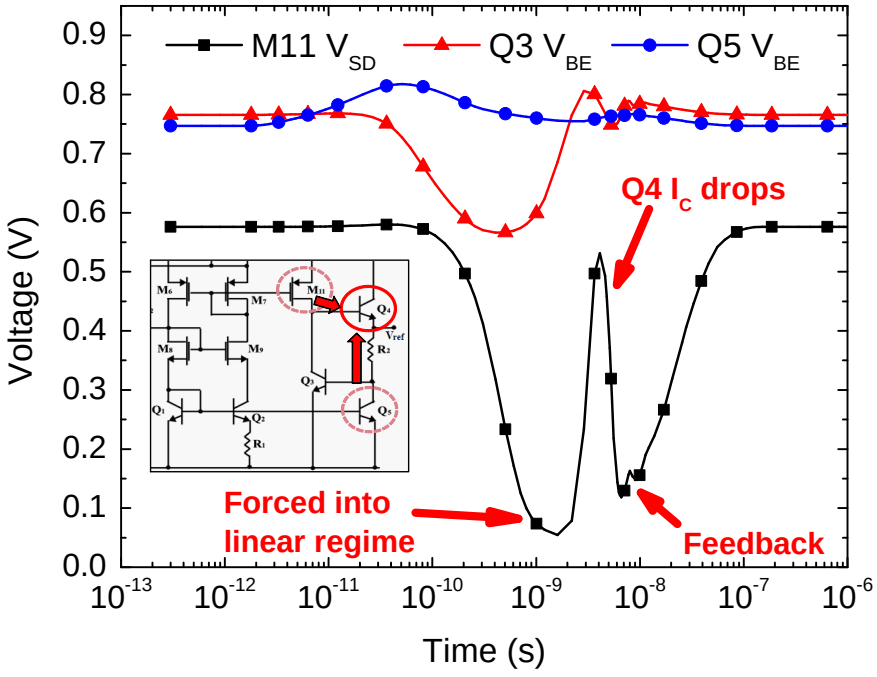

Figure 57: Mixed-mode voltage bias transients for transistors M11, Q3, and Q5.

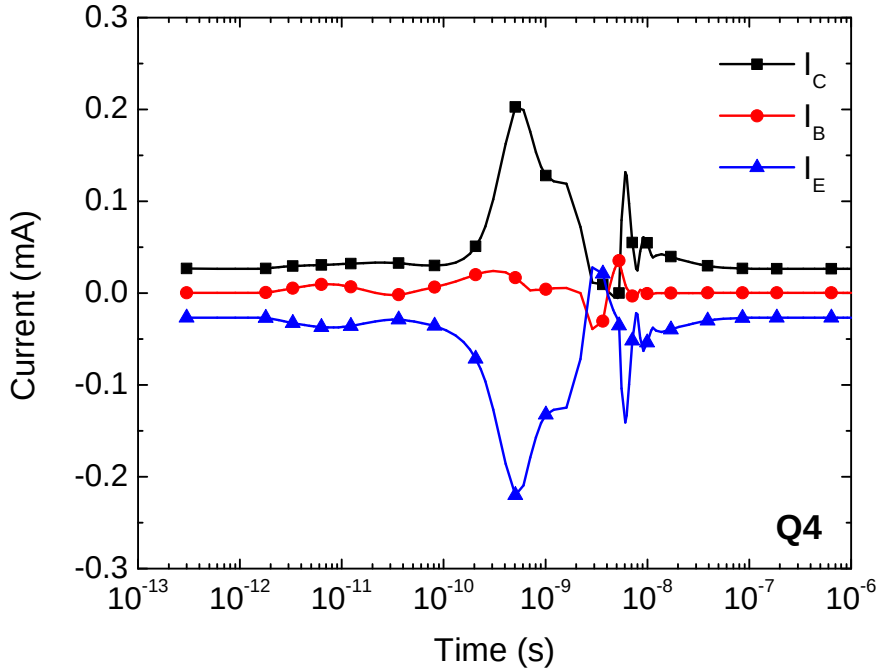

Figure 58: Mixed-mode current transients at the terminals of transistor Q4.

flows into Q4, as shown in Figure 58; this forces a reduction of the emitter current of Q4, followed by a decrease in the output voltage as less current flows through R2. However, as soon as the output voltage begins to decrease, M11 reenters saturation and experiences a sharp increase in its drain current around 3 ns. The collector and base currents of Q3 increase accordingly, and the emitter current of Q4 increases sharply as more current flows into its base, causing the Q5 collector current to increase. Together these currents cause the output voltage to rise again until M11 once again approaches linear operation. The feedback from the drain current of M11 is less severe at this point, since the current flowing through R2 peaks at a lower magnitude. The circuit output transient is then controlled by decreasing oscillations of the Q3 base current and a steady decrease of the Q5 collector current.

Since the BGR output voltage is the sum of the V_{BE} across Q3 in addition to the voltage drop across R2, the transient response of V_{BE} must also be considered. Figure 59 plots the BGR output transient along with these components, with the DC offset removed to identify the contributions of each component. This plot reveals that the

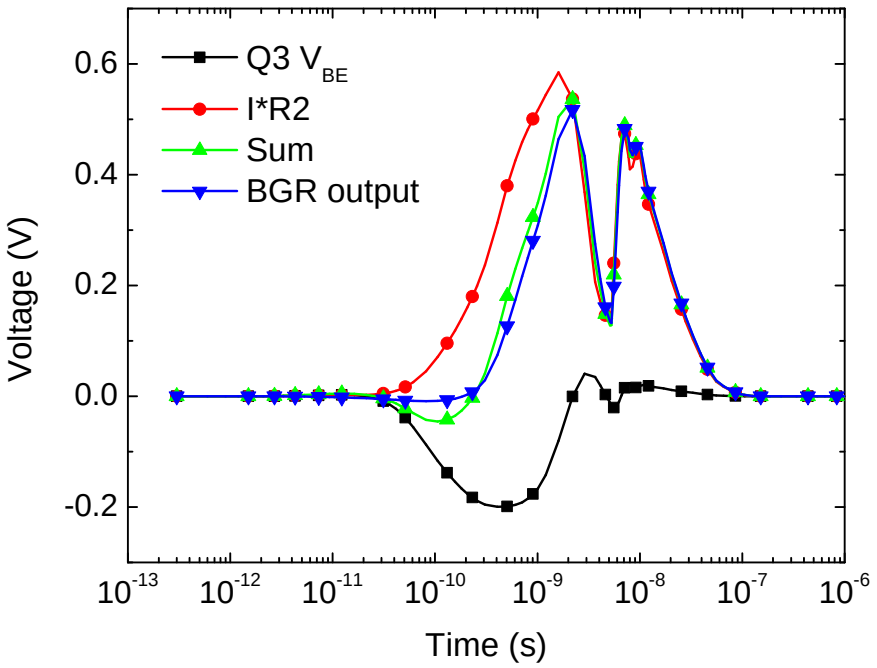

Figure 59: Mixed-mode BGR output transient along with its key components: the resistive drop across R2 and the V_{BE} across Q3. The DC offset of each curve has been removed to highlight the contributions of each component to the overall SET.

output voltage transient is defined primarily by the current component, as the V_{BE} transient is significantly smaller than the voltage drop across R2. The V_{BE} transient corresponds to the shape of the Q3 base current and serves primarily to retard the rising edge of the output transient. In Figure 59, the current and voltage components are summed together to demonstrate a close match to the simulated voltage transient at the output, indicating that it is valid to assume that the current component of the output transient can be represented by current simulated at the top of R2 times the total resistance of R2.

In summary, the initial rising output voltage results from a rising Q5 collector current combined with a subsequent increase in the Q3 base current. The sharp roll-off and subsequent increase in the output voltage is caused by a feedback loop: the roll-off occurs when M11 enters linear operation, reducing the Q4 base current and therefore its emitter current that flows through R2; the decreasing output voltage enables M11 to reenter saturation, resulting in a large signal increase in its drain current, which

causes Q3 and Q4 directly and Q5 indirectly to conduct more current. The output voltage then peaks a second time, followed by decreasing oscillations caused by the Q3 base current and a steady decrease caused by the Q5 collector current.

5.2.3.3 Comparison of Simulation Approaches

Having decomposed the BGR output transient, the origin of the similarities and differences between simulation approaches can be identified. The results show that the analog circuit output is primarily sensitive to the base transient on Q2, unlike traditional digital circuits; as such, although accounting for the actual circuit biases yields the most accurate results, a qualitative match is still achieved by Method #2 since it has a comparable base transient. On the other hand, the reduction in collector and emitter transients shown in Figure 48 cannot explain the difference in the output transients between the mixed-mode simulation and Method #3. The chief deviation between the mixed-mode and current-injection results is that the negative base transient is much larger in the current-injection simulations. Figures 60-61 shows the injected and net transients at Q2 for Methods #1 and #3. The larger base current results in a larger Q5 collector current, thus increasing the base current drawn from Q3 and together causing a much larger initial voltage peak at the BGR output. Subsequently, the larger output transient forces M11 further out of saturation than in the mixed-mode simulation. Since a larger initial Q5 collector transient pulls current from both R2 and the base terminal of Q3, there is only a marginal increase in the base transient of Q3 and its V_{BE}. Consequently, each simulation approach shows a similar reduction in the output voltage as a result of the feedback loop. Since the first output peak is a direct result of current flowing from the Q2, the differences are striking for different base transient magnitudes. However, the second output voltage peak is only affected indirectly by the base transient current, since it is a determined by the feedback loop of Q3, Q4, Q5, and M11. As a result, the current

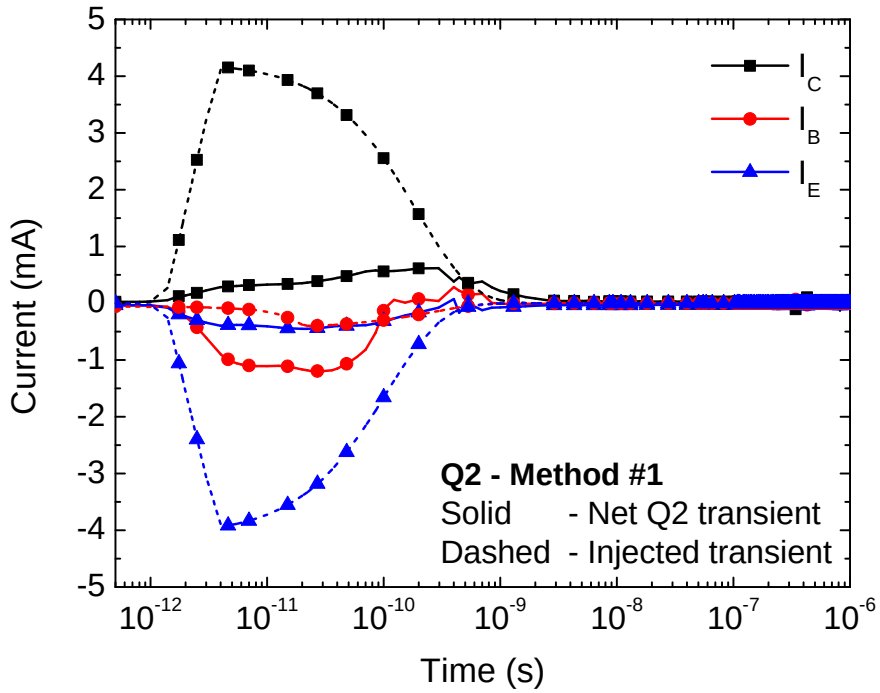

Figure 60: Transient currents at Q2: injected current transients (dashed) and net transient currents seen by the remainder of the circuit (solid), computed according to Method #1.

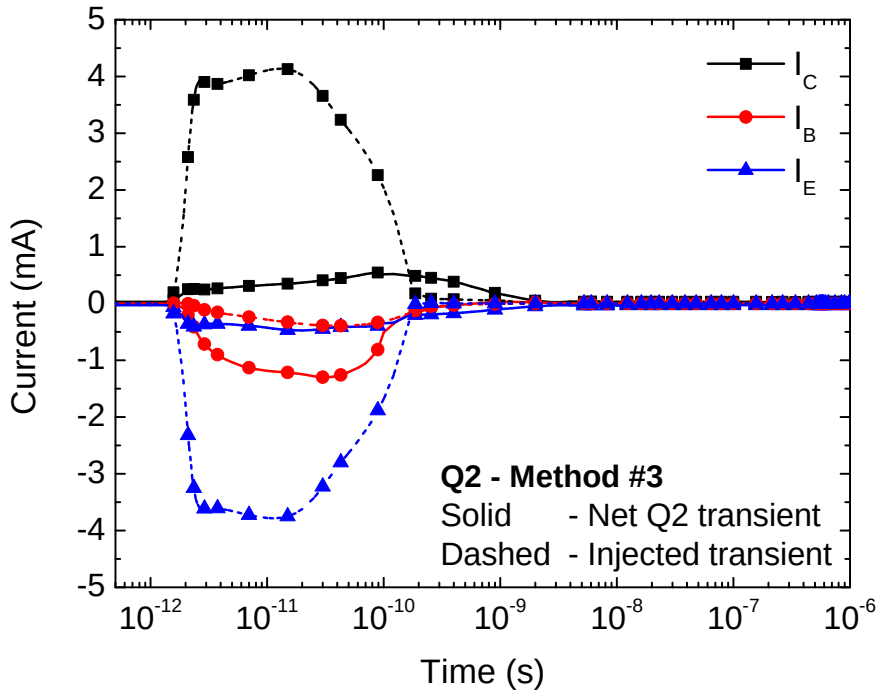

Figure 61: Transient currents at Q2: injected current transients (dashed) and net transient currents seen by the remainder of the circuit (solid), computed according to Method #3.

injection approaches still estimate a larger second peak of the output voltage than the mixed-mode simulation, but the differences are less prominent. Since the base transient of Method #1 subsides before that of Method #3, its simulated output is closer to that of the mixed-mode simulation at the second peak. Nevertheless, all four approaches overestimate the second peak of the measured output voltage. This is possibly due to the fact that some of the underlying 3-D TCAD simulations of an HBT ion strike overestimate the measured SET duration for a single device [107]. Considering Figure 53, if more transient base current is forced into Q5 for a longer period of time, a larger forward bipolar current will be induced, leading to a stronger feedback mechanism and larger BGR output transient. Discrepancies at the device-level will likely be reflected at the circuit-level if they are due to inaccuracies in the physical models. Further work needs to be performed to reconcile all simulation to data discrepancies at the device-level and to assess their impact on circuit-level SET simulations.

5.2.4 Conclusions

Four different approaches to modeling of SET in circuits have been applied to a precision SiGe voltage reference circuit. The importance of modeling the bias dependence of the single device transient was shown by comparing 3-D TCAD transients of a SiGe HBT at negative substrate bias with those of a SiGe HBT mirroring the biases of transistor Q2 from the BGR. The ion shunt effect was identified to be the probable cause of the large collector to emitter transients from Q2. The primary limitation of current injection approaches is that they do not account for the loading of the device terminals that shifts the terminal biases throughout the duration of the SET, as evidenced by the difference between the HBT terminal transients simulated using TCAD alone and those simulated within a full mixed-mode circuit simulation (Figure 48).

With this knowledge of both the device-level response and circuit loading, the

effectiveness of each modeling approach was tested at the circuit level against measured transients in a SiGe voltage reference. Although the mixed-mode simulation is marginally closer to the measured transient, all three current injection approaches also capture the basic shape of the circuit SET. The reasons behind this were illuminated by a detailed analysis of the transient response of the circuit, which enabled the mechanisms that drive the major components of the circuit SET to be traced back to the original device transients at Q2, revealing that the base transient of Q2 is the driving force behind the circuit SET. For this particular BGR circuit, current injection approaches provide similar qualitative insights, since in this case the circuit loading and bias conditions strongly affect the collector and emitter transients, but only marginally affect the base transient that drives the circuit SET. However, this result is entirely dependent on the circuit topology and must be examined for other analog and RF circuits in which loading effects may influence the critical device-level transients (e.g., in dynamic circuits such as voltage controlled oscillators). Loading effects need to be examined at more aggressive scaling nodes (e.g., 130 nm) in which circuit response times are comparable to single-device SET durations. Moreover, for novel device structures such as [82], for which no calibrated compact model presently exists, SET can be realistically modeled only by full mixed-mode simulation.

5.3 Understanding Single-Event Transients in Gb/s SiGe Digital Logic

Significant differences between current-injection and mixed-mode approaches were demonstrated within a SiGe voltage reference circuit in the previous section. Similar discrepancies have also been reported in CMOS logic [106]. In the SiGe voltage reference, differences in the internal voltage and current transients were shown to be caused by feedback from the circuit that influenced the originating device transient as it evolved in time. The feedback mechanism and its impact on the circuit output SET were determined by the particular circuit topology, suggesting that current-injection and mixed-mode approaches may produce greatly differing answers in other circuit types and topologies. Consequently, clear guidelines must be established as to which approaches to modeling circuit-level SET are valid for various conditions, including circuit type and topology, technology node, device geometry, and operating environment. In this section, I apply both current-injection and mixed-mode SET modeling approaches to a standard Gb/s SiGe master/slave D flip-flop (DFF). My goal is to assess, for the first time, the limitations of decoupled TCAD simulations in predicting single-event upset (SEU) in high-speed SiGe digital circuits, and to better understand the fundamental mechanisms and corresponding metrics that are suitable for accurately predicting SEU in high-speed SiGe latches, shift registers, and digital logic.

5.3.1 Circuit and Simulation Details

CFDRC's NanoTCAD tool was used to perform simulations of normally-incident emitter-center ion strikes to off-state SiGe HBTs within the differential pairs of the DFF, since these transistors have been identified as the most sensitive nodes of the DFF [62]. A schematic of the circuit is given in Figure 62. All transistors have the same geometry (0.5 µm x 2.15 µm SiGe HBT). The results presented here include

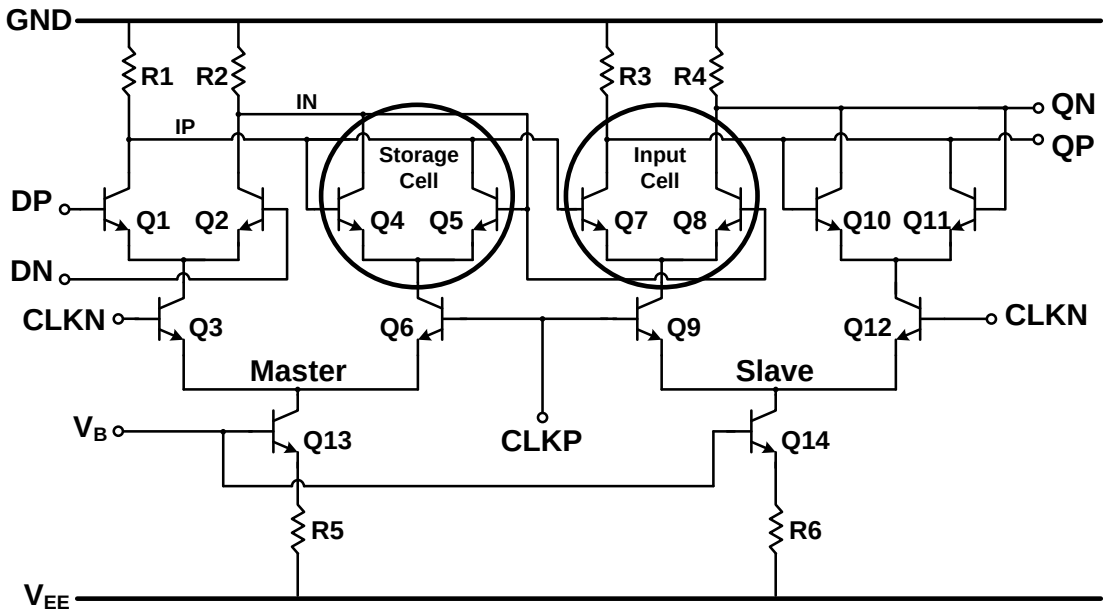

Figure 62: Standard SiGe master/slave D flip-flop schematic.

transient simulations with static input biases as well simulations with dynamic clock biases. For all static-bias simulations, the differential clock and input signals are set such that the master storage and slave input stages are active, Q4/Q7 are in the "on" state, and Q5/Q8 are in the "off" state. The off-state transistor biases in each differential pair are $V_B = -0.25\,V$, $V_E = -0.91\,V$, $V_C = 0\,V$, and $V_S = -5.2\,V$. The on-state collector current is 1 mA. To model the complex ion track, the SRIM software tool was used to compute the energy-deposition vs. depth profile for a 36 MeV oxygen ion, taking into account the back-end-of-line (BEOL) interconnect layers present above Q5 and Q8. This variable linear energy transfer (LET) profile was then imported into NanoTCAD using its automated ion track meshing capability. The LET at the emitter surface is approximately $7\,MeV \cdot cm^2/mg$. The peak of all TCAD ion strikes presented here occurs at 2 ps. The ion-induced current transients computed for fixed biases are shown in Figure 63. At these bias conditions, the large forward shunt current up to about 0.1 ns, followed by diffusive charge collection through the subcollector-substrate junction until roughly 3 ns.

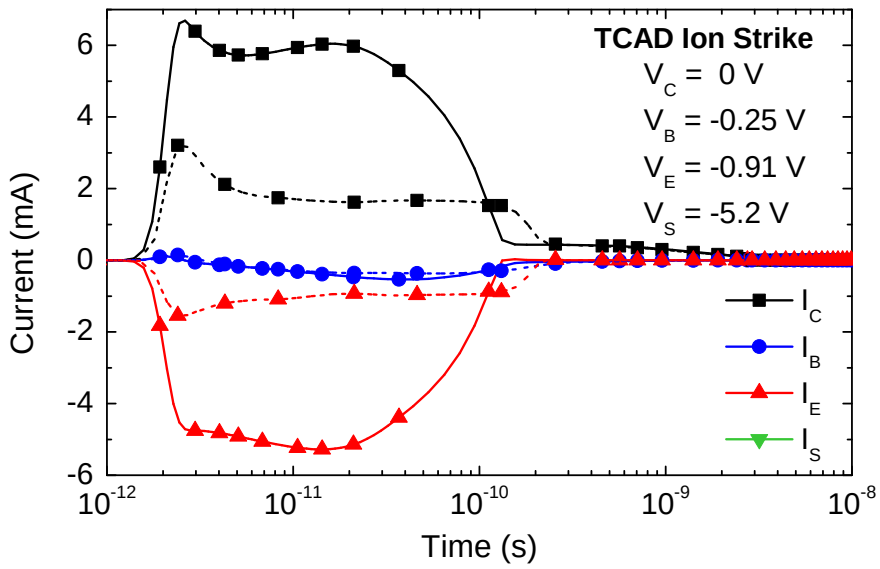

Figure 63: 3-D TCAD current transients for a 36 MeV oxygen ion strike to a standalone SiGe HBT with fixed voltage biases corresponding to an off-state HBT. Dashed lines represent the mixed-mode current transients for the off-state HBT operating within the DFF.

Current-injection simulations were performed by placing piecewise-linear independent current sources at each node of the SiGe HBT that was irradiated, in the same manner as in [76]. The injected current waveforms were taken from 3-D TCAD simulations of the off-state SiGe HBT. Mixed-mode simulations were performed using CFDRC's MixCad tool (3-D NanoTCAD interface to Cadence Spectre) [108], using a calibrated 3-D TCAD model of a first-generation SiGe HBT. Whereas mixed-mode studies such as [112] are limited to basic SPICE passives and compact models, the unique interface between NanoTCAD and Spectre allows this mixed-mode tool to be used directly with the compact models included in commercial process design kits (PDKs). The current-injection simulations require an initial TCAD transient simulation for each transistor and bias case of interest (one of which typically completes within four hours), followed by circuit transient simulations of minimal duration. In contrast, a single mixed-mode transient simulation requires no initial TCAD transient simulation and typically completes within 20-30 hours, depending on the simulation conditions.

5.3.2 Static Bias Simulation Results

In both the DFFs examined here and shift registers employing the same topology, SET can be classified into two categories to simplify analysis: 1) input cell and 2) storage cell ion strikes, since the loading conditions are identical within each group, not only between master and slave stages of each DFF, but also between each element of a full shift register. A strike to any active storage cell interacts with the active input cell of the next stage; likewise, a strike to any active input cell interacts with the active storage cell of the previous stage. To account for both cases, static bias simulations were performed of ion strikes to the off-state transistors of the master storage and slave input cells (Q5 and Q8).

5.3.2.1 Master Storage Cell Ion Strike

The ion-induced voltage transients for a strike to Q5 from both current-injection and mixed-mode simulations are illustrated in Figure 64. An SEU is predicted for both simulations. Moreover, the similarity of the output voltage transients (QP/QN) between simulation approaches suggests that the net effect of circuit feedback on device-level transients is minimal. However, the intermediate node voltages that are set by the Q4/Q5 collector currents exhibit much greater disparity, which can be explained by considering the Q4/Q5 current transients that are shown in Figure 65.

The mixed-mode result shows a straightforward sequence of events. Immediately following the ion strike, the ion-induced currents at Q5 essentially act as forward bipolar current. The increasing $I_{C,Q5}$ pulls down IP and at the same time induces a negative base current on Q4 around 3 ps. Moreover, Q6 limits the total differential pair current, which forces Q4 to cease conducting and prolongs the $I_{E,Q5}$ transient at a lower magnitude than in the device-only TCAD results of Figure 63. IN rises, and by the end of the SET, the opposite logic state is firmly set. The current-injection result, however, shows several notable differences. By forcing a predefined current

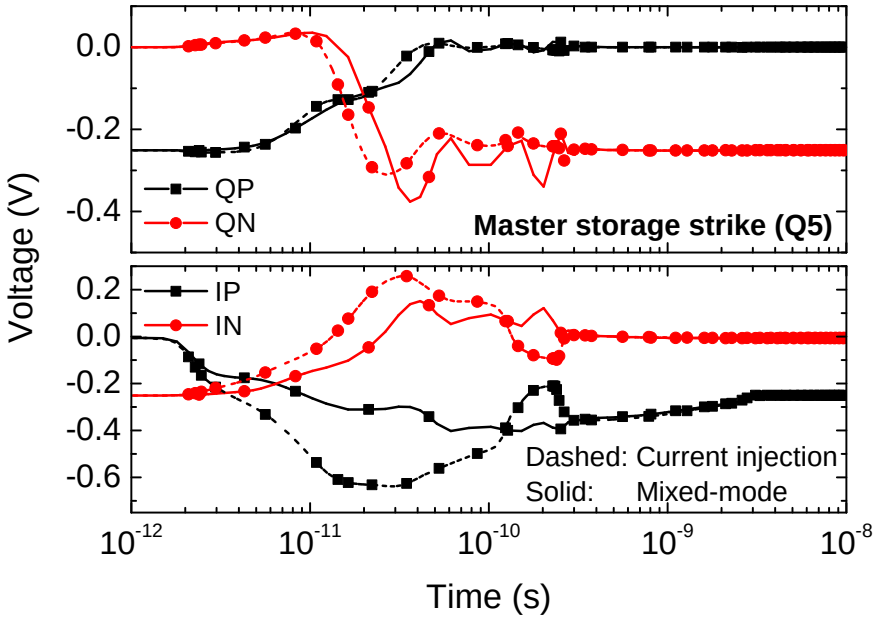

Figure 64: Mixed-mode and current-injection voltage transients for an ion strike to transistor Q5.

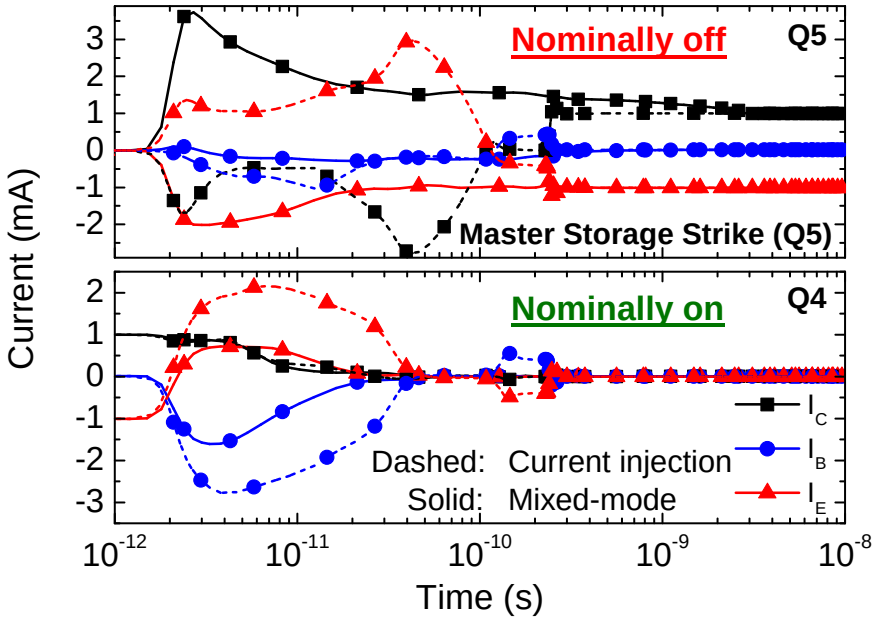

Figure 65: Mixed-mode and current-injection current transients for an ion strike to transistor Q5.

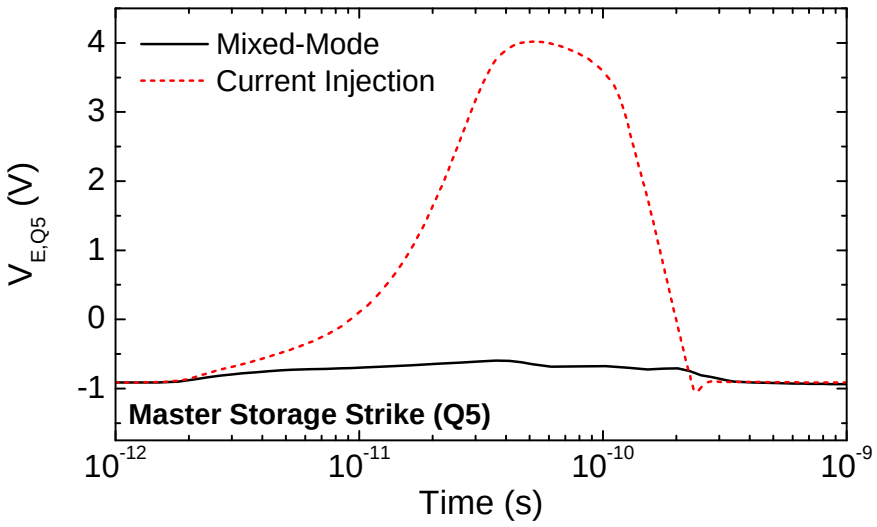

Figure 66: Mixed-mode and current-injection emitter voltage transients for an ion strike to transistor Q5.

into the circuit nodes, the current magnitudes are larger, driving IP down much further in a shorter time period. Moreover, the large emitter transient cannot be supplied entirely by current source Q6, compelling the lower-impedance Q4 emitter-base (EB) junction to become reverse-biased and accumulate charge until 50 ps, as evidenced by the unphysical rise of the emitter voltage to 4 V at 50 ps in Figure 66, and the subsequent discharging base/emitter currents from Q4 after 100 ps. The same capacitive charging/discharging of the EB junction occurs in Q5, which combines with the injected collector transient to induce inverse-mode operation up to 100 ps. Finally, the logic state remains upset at the end of the SET, since Q7/Q8 have already switched states along with all of the node voltages.

To further probe the differences between the two simulation approaches, additional simulations were performed with a range of different LET values for the ion track (from 0.5 to 10 MeV $\cdot$ cm^2/mg), as illustrated in Figure 67. Mixed-mode and current-injection output voltage transients exhibit a high degree of correlation across LET, both showing an LET of 1 MeV $\cdot$ cm^2/mg as the threshold for SEU for strikes to the storage cell with static bias conditions. The fact that SEU occurs before 100 ps in all

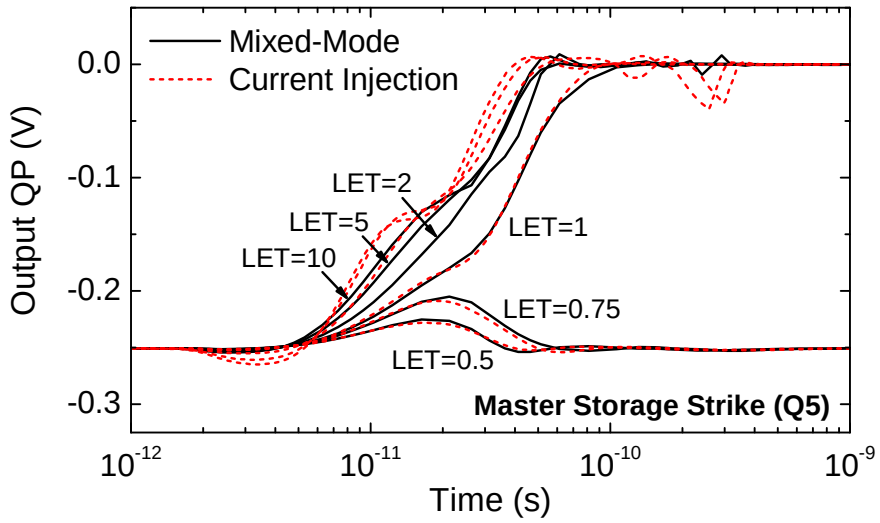

Figure 67: Mixed-mode and current-injection output voltage transients for ion strikes to transistor Q5, with the LET varied from 0.5 to 10 MeV $\cdot$ cm^2/mg.

cases indicates that forward shunt current drives this sensitivity. For a static-bias SEU to occur, the storage cell must be upset, otherwise it will simply reset the following input stage as the SET currents subside. With that in mind, the ion-induced currents at the emitter and collector of Q5 are both needed to cause SEU, since the emitter current takes up an increasing share of the fixed tail current while the collector current pulls down the high-state output and lowers the V_{BE} of the on-state differential pair transistor. Figure 68 shows the emitter currents in the differential pair HBTs for LET values around the SEU threshold, revealing that SEU occurs when the forward shunt emitter current of Q5 exceeds that of Q4. This also explains why the mixed-mode and current-injection results closely agree about the SEU threshold, since it occurs when the forward shunt currents are roughly half of the supply current; for this current magnitude, circuit feedback is minimal.

5.3.2.2 Slave Input Cell Ion Strike

The ion-induced voltage transients for a strike to Q8 from both current-injection and mixed-mode simulations are illustrated in Figure 69. As with the strike to Q5, the current-injection simulation results in an upset. In contrast, the output voltages from

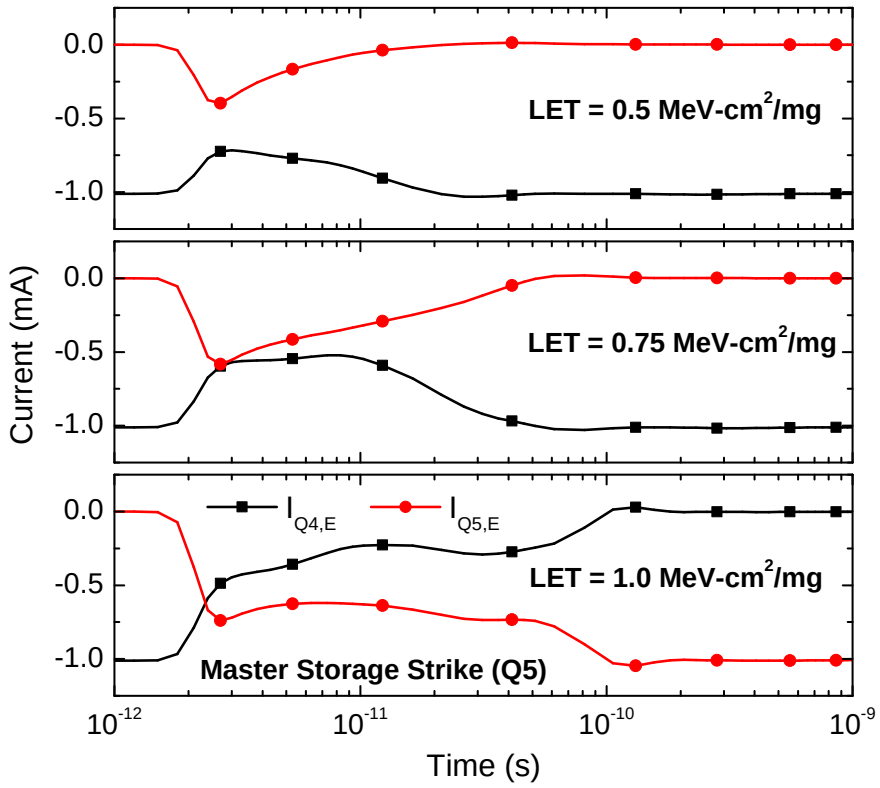

Figure 68: Mixed-mode emitter current transients for ion strikes to transistor Q5, for LET values of 0.5, 0.75, and 1.0 $\mathrm{MeV \cdot cm^2 \cdot mg^{-1}}$.

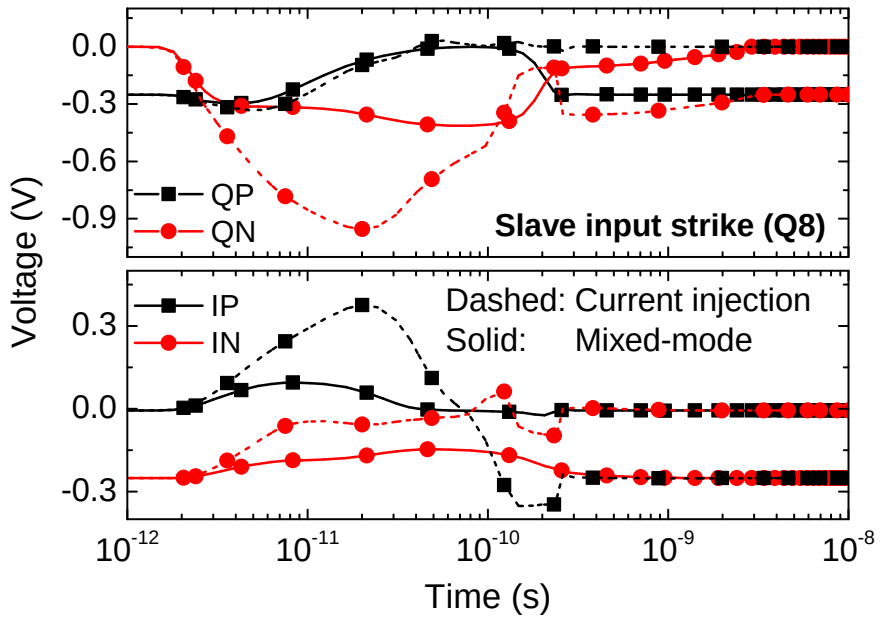

Figure 69: Mixed-mode and current-injection voltage transients for ion strike to transistor Q8.

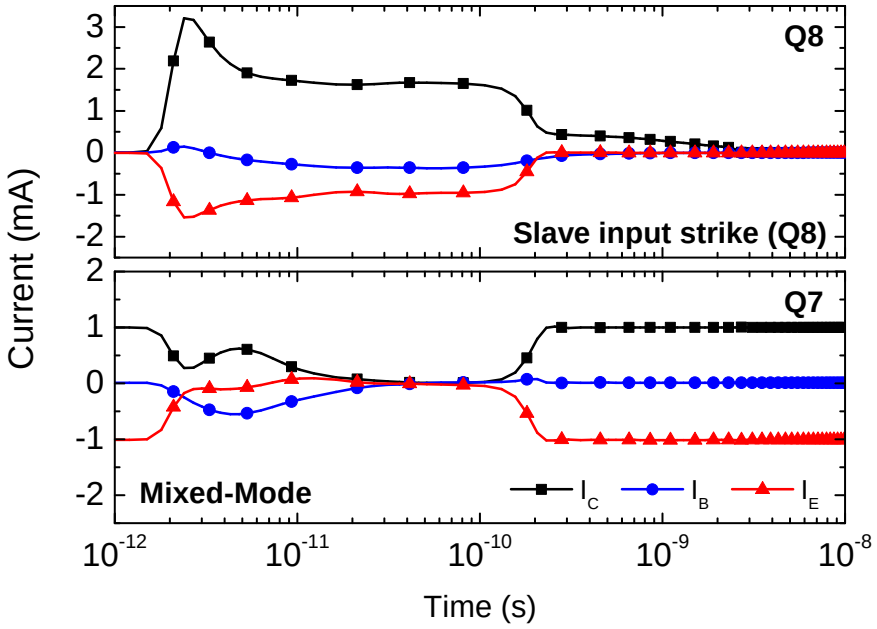

Figure 70: Mixed-mode current transients for Q7/Q8 following an ion strike to transistor Q8.

the mixed-mode simulation recover and do not result in an upset. The differential pair currents again illuminate the mechanisms underlying this discrepancy, with the mixed-mode Q7/Q8 current transients given in Figure 70 and the current-injection Q7/Q8 current transients given in Figure 71. Since the logic state of the master storage cell is critical for causing an upset in the slave input cell at static bias, the current-injection Q4/Q5 current transients are shown in Figure 72. Below 100 ps, the mixed-mode QP/QN transients of Figure 69 correspond closely to those of Figure 64. However, the key difference is that IP/IN never upsets, which ensures that the slave input cell resets by the end of the device-level transients. The slave input cell can affect IP/IN through two paths: the ion-induced negative $I_{B,Q8}$, which causes the moderate increase in IN, and capacitive coupling of a small portion of the Q8 emitter transient through the Q7 EB capacitance, which causes the brief increase in IP. Interestingly, QP returns to its original state at 200 ps, the end of the forward shunt dominated region of the device-level transient, suggesting that the shunting effect plays an important role in SiGe SEU, not just subcollector-substrate charge collection,

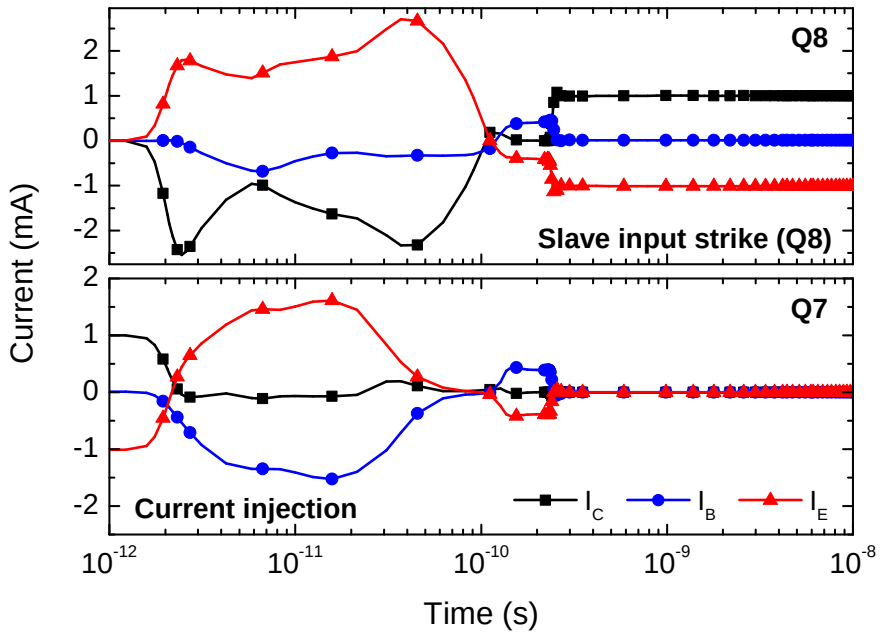

Figure 71: Current-injection current transients for Q7/Q8 following an ion strike to transistor Q8.

as traditionally believed. Nevertheless, after the forward shunt current ceases, QN still remains partially pulled down below its original voltage for the duration of the diffusive collector current. The current-injection QP/QN voltage transients follow comparable trends to those of Figure 64, indicating that similar mechanisms operate within this differential pair when it is irradiated. The Q7/Q8 current transients in Figure 71 closely follow the current-injection transients in Figure 65, demonstrating the same EB charging up to 50 ps that results in an unphysically high emitter voltage, accompanied by inverse-mode operation of Q8 due to the injected collector transient, followed by discharging of the EB capacitance between 130 and 250 ps, and ending with a static SEU at the output. The strike cases differ, however, in the reason why the upset persists following the SET. The master storage cell remains upset because the V_{BE} of Q4 and Q5 are cross-coupled to their collector currents. In contrast, the slave input cell should return to its original state as in the mixed-mode simulation, because the V_{BE} of Q7/Q8 are set by the outputs of the preceding state, unless the current-injection simulation shows that the master storage cell is itself upset by a

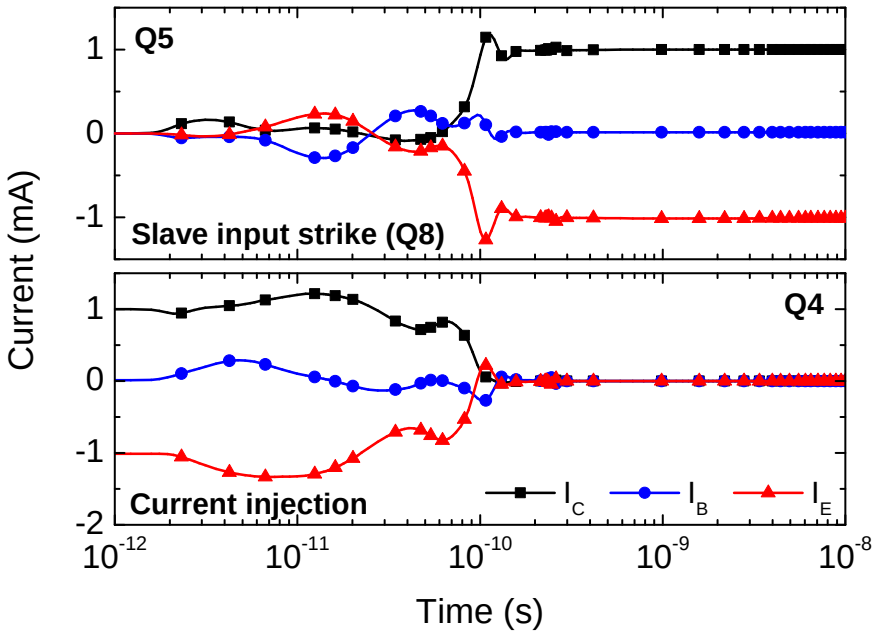

Figure 72: Current-injection current transients for Q4/Q5 following an ion strike to transistor Q8.

strike to the slave input cell. This is indeed the case, as shown by the Q4/Q5 current transients in Figure 72. Current is forced into the base of Q5, inducing forward bipolar conduction and causing the cross-coupled Q4 to cease conducting by 100 ps, matching the point at which IP and IN switch states. By the time the SET ceases in the slave input cell, QP/QN are held in the upset state due to the prior upset of IP/IN.

Additional simulations were performed with a range of different LET values for the ion track (from 0.5 to 10 MeV $\cdot$ cm^2/mg), illustrated in Figs. 73-74. Current-injection results show an LET of 5 MeV $\cdot$ cm^2/mg as the threshold for SEU with static bias conditions. Below the threshold, the base current pushed into Q5 is not sufficient to induce forward-active operation and cause an upset of IP/IN and subsequently QP/QN. Mixed-mode results show that no upset occurs across the entire LET range for a static bias case, yet the deviation of QP/QN for tenths of nanoseconds can result in SEU should a clock edge occur during the window of the output voltage SET. Current-injection simulations not only predict a different threshold LET for SEU in the static bias case, but also predict significantly different output voltage transients

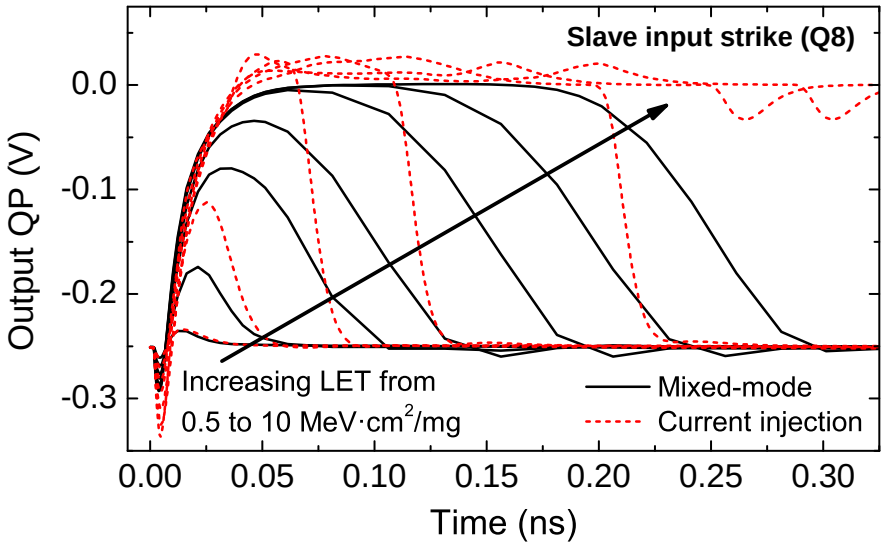

Figure 73: Mixed-mode and current-injection output voltage (QP) transients for ion strikes to transistor Q8, with the LET varied from 0.5 to 10 MeV · cm^2/mg.

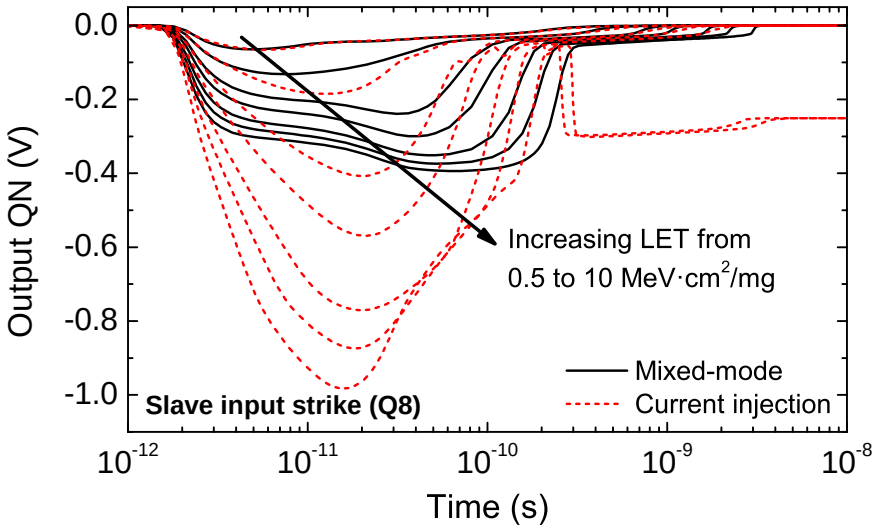

Figure 74: Mixed-mode and current-injection output voltage (QN) transients for ion strikes to transistor Q8, with the LET varied from 0.5 to 10 MeV · cm^2/mg.

below that threshold, highlighting the stark difference in SEU prediction between mixed-mode TCAD and current-injection.

5.3.3 Clocked Simulation Results

Mixed-mode simulations of a clocked DFF were performed to explore the impact of the timing between the ion strike and clock edge. All clocked SET simulations were performed with the variable LET profile computed by SRIM. The mixed-mode voltage transients of Figure 69 show that QP returns to its original state approximately 200 ps after the ion strike. The deviation of QP results from transistor Q7 turning off during the forward shunt emitter transient of Q8; as the $I_{E,Q8}$ transient subsides, Q9 again sources its current through Q7 and QP is pulled down. The deviation of QN, however, is caused directly by the $I_{C,Q8}$ transients. Thus, it reflects the two dominant components of the collector transient: (1) forward shunt current between the emitter and collector, and (2) diffusive current from charge collection at the subcollector-substrate junction. During the forward shunt portion of the transient, QN is pulled down to the full voltage swing of the DFF. When the shunting effect ceases, $I_{E,Q8}$ becomes negligible and the 1 mA design current is again sourced by Q9 through Q7. However, the Q8 subcollector-substrate diffusion current remains superposed on the original circuit operating currents, pulling QN below 0 V for approximately 3 ns. This extended QN transient can be seen across LET in Figure 74.

Clocked simulations were first performed with a single clock transition, where the delay from the ion strike to the clock edge was varied from 0.2 ns, near the end of the forward shunt current, to 2.0 ns. The nominal output voltages are QN = 0 V and QP = −0.25 V. Figure 75 overlays the output QN and CLKN signals for each delay condition, demonstrating that SEU occurs for strikes occurring up to 0.9 ns prior to the clock edge. This indicates that the DFF is sensitive not only during the forward shunt region, but continues to be sensitive when the only remaining transient

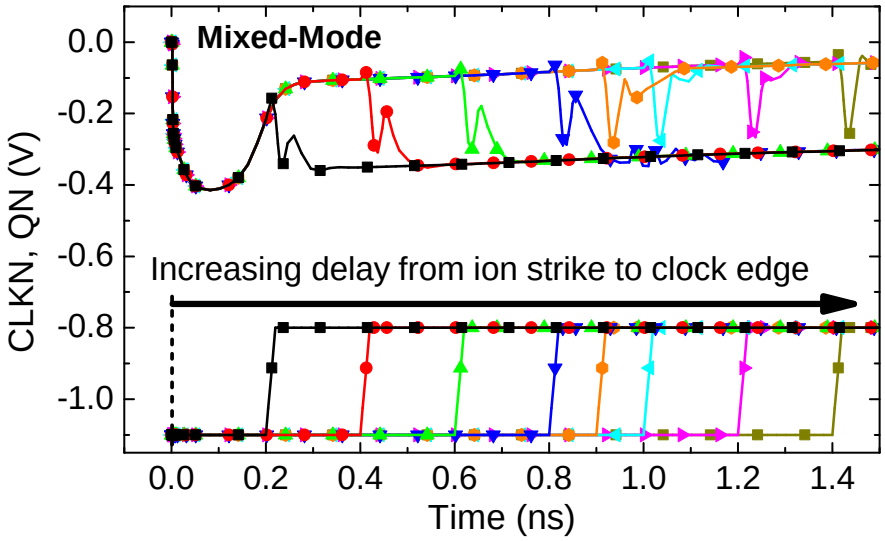

Figure 75: Mixed-mode voltage transients for ion strikes to transistor Q8, with increasing delay to the clock edge. The strike time is held constant at 2 ps.

current is diffusive collector current that partially pulls down QN. Although this result reinforces the motivation for mitigating subcollector-substrate charge collection in SiGe HBTs, it also suggests why transistor-level RHBD approaches have only been partially effective and SiGe SEE sensitivity remains even in SOI technologies where there is no appreciable collector diffusion current. Even without the collector diffusion current, clock transitions during the forward shunt region will result in SEU.

The results of Figure 75 indicate that the SEU sensitivity is not a function of total collector charge collection, but is rather caused by the value of QN at the clock transition. As the delay from the ion strike exceeds 0.9 ns, the total charge collection continues to increase, but SEU ceases to occur. This suggests that the exact value of QN determines whether SEU will occur. With this in mind, it might be surprising that SEU would occur for a designed voltage swing of 250 mV when QP has already returned to its original value of −250 mV and QN is only pulled down to −100 mV. This result can be understood by comparing it to the normal current and voltage transients surrounding a clock transition at 0.4 ns. Figure 76 shows the results for a mixed-mode TCAD simulation of a 2 ps ion strike, whereas Figure 77 shows the

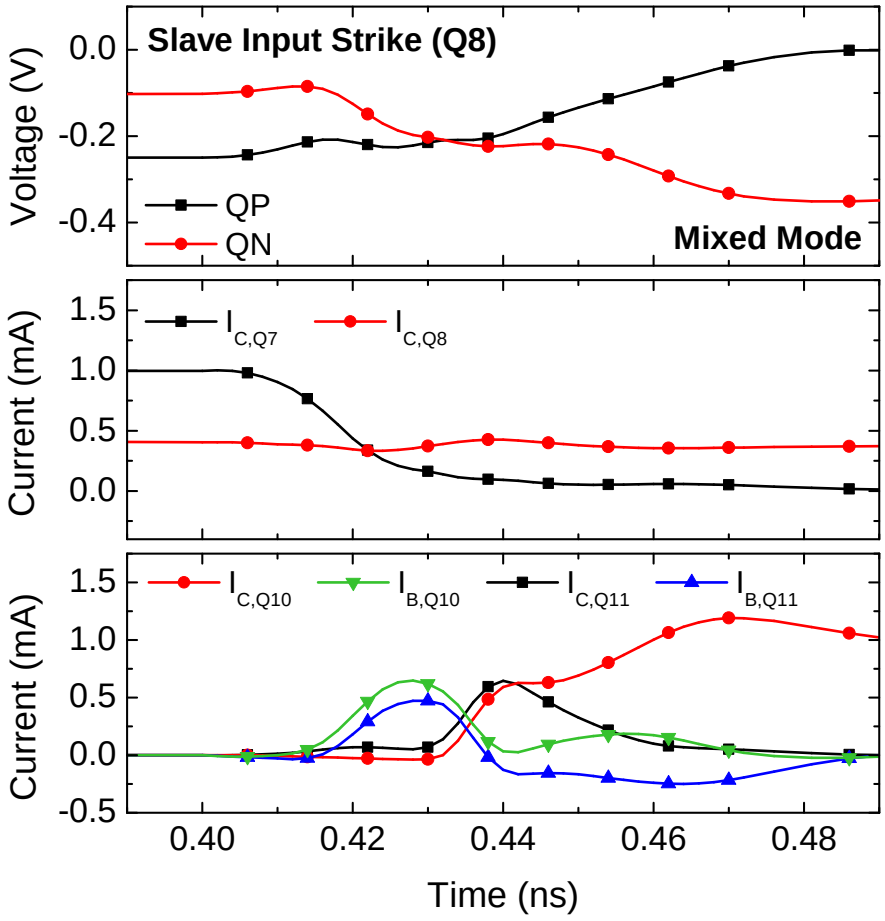

Figure 76: Mixed-mode voltage and current transients for an ion strike to transistor Q8 at 2 ps, with the clock transition occurring at 0.4 ns.

Spectre-computed results for no ion strike.

Prior to the clock transition, the only difference resulting from the ion strike is the non-zero $I_{C,Q8}$ and correspondingly lower QN. At 0.4 ns, CLKP falls and CLKN rises, turning off Q9 and turning on Q12. As Q12 turns on, its collector node voltage drops, inducing a positive base current around 0.42 ns on both Q10 and Q11 as their EB junctions become more forward biased. The base currents flow through QP and QN; since $I_{C,Q7}$ is rapidly decreasing, the net result is a slight increase in QP and moderate decrease in QN for both simulations. Following the initial base transients of Q10/Q11, the two simulations diverge significantly. In the standard case of Figure 77, the V_{BE} of Q11 is much larger than that of Q10 at 0.42 ns. Consequently, as

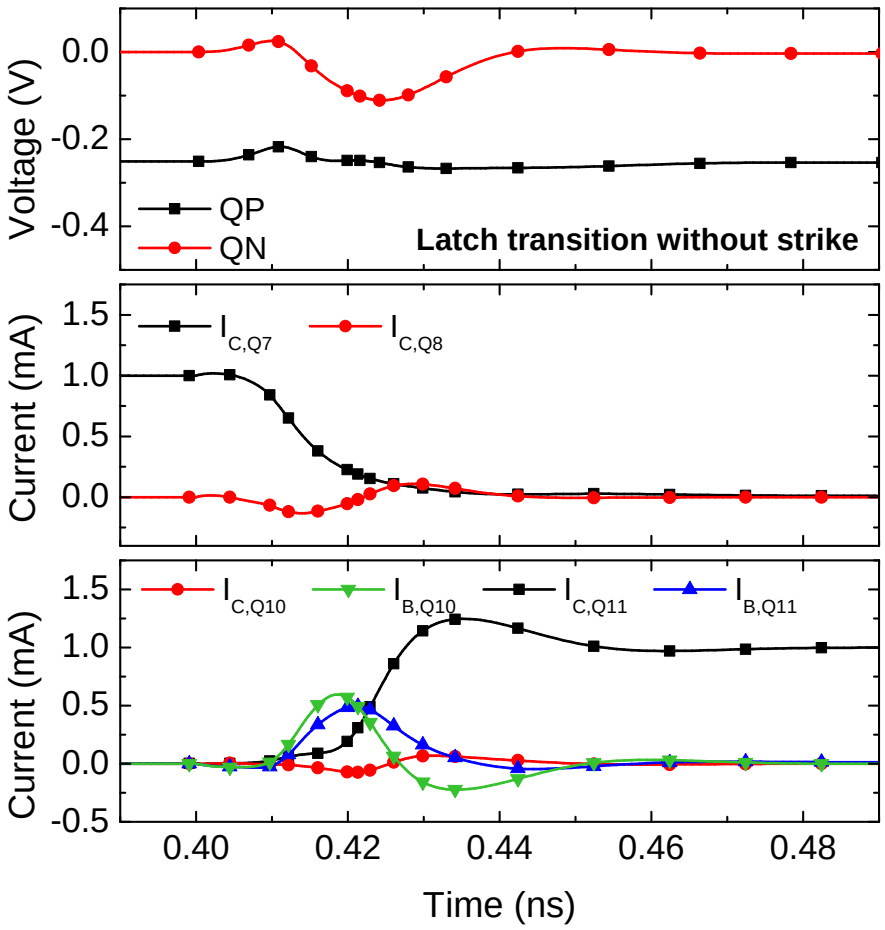

Figure 77: Spectre-computed voltage and current transients for a clock transition occurring at 0.4 ns, with no ion strike.

their shared emitter voltage lowers, Q11 begins to conduct forward bipolar current that is exponentially larger than that of Q10, holding QP constant in the low bias state as Q11 matches the tail current of Q12. As both the emitter voltage and the EB junction charge of Q10 stabilize, QN returns to zero along with $I_{B,Q10}$. In the mixed-mode ion strike case of Figure 76, the key difference is the additional voltage drop across R4 due to the collector SET current. At the peak of the base transients, this additional voltage drop causes QN to approach the level of QP. With comparable V_{BE}, both Q10 and Q11 begin to turn on to match the tail current, evidenced by the parallel rise in their collector currents. However, the voltage drop across R3 at 0.44 ns is caused solely by $I_{C,Q11}$, whereas the voltage drop across R4 is determined by the

summation of $I_{C,Q10}$ and ion-induced $I_{C,Q8}$ transient. Consequently, since there are only two stable operating states (only one of Q10/Q11 can turn on), the additional voltage drop across R4 reduces QN sufficiently to drop below QP during this transition period. Subsequently, Q11 turns back off and $I_{C,Q10}$ proceeds to increase to the full supply current of the DFF. The small negative $I_{B,Q11}$ after 0.44 ns is evidence of Q11 turning completely off as its V_{BE} reduces.

For clock transitions during the diffusion region of the collector SET, the SEU sensitivity is linked to the magnitude of the collector SET current and the output voltages at the moment of the clock transition. As such, it is logical to expect that the SEU sensitivity will change if the operating conditions of the circuit are varied (voltage swing, bias current, etc.). Previously, increasing the bias current has been shown to reduce SEU sensitivity [50], assuming that the device-level SET magnitude remains constant. However, the SEU sensitivity during the forward shunt region will not be affected by the same changes, since the large forward shunt current directly forces the input cell to assume the opposite state, whereas the collector SET diffusion current is simply superposed onto the correct input cell state. Thus, two different circuit SEU mechanisms accompany the two distinct regions of the device-level SET. This has strong implications for SEU-hardening approaches; existing approaches that strictly target subcollector-substrate charge collection cannot mitigate SEU due to forward shunt current. Further analysis and experimental testing is needed to assess the statistical significance of each SEU mechanism, so that more effective SEU-hardening techniques can be developed. Mixed-mode TCAD is an essential tool for this area of research, since decoupled circuit SET modeling cannot capture the essential physics of the SEU mechanisms. This is highlighted by the fact that above a moderate LET of $5\,\mathrm{MeV \cdot cm^2/mg}$, decoupled simulations predict SEU to occur at all times in the circuit examined here, even without an active clock.

5.3.4 Conclusions

The simulation results demonstrate that strikes to the off-state storage cell SiGe HBT of a DFF can cause SEU even with static biases. This static-bias sensitivity is attributed to the forward shunt region of the device SET. Moreover, fundamental limitations of conventional decoupled circuit SET simulations lead to inaccurate and, at times, unphysical transient behavior when the circuit response influences the device-level transient. With static biases, decoupled simulations erroneously predict SEU for ion strikes to the DFF input cell, in contrast to the mixed-mode TCAD results. Mixed-mode TCAD simulations reveal that strikes to the off-state input cell SiGe HBT cause SEU in two ways: (1) large sub-nanosecond current shunts from emitter to collector and directly causes the input cell to upset, resulting in SEU when coincident with clock transitions, and (2) diffusive charge collection at the subcollector-substrate junction causes an additional voltage drop that pulls down the high-state output voltage, which if large enough can cause SEU during clock transitions. These factors will help to explain actual SEU data for various RHBD approaches that have been developed for Gb/s SiGe logic.

5.4 Best Practices in Circuit-Level SEE Simulation

Despite the intrinsic advantage in accuracy that coupled mixed-mode TCAD holds over decoupled modeling approaches, there are several important considerations of which a designer needs to be aware of when modeling SEE. An ideal SEE modeling tool at the circuit and system level would minimize or avoid any changes to the existing circuit netlists, minimize simulation runtime, maintain fidelity to measured results, and provide accurate predictions for SEE behavior in circuits for which no measured data has been yet obtained. However, since no single tool can accomplish all of these goals, the best tool in a given situation depends entirely on the designers goals and intended use.

As with any simulation and modeling application, the basic trade-off in SEE modeling is between computation time and accuracy. At higher levels of abstraction (e.g., system level design), some error margin may be acceptable if it keeps the simulation time reasonably short, since the primary purpose is to validate the overall sensitivity of the system. However, within individual circuit blocks, an accurate representation of the underlying physical mechanisms is critically important, because it informs the designers who develop and employ RHBD techniques within the lowest levels of the system design.

Since mixed-mode TCAD includes physical simulations of a 2-D or 3-D semiconductor device, it is quite computationally complex. Although the runtime depends heavily on the specific software, complexity of the TCAD device model, and simulation settings, the general runtime for a mixed-mode TCAD SET simulation can be expected to be anywhere from a few hours to a few days for a standard analog or digital circuit block. Multithreading can reduce this time considerably, but multithreading support varies from tool to tool. Also, since convergence of the compact models is relatively quick, the overall mixed-mode TCAD runtime is primarily determined by convergence of the TCAD device solver and is not significantly affected by the size of

the surrounding compact model circuit. Compact model approaches perform much faster than mixed-mode TCAD (e.g., a few seconds to a few minutes for a standard analog or digital circuit block), but they typically require much more preparation. As shown in the literature, the most accurate compact model approaches typically use TCAD simulations as the basis for the device-level transients, and often calibrate the SEE model parameters using mixed-mode TCAD simulations. Even though the end results produced using SEE compact models can be computed in a fraction of the time, there is a significant one-time cost for the initial model development. The advantage, however, is that compact model approaches are typically more scalable and well-suited for iterative simulations, albeit applicable primarily to just one technology and type of design (e.g., SOI CMOS shift registers), since the SEE compact model is calibrated using TCAD results and/or data from a specific circuit. In contrast, the advantage of mixed-mode TCAD approaches is that the initial simulation setup is relatively simple, enabling a quick analysis of SEE in new circuit designs and technologies.

Other considerations linked to the trade-off of time vs. accuracy include the type of information the tool can provide as well as the ease with which a modeling tool can be integrated within the standard circuit design tool flow. At the system level, the SEE model should be integrated in a way that is transparent to the designer, which is a reasonable goal, since in this case a tool only needs to provide overall SEE rates. One important concern is to avoid modifying netlists and schematics that are considered golden [33]. Yet at the same time, device and circuit RHBD development necessitates accurate predictions of SEE behavior based on the actual physical mechanisms, and physics-based TCAD modeling tools tend to be much more disruptive of the standard design tool flow. In fact, this often requires the designer to convert existing circuit netlists over to a custom format used by the TCAD software, a process which can be complicated by the fact that some TCAD tools do not support the compact models used by the PDK. Nevertheless, creative implementations of

mixed-mode TCAD, such as CFDRC's MixCad tool [10], demonstrate the ability to integrate a physics-based TCAD solver with standard circuit simulation software (Cadence Spectre) and commercial PDKs. In this case, existing circuit netlists can be left unchanged with all of the PDK compact models, and the designer can swap out solely the devices of interest with physical TCAD models.

The most natural application of mixed-mode TCAD is for fundamental research and development of new devices and technologies. When scaling the physical device structure and doping profiles for a new technology node, the easy path is to continue to use the same modeling tools. However, the fast and easy solution is pointless if it gives incorrect or unphysical results. Thus, in the early stages of technology development, it is vital that modeling tools be tested to ensure that all underlying assumptions remain valid. Numerous examples of the shortcomings of conventional decoupled modeling techniques have been reported over the last decade, providing direct comparisons to data and mixed-mode TCAD simulations [33, 63, 65, 106, 68].

Ultimately, the considerations for SEE modeling are much the same as for modeling of normal device and circuit operation. Compact models are created because they provide fast, efficient simulations of circuit and system performance for designers, whereas TCAD is prohibitively slow, especially at the system level. Compact modeling of SEE can no longer be done in quite the same way, because of the strong dependencies on circuit type, topology, and the unique interactions with device-level effects that change with each technology and node. Although SEE compact models are necessary for system-level simulations, a deep knowledge of device-level SET mechanisms is necessary before a valid SEE compact model can be created. In traditional IC technology development, measured data can be sufficient to inform the compact model developers, but since SEE experiments are much more costly and expensive than simple DC and RF characterization, mixed-mode TCAD tools can play a vital role in reducing the number of experiments while evaluating the underlying physical SEE

mechanisms in a new technology. The results discussed in this chapter demonstrate the ability of mixed-mode TCAD to uncover these mechanisms. Moreover, mixed-mode TCAD can be an efficient tool to evaluate device-level RHBD approaches in which the device structure, doping, and/or layout are modified, rendering existing compact models insufficient. In the big picture, what is most important is to ensure that any simulation tool a designer chooses has a foundation that traces back to a fundamentally physical understanding of SEE in that particular technology and environment.

CHAPTER VI

PHYSICS-BASED TCAD FRAMEWORK FOR PREDICTIVE MODELING OF BIPOLAR CIRCUIT RELIABILITY

As technologies evolve to meet the demands of modern broadband applications, operating voltages inevitably shrink, and circuit designers are compelled to operate devices closer and closer to the classical "safe operating area" (SOA) boundaries that define the maximum voltage and current levels that provide robust and stable operation. Defining these SOA boundaries is quite problematic, since the underlying physics of the various damage mechanisms is extraordinarily complex, and conventional SOA definitions based on DC measurements do not necessarily reflect the actual SOA for devices within mixed-signal circuits [15, 17, 22, 36, 95].

In addressing circuit reliability associated with transient radiation effects, I have improved and validated physics-based TCAD models of ion-strike charge collection, then embedded these models within a compact model circuit environment to assess the impact on circuit operation. Similarly, physics-based TCAD models can be used to model device operation at or beyond classical SOA limits, which gives rise to complexities such as pinch-in instabilities, thermal effects, current-dependent breakdown phenomena, complex 3-D effects, and stress-induced leakage. Whereas circuit simulations with compact models alone simply cannot account for these issues, TCAD simulations themselves cannot adequately predict the impact on circuit operation based only on single device simulations. Consequently, my objective here is to apply the knowledge gained from mixed-mode simulations of ion-induced transients to other reliability issues facing advanced SiGe HBTs.

Since this is a complex problem with many applications, the scope of this task will be limited to proving the feasibility of using mixed-mode TCAD by developing calibrated physics-based degradation models that are compatible with mixed-mode TCAD simulations of mixed-signal circuits.

6.1 Physics-Based Trap Degradation Model

The mixed-mode degradation mechanism in SiGe HBTs occurs when a high collector current density and large collector-base reverse-bias are simultaneously applied. The basic process is illustrated in Figure 78 and proceeds as follows. First, minority carriers traverse the base and enter the collector-base junction space-charge region. For sufficiently high electric fields, impact-ionization will occur, resulting in additional electron-hole pairs. These generated electrons and holes will drift toward the base and collector regions based on the polarity of the electric field, and can potentially participate in further impact-ionization. Some of these energetic carriers can then travel to the various oxide interfaces within the device, arriving with sufficient energy to create traps. In the SiGe HBT, traps at the emitter-base (EB) spacer oxide interface can cause a large increase in non-ideal forward-mode base current due to Shockley-Read-Hall recombination. Traps at the shallow-trench isolation (STI) can likewise be revealed by an increase in non-ideal base current when biasing the transistor in inverse-mode operation. For typical circuits, the transistors are biased in the forward mode and EB spacer traps are of greatest concern, as they reduce the current gain and can shift the operating point of a circuit outside of its functional range.

There are two distinct components of the mixed-mode degradation mechanism: (1) the process by which hot carriers are generated and travel to an oxide interface, and (2) the dynamic process of trap formation and annealing at the interface. The first component can be modeled following the same approach as the lucky-electron model that has been widely used for hot-carrier injection in MOSFETs [38, 41]. The

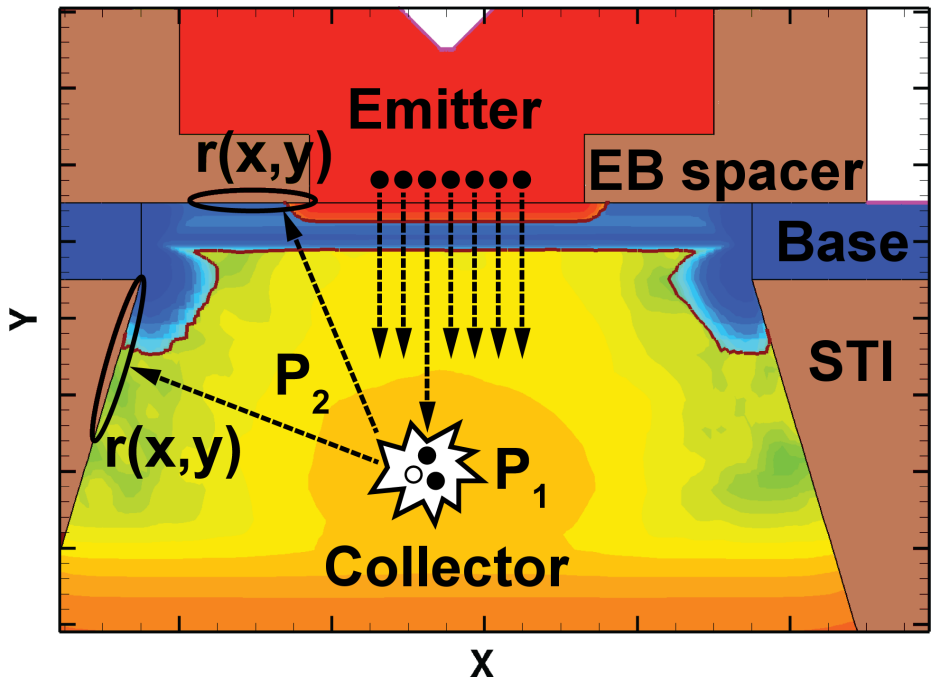

Figure 78: Cross-section of 2-D device model used for transient degradation simulations, annotated with the basic process behind the mixed-mode degradation mechanism.

lucky-electron model uses a probabilistic approach to compute the total rate of hot carriers that satisfy the necessary conditions to be injected from the MOSFET channel region across the gate oxide and into the gate contact. To apply the same approach here, the individual probabilities are adjusted to instead compute the rate of carriers that become highly energetic and are directed all of the way to an oxide interface while retaining sufficient energy to cause damage. The second component of the mixed-mode degradation mechanism can be modeled using the well-known reaction-diffusion process, which was first applied in the context of transistor degradation by Jeppson et. al [42] to explain bias temperature instability (BTI) in MOSFETs as a hydrogen-diffusion controlled interface state creation mechanism. The same fundamental approach applies here, with the major differences being the oxide thickness and geometry. What links the two components of mixed-mode degradation is the fact that trap formation at the oxide interface is determined by the rate of hot carriers reaching the interface.

For a device model comprised of a 2-D discretized mesh, the forward trap formation rate at an oxide interface can be defined as

$$K_F(x_0, y_0) = \sum_{(x,y)\in V} r_e(x, y, x_0, y_0) + \sum_{(x,y)\in V} r_h(x, y, x_0, y_0), \tag{10}$$

where (x_0, y_0) is an interface vertex along the oxide interface and $r_e(x, y, x_0, y_0)$ and $r_h(x, y, x_0, y_0)$ are the impinging hot electron and hot hole rates, respectively, as a function of origin within the semiconductor. These rates are summed across the entire semiconductor volume to give the net rates of hot electrons and hot holes that reach the given position along the oxide interface.

The hot carrier rates are defined as a function of vertex position within the semiconductor to be

$$r_{e/h}(x, y, x_0, y_0) = \frac{||\vec{J}_{n/p}(x, y)||}{q} P_{1,e/h}(x, y) P_2(x, y, x_0, y_0) M(x, y), \tag{11}$$

where $||\vec{J}_{n/p}(x, y)||$ is the local electron or hole current density, $P_{1,e/h}$ is the local probability that a carrier will gain at least a threshold energy of ϕ_{hot} and will be directed towards the interface vertex at (x_0, y_0) with sufficient momentum to create a trap, $P_2(x, y)$ is the local probability that a hot carrier will traverse the entire distance between (x, y) and (x_0, y_0) without suffering a momentum-robbing collision, and $M(x, y)$ is the measure of the local vertex. The measure is defined to be the equivalent 2-D area of that vertex and is needed to weight the hot carrier rate at each origin vertex based on the local density of the mesh.

The hot carrier redirection probability can be defined after [38] as

$$P_{1,e/h}(x, y) = \int_{\phi_{hot}}^{\infty} P_{hot,e/h}(\epsilon, x, y) P_{red}(\epsilon), \tag{12}$$

where $P_{hot,e/h}(\epsilon, x, y)$ is the probability that a carrier gains energy ϵ, defined as

$$P_{hot,e/h}(\epsilon, x, y) = \frac{1}{\lambda F_{eff,e/h}(x, y)} e^{\epsilon/\lambda F_{eff,e/h}(x,y)} d\epsilon, \tag{13}$$

and $P_{red}(\epsilon)$ is the probability per unit length that that same carrier will undergo a redirecting collision such that it has sufficient momentum in the direction toward the oxide interface, with

$$P_{red}(\epsilon) = \frac{1}{2\lambda_r}\left(1 - \sqrt{\frac{\phi_{hot}}{\epsilon}}\right). \tag{14}$$

$F_{eff,e/h}$ is the effective electric field experienced by electrons and holes, respectively, and can be defined in several ways, either as the component of the electric field in the direction of the current flow, or more realistically based on the local carrier temperature, although this second option requires use of the hydrodynamic transport model. λ is the mean free path between collisions for a carrier in Si, λ_r is the mean free path between redirecting collisions, and ϕ_{hot} is the threshold energy required to break a silicon dangling bond at the interface, depassivating an interface trap (2.3 eV [27]).

The probability that a hot carrier will reach the oxide interface without a collision is given as

$$P_2(x, y, x_0, y_0) = e^{-d/\lambda}, \tag{15}$$

where d is the distance between (x, y) and (x_0, y_0).

Once the net hot carrier rates have been accumulated at a given interface vertex, the interface trap density can be computed following the classical reaction-diffusion equation:

$$\frac{\partial N_{it}}{\partial t} = K_F(N_0 - N_{it}) - K_R N_{it} H_2, \tag{16}$$

where K_F is the trap depassivation rate constant, N_0 is the initial concentration of dangling bonds (10^{13}–$10^{14}\,\mathrm{cm}^{-2}$ [101]), N_{it} is the surface density of unpassivated dangling bonds, K_R is the trap passivation rate constant, and H_2 is the surface density of hydrogen. In the context of this problem, several assumptions are made. The majority of silicon dangling bonds at the oxide interface (P_B centers, denoted as $^0\mathrm{Si}{\equiv}\mathrm{Si}_3$) are passivated by atomic hydrogen to result in a diamagnetic neutral defect [101]. Under this theory, when hot carriers strike the interface, the P_B centers are

depassivated and a hydrogen atom is liberated. The free hydrogen atoms can then proceed to repassivate a P_B center, or they can bond together to form molecular hydrogen, H_2. If they remain at the interface, hydrogen molecules will continue to participate in the repassivation process, but migration of hydrogen away from the interface will result in a net increase in unpassivated traps. Hydrogen cannot migrate into Si, but will slowly diffuse into the oxide away from the interface. The passivation reaction is assumed to occur instantaneously, such that the reaction is diffusion-controlled rather than reaction-rate-controlled. If the reaction is far from saturation ($N_{it} << N_0$) and the oxide is thick ($W^2/4Dt > 1$), then the solution for the interface trap density during hot carrier degradation can be approximated as

$$N_{it}(t) \approx 1.16\sqrt{\frac{K_F N_0}{K_R}}(Dt)^{\alpha}, \tag{17}$$

where D is the diffusion coefficient of hydrogen in the oxide ($0.01\,\mu m^2/s$ [16]) and α sets the power-law time dependence of trap formation, fixed here at 0.25 [42]. For these simulations, K_R is held constant at $10^{-7}\,s^{-1}$ [16]. The equation above will not hold during the annealing process that occurs when the hot carrier rate becomes negligible and hydrogen molecules diffuse back to the interface to passivate the P_B centers.

6.2 Transient Degradation Simulation Methodology

The model described in the previous section was implemented for use in the Synopsys Sentaurus TCAD suite, using Sentaurus Device (sdevice) for 2-D electrical device simulations and Sentaurus Workbench to parameterize the model and manage the simulations. The model was built in C++ and integrated into transient device simulations using the sdevice physical model interface (PMI). Prior to the degradation simulation, the HBT is ramped to a specified stress bias condition ($J_{E,stress}$ + $V_{CB,stress}$) in the common-base configuration. Next, the transient degradation simulation is performed following the basic simulation flow in Figure 79. The dynamic interface

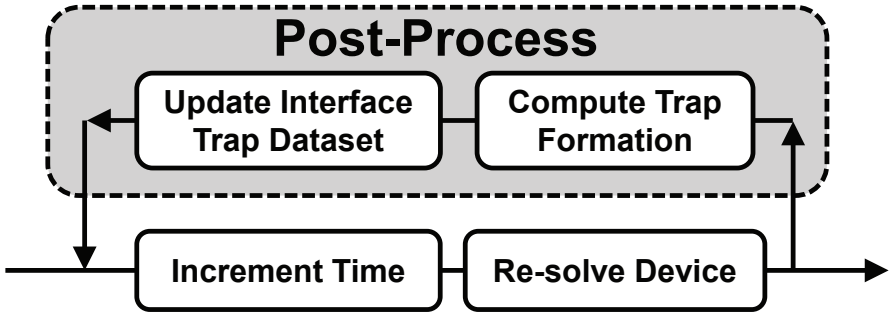

Figure 79: Simulation flow during transient simulation with PMI degradation model enabled.

trap densities are computed in a post-process step by the PMI degradation model and are then coupled back into the device simulation prior to each new time step using a separate PMI model. At specified stress durations, the device state is saved for further analysis, similar to how in stress measurements, the Gummel characteristics are typically measured at specified stress intervals to capture the change in current gain and base current. In this case, following the transient stress, the post-stress forward and inverse Gummel characteristics are simulated at the saved stress states and a change in base current vs. stress time is extracted.

Although stress measurements provide the final word on device reliability, TCAD simulations can provide a unique perspective on the internal workings of the degradation mechanisms. Here, the degradation process can be dissected into its separate components, aiding in the analysis of which specific aspects of the device profile drive the overall degradation of device FoM. To demonstrate the degradation process as it is captured by this model, an example stress simulation is performed using a calibrated 2-D TCAD model of a 50 GHz peak-f_T npn SiGe HBT, the cross-section of which is given in Figure 78. The stress conditions for this example are $J_E = -1\,\mathrm{mA}/\mu\mathrm{m}^2$ and $V_{CB} = 8\,\mathrm{V}$, with the results shown in Figures 80-89 taken after a stress duration of 1000 s. The first element of the hot carrier rate is the redirection of hot carriers with sufficient momentum to create interface traps. The probability of this occurring is dependent on the mean free paths and is a function of the local effective electric

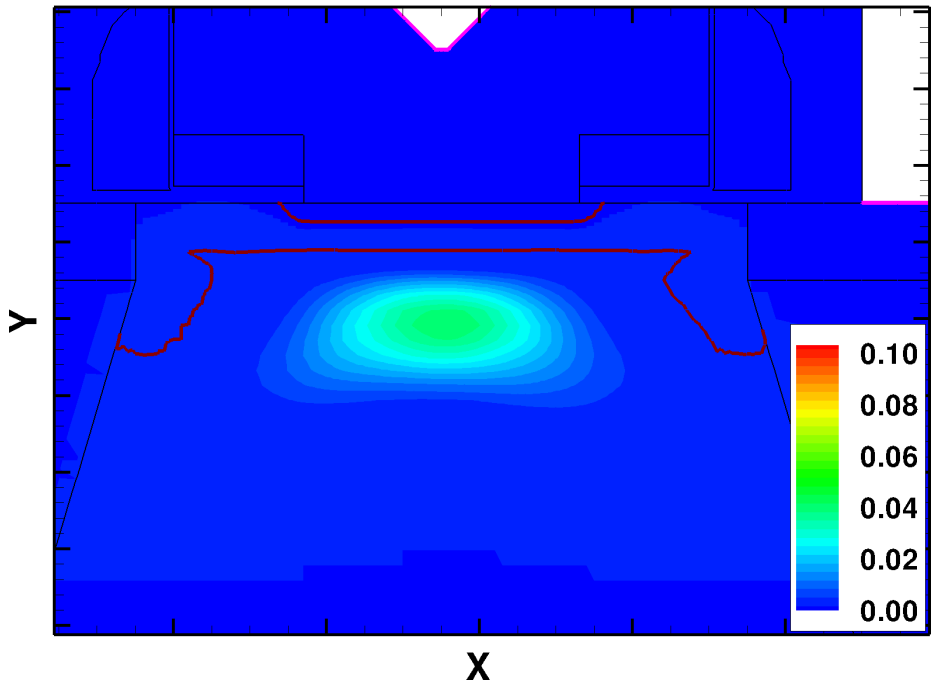

Figure 80: Hot electron redirection probability at $t = 1000s$, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

field, which differs for electrons and holes. In this simulation, hydrodynamic carrier transport models were enabled and the effective electric field was computed from the individual carrier temperatures. Figures 80-81 show the hot electron and hot hole redirection probabilities, respectively, as a function of position within the semiconductor. The greatest probability for hot carriers to occur lies as expected within the large electric field of the collector-base space-charge region. The differences between the electron and hole probabilities follow those of their respective effective electric fields and derive directly from the differences between the respective carrier temperatures.

The second component of the hot carrier rate is the probability that a redirected hot carrier will travel all of the way to the oxide without suffering a energy-robbing collision. This probability is dependent on the mean free path and is otherwise only a function of the hot carrier origin and the interface position. Figure 82 shows the 2-D probability distribution as a function of hot carrier origin for the interface vertex nearest the emitter window.

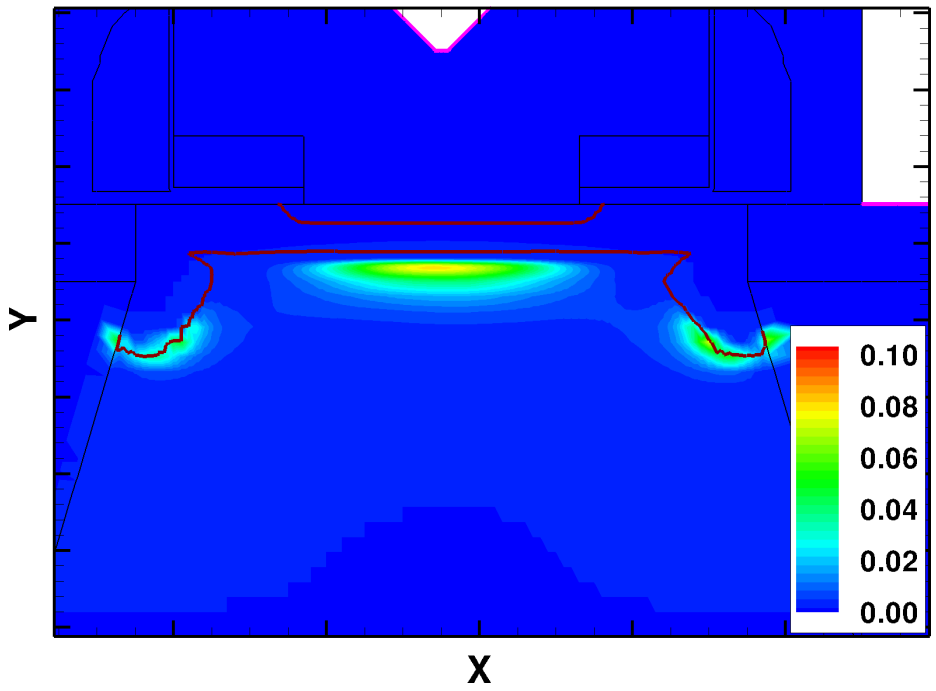

Figure 81: Hot hole redirection probability at $t = 1000s$, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

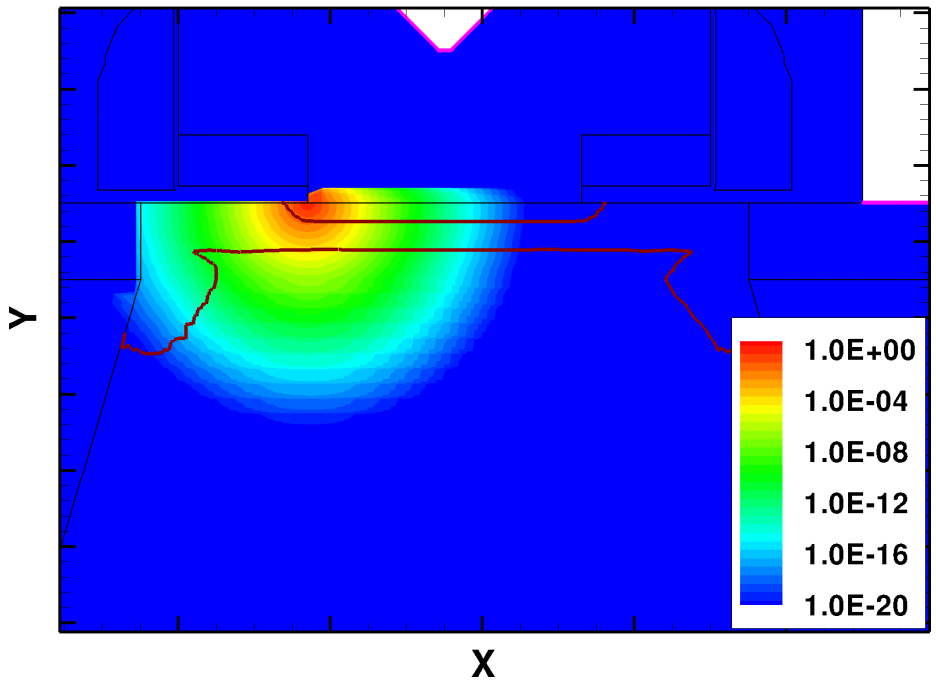

Figure 82: P_2 as a function of position through the semiconductor volume, computed for the EB spacer oxide interface vertex nearest the emitter window.

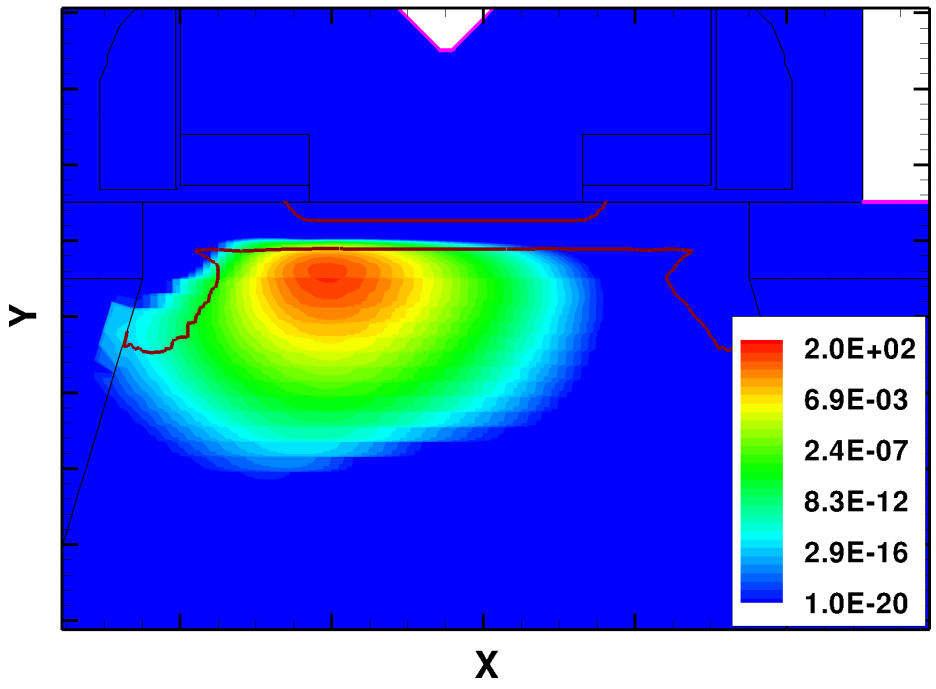

Figure 83: Rate of hot electrons reaching the left EB spacer oxide interface vertex nearest the emitter window at $t = 1000s$, as a function of the hot electron origin, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

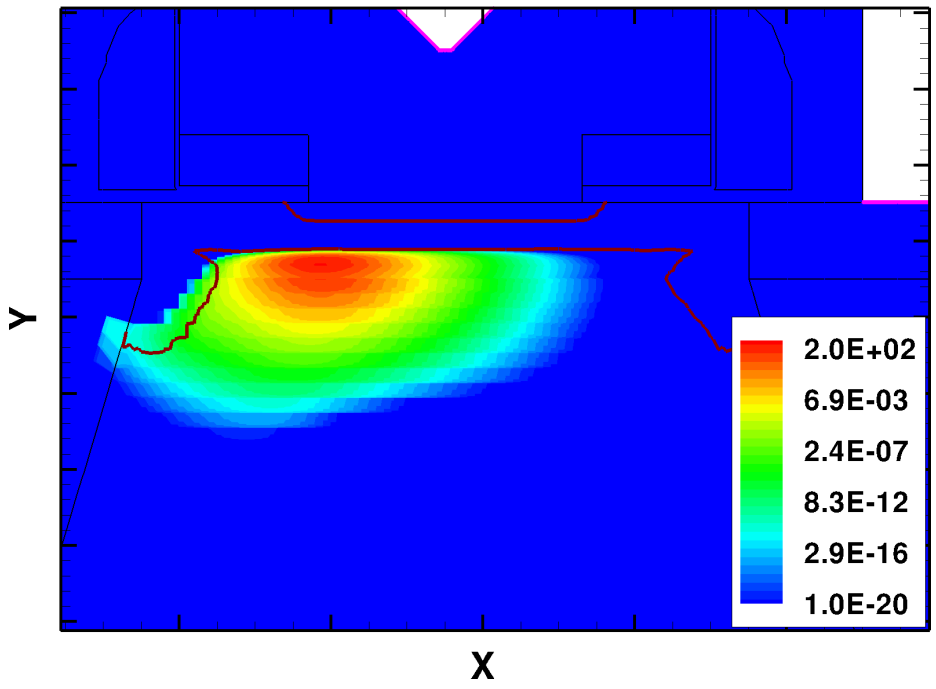

Figure 84: Rate of hot holes reaching the left EB spacer oxide interface vertex nearest the emitter window at $t = 1000s$, as a function of the hot hole origin, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

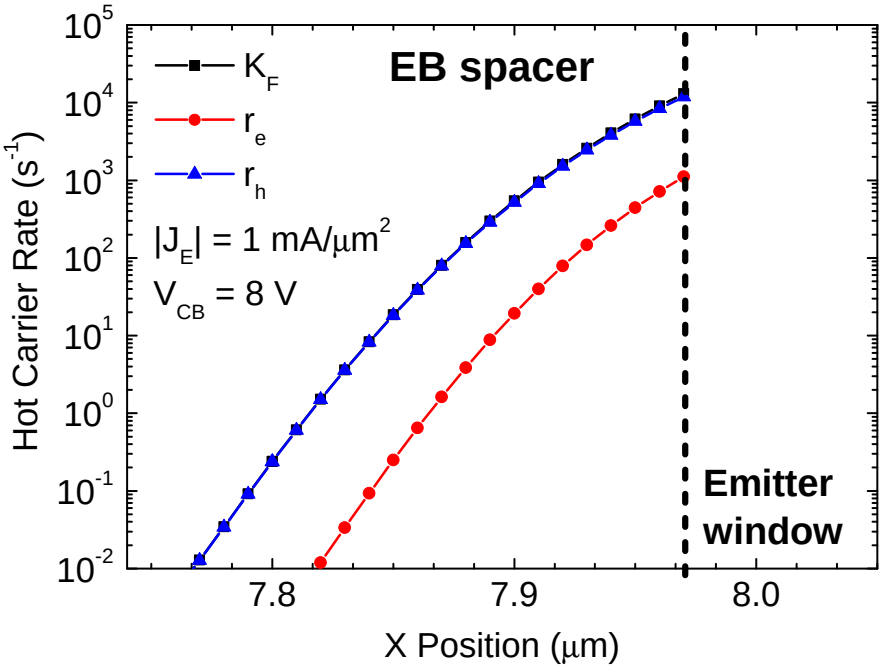

Figure 85: Rate of hot carriers reaching the left EB spacer oxide interface at $t = 1000s$ as a function of the x coordinate along the interface, for a constant stress condition of $\mathrm{J_E} = -1\,\mathrm{mA/\mu m^2}$ and $\mathrm{V_{CB}} = 8\,\mathrm{V}$.

The hot carrier rate at a particular position within the semiconductor is determined by the above probabilities and the local current density. The result is shown in Figure 83 for hot electrons and Figure 84 for hot holes (again for hot carriers reaching the single EB spacer interface vertex nearest the emitter window). As expected, the majority of hot carriers striking the EB spacer oxide originate within the collector-base space-charge region, concentrated on the side nearest the interface.

The net hot carrier rate for a particular interface vertex is the sum of all hot carrier rates computed throughout the semiconductor (i.e. the integration of the data in Figures 83-84). That rate is then calculated for each interface vertex. Figure 85 gives the forward trap formation rate (i.e. the total hot carrier rate) along with the hot electron and hot hole rates, as a function of x position along the left EB spacer oxide interface. Intriguingly, it is not hot electrons, but hot holes that dominate the overall hot carrier rate at the EB spacer for this npn SiGe HBT. This agrees directly

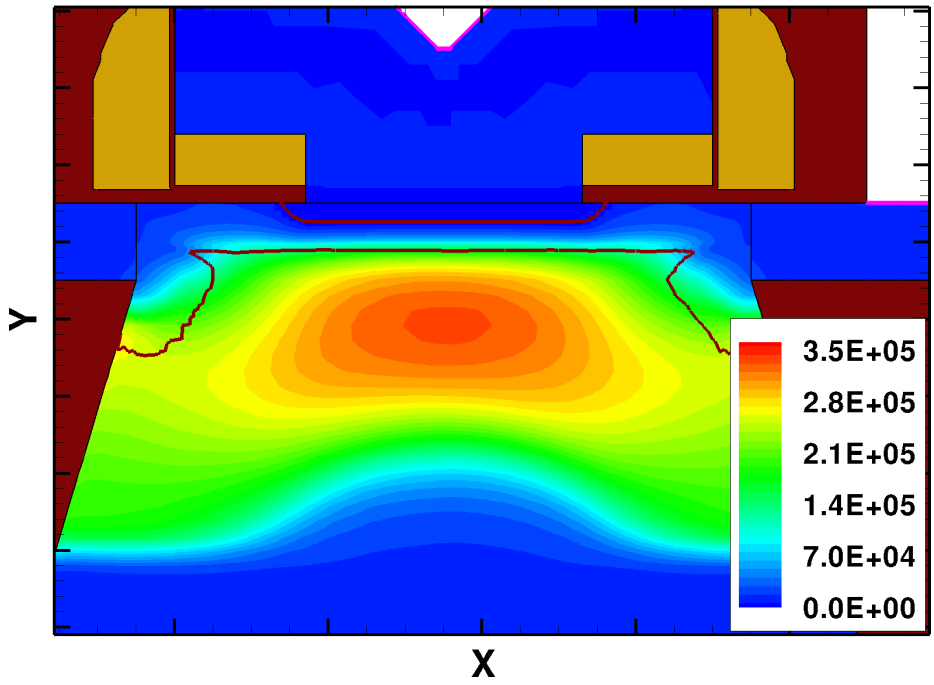

Figure 86: Electron effective electric field computed from the electron temperature distribution, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

with previous published results of full-band Monte Carlo simulations that showed that secondary holes produced by impact-ionization were the dominant carrier to reach the EB spacer oxide interface with sufficient energy to activate traps [110]. Considering the effective electric fields (shown in Figures 86-87) that drive the redirection probabilities of Figures 80-81, it is not surprising to see hot holes dominate, since hot holes are generated closer to the EB spacer oxide interface. This difference in hot carrier origin can be understood based on the direction of electron vs. hole flow combined with the fact that the effective electric fields used here are dependent on the local carrier temperatures. Since electrons are flowing toward the collector, they gain energy as they move away from the EB spacer; thus, hot electrons occur farther from the EB spacer. In contrast, holes that are generated by impact ionization are driven towards the EB spacer due to the polarity of the electric field and are therefore more energetic closer to the EB spacer. These differences drive the effective electric field distributions of Figures 86-87.

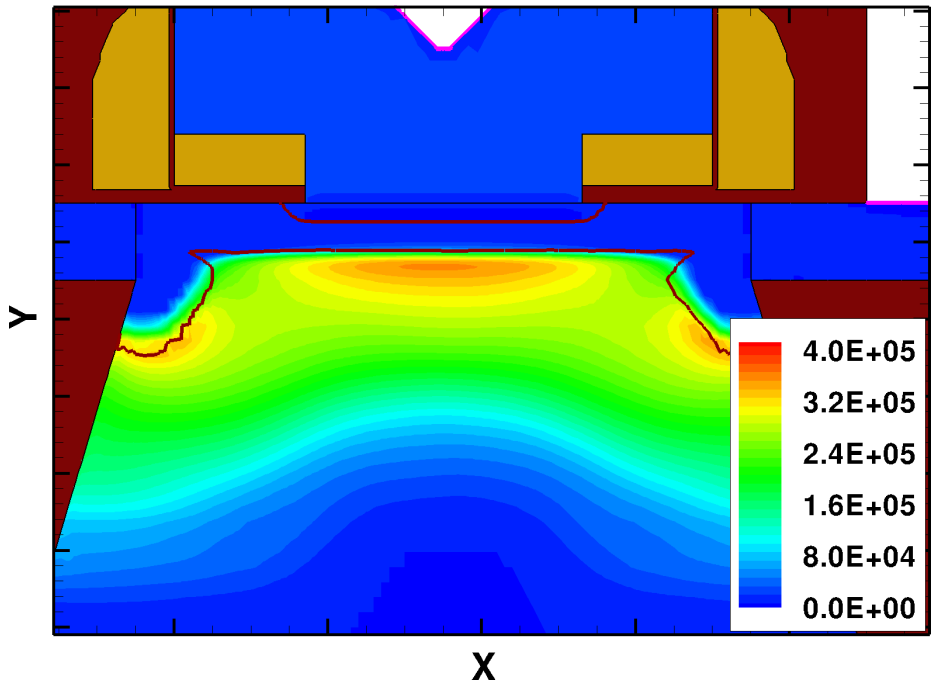

Figure 87: Hole effective electric field computed from the electron temperature distribution, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

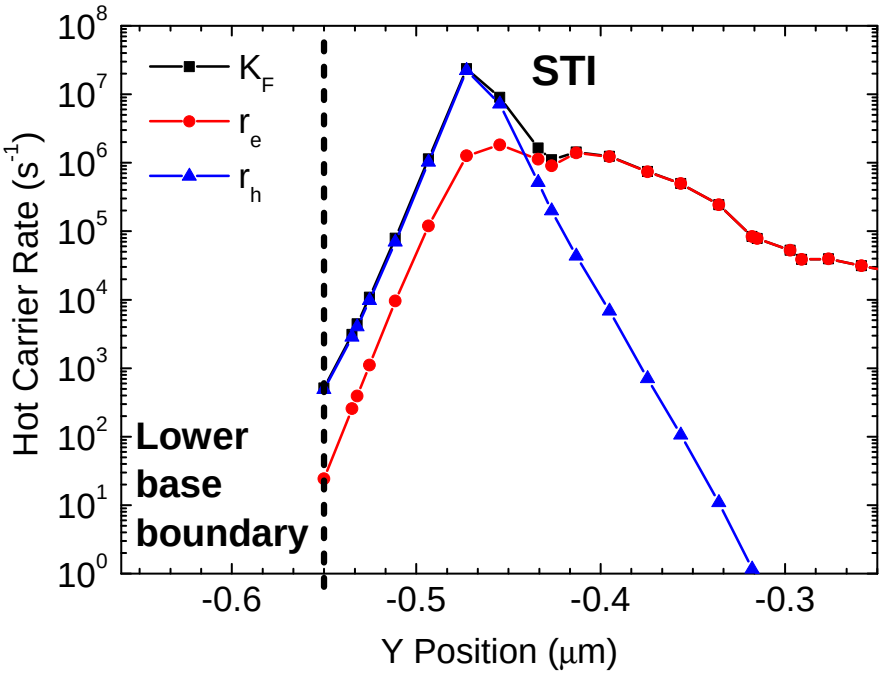

Figure 88: Rate of hot carriers reaching the left STI interface at $t = 1000s$ as a function of the y coordinate along the interface, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

Figure 88 shows the trap formation, hot electron, and hot hole rates as a function of y position along the left STI interface. Here, hot holes dominate the hot carrier rate only for interface positions close to and above the metallurgical junction where the extrinsic base doping has diffused outward into the active region of the device. This can be traced back to Figure 81, which reveals a moderate hot hole redirection probability in that region. This moderate probability arises from the large hole effective electric field that is directly adjacent to the oxide interface at the metallurgical junction. Although the electric field is much smaller than in the collector-base space charge region, the hot hole rate remains significant because of its proximity to the oxide interface. Above this region, hot holes dominate because the region of large hole effective electric field is in closer proximity. Below the metallurgical junction, the hot electron rate dominates the overall hot carrier rate because the hot hole rate drops exponentially with distance from the metallurgical junction, whereas the rate of hot electrons produced by the large electron effective electric field of the collector-base junction remains relatively high. The net hot electron and hot hole rates are much higher for the STI interface than for the EB spacer interface, and this results in much higher trap densities, as can be seen in Figure 89. This falls in line with the well-established fact that the mixed-mode degradation mechanism more severely affects inverse-mode vs. forward-mode device operation [16, 18, 118, 120]. These results also indicate that modeling of the outdiffusion of the extrinsic base doping is critical to capturing the distribution of traps along the STI, since the hot carrier rate is orders of magnitude higher at the location of the metallurgical junction than anywhere else along the STI.

After the net hot carrier rates have been summed following each time step, the interface trap densities can be computed. Figure 90 shows the time evolution of the interface trap density as a function of position along the left EB spacer, where the peak trap density occurs as expected at the emitter window. Considering only the

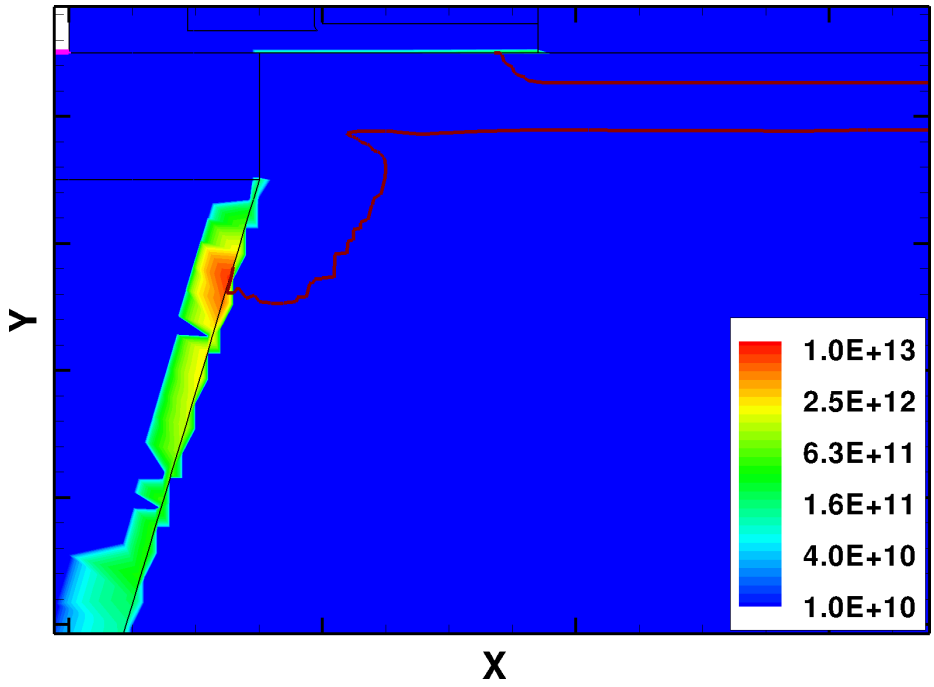

Figure 89: Interface trap density at $t = 1000s$, for a constant stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$.

peak trap densities, the time dependence of trap formation at the EB spacer and STI interfaces is compared in Figure 91. For this stress condition, the peak STI trap density increases more rapidly than the peak EB trap density, which must result from an increase in K_F with time, most likely caused by changes in the current density near the STI interface that result from increased recombination current. This difference in the rate of trap creation decreases as the stress time increases.

Directly correlating with the trends established in mixed-mode stress data [16, 18, 118, 120], the damage caused to the EB spacer and STI interfaces can be separated by performing both forward and inverse Gummel measurements following various stress intervals. Since the majority of the base current flows nearer to the EB spacer in forward-mode operation, the change in non-ideal I_B can be linked to traps at the EB spacer interface. Likewise, the change in inverse-mode I_B can be attributed to traps at the STI interface, at which point the majority of the base current flows near that interface. The forward and inverse-mode Gummel characteristics are shown in Figures

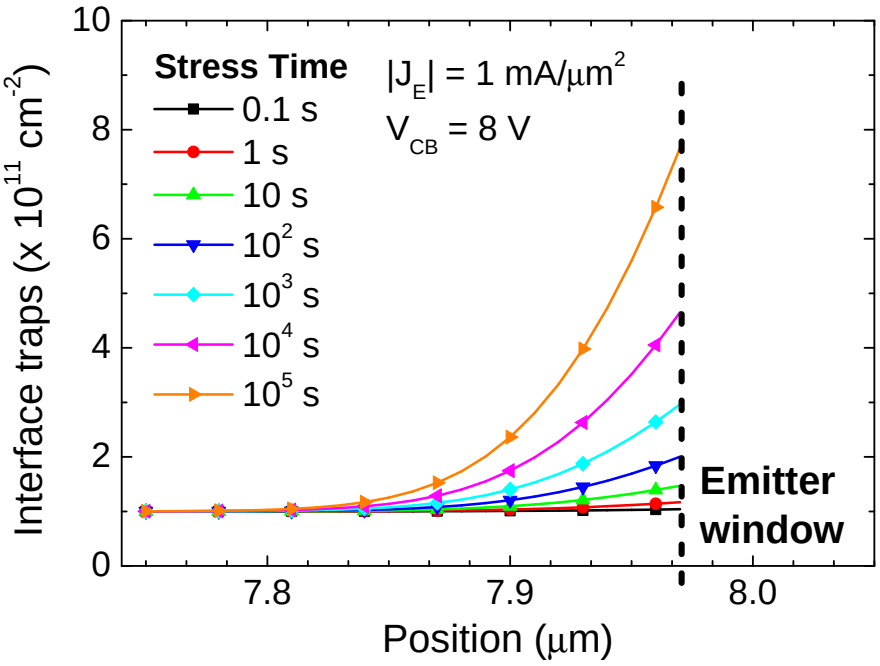

Figure 90: Interface trap density along the EB spacer oxide for increasing stress time. The npn SiGe HBT is stressed at $J_E = -1\,mA/\mu m^2$, $V_{CB} = 8\,V$.

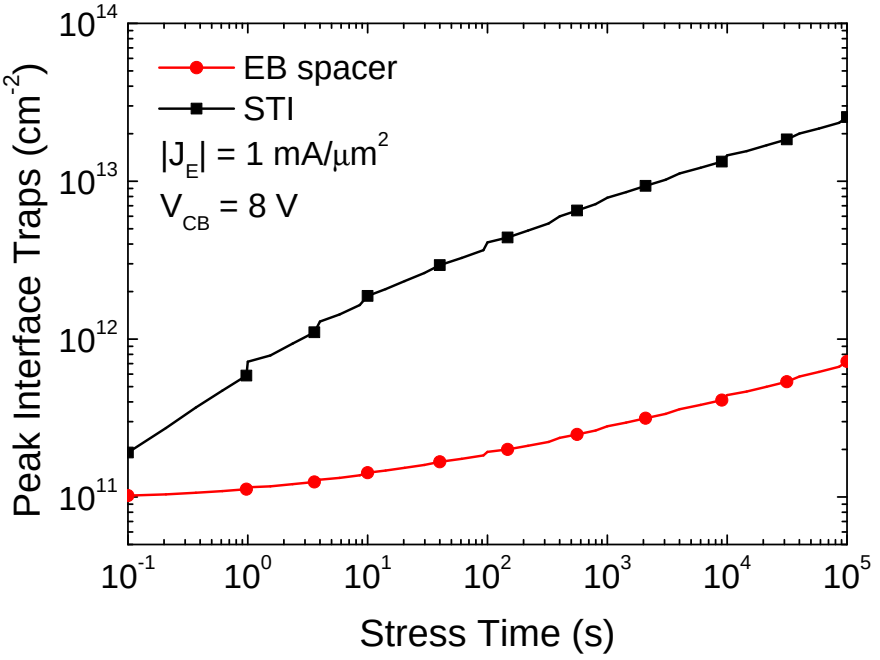

Figure 91: Peak interface trap density at the EB spacer and STI oxides as a function of stress time. The npn SiGe HBT is stressed at $J_E = -1\,mA/\mu m^2$, $V_{CB} = 8\,V$.

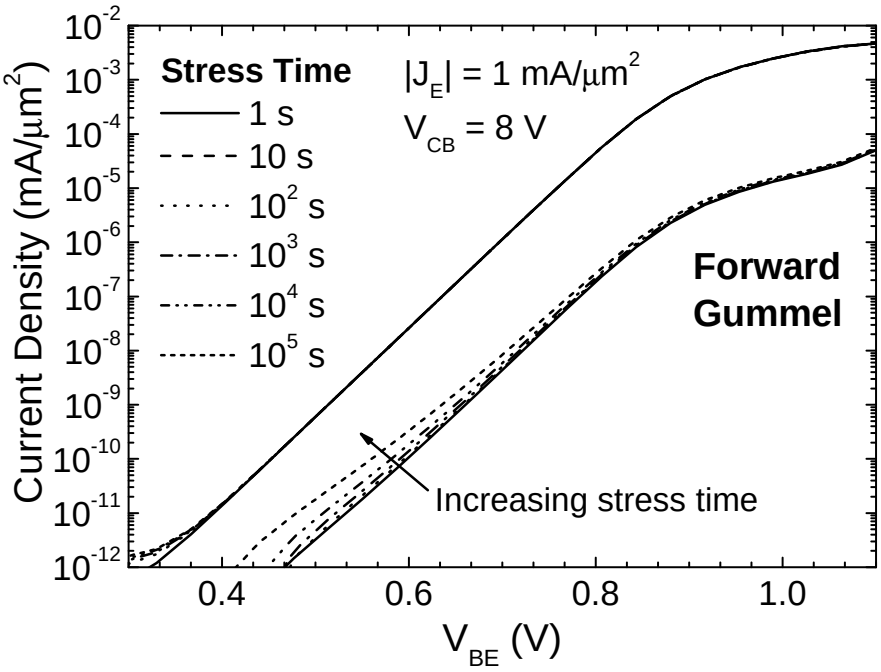

Figure 92: Post-stress forward gummel characteristics for stress times ranging from 1 s to 10^5 s, with a stress condition of $J_E = -1\,mA/\mu m^2$, $V_{CB} = 8\,V$.

92-93, respectively, for stress times increasing from 1 s to 10^5 s. The effect of the traps at each interface can be clearly distinguished by the two measurements, since the inverse-mode I_B exhibits a much larger change due to the much higher interface trap density at the STI interface.

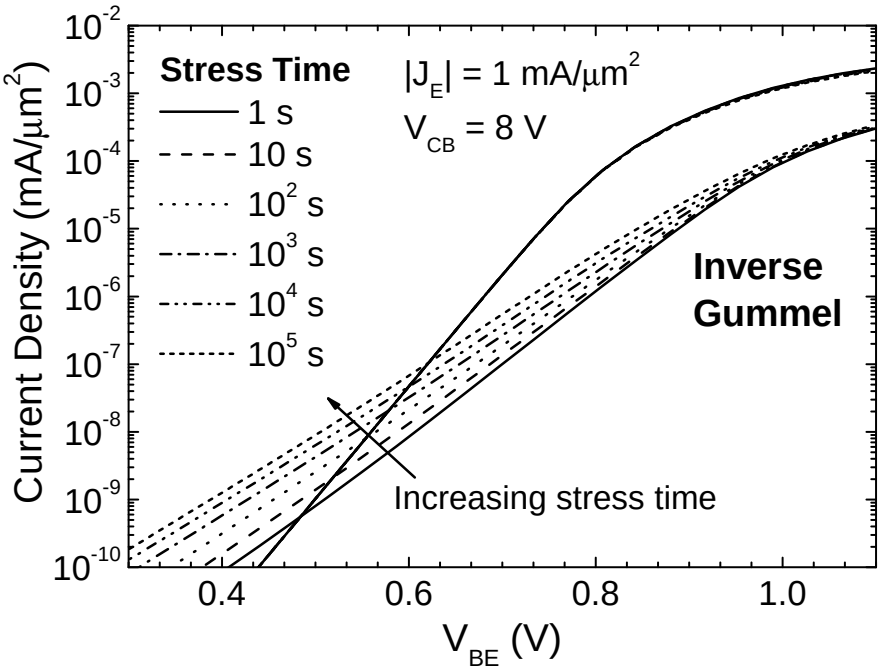

Figure 93: Post-stress inverse gummel characteristics for stress times ranging from 1 s to 10^5 s, with a stress condition of $J_E = -1\,mA/\mu m^2$, $V_{CB} = 8\,V$.

6.3 Bias Dependence and Calibration to Data

Now that the PMI degradation model has been demonstrated with a stress condition of $J_E = -1\,mA/\mu m^2$ and $V_{CB} = 8\,V$ as a test case, the bias dependence of the mixed-mode degradation can be simulated. Figure 94 shows the peak trap density at the EB and STI interfaces as a function of the stress condition, taken after a stress interval of 10^3 s. As expected, for a greater collector-base reverse bias, more traps are formed at both interfaces due to the greater electric field. Interestingly, however, the maximum traps at the EB spacer are formed for a current density near peak-f_T, whereas the peak STI traps continue to increase with increasing current density. This can be attributed to the fact that when high-injection effects onset at current densities above peak-f_T, the collector-base electric field is pushed further into the physical collector, as demonstrated by the evolving electron and hole effective electric fields in Figure 95. Since this is in the direction directly away from the EB spacer, a dramatic decrease in hot carrier damage is observed. In contrast, the peak electric field remains roughly

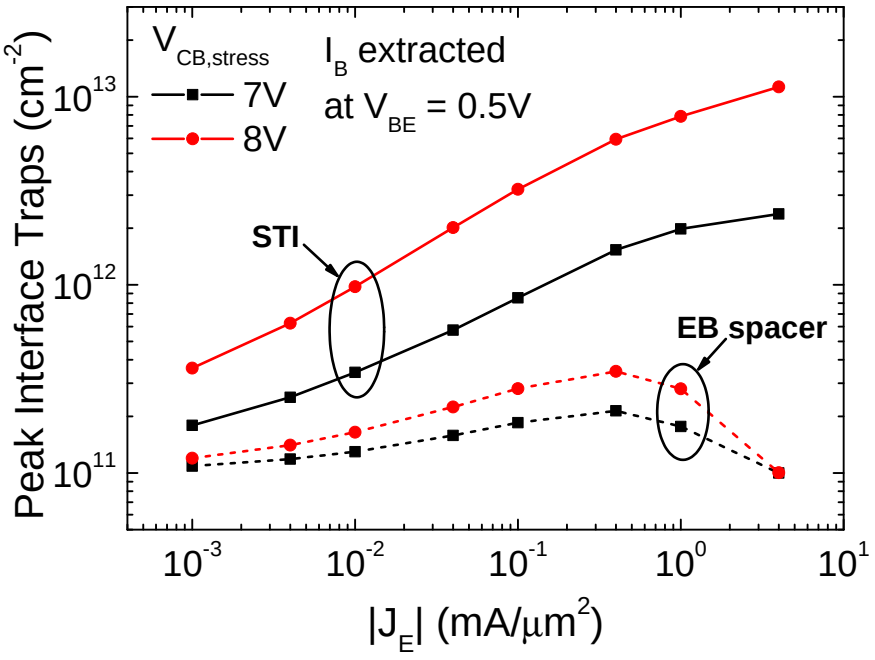

Figure 94: Peak interface trap density at the EB spacer and STI oxides for increasing magnitude of emitter current stress. The trap density is extracted after 10^3 s stress at $V_{CB} = 8$ V.

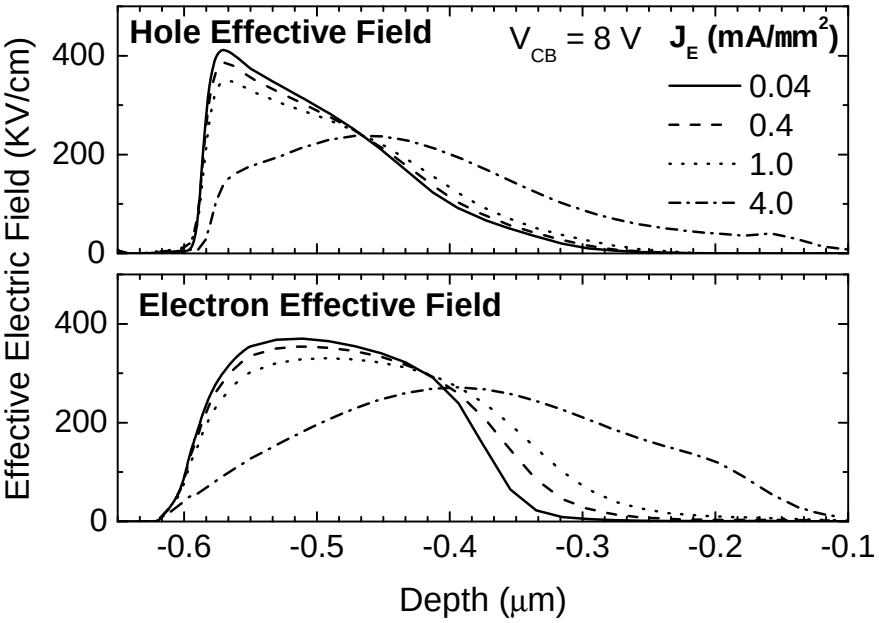

Figure 95: Hole and electron effective electric fields, taken from a vertical cut through the center of the npn SiGe HBT, for a range of emitter current stress conditions with $V_{CB} = 8$ V.

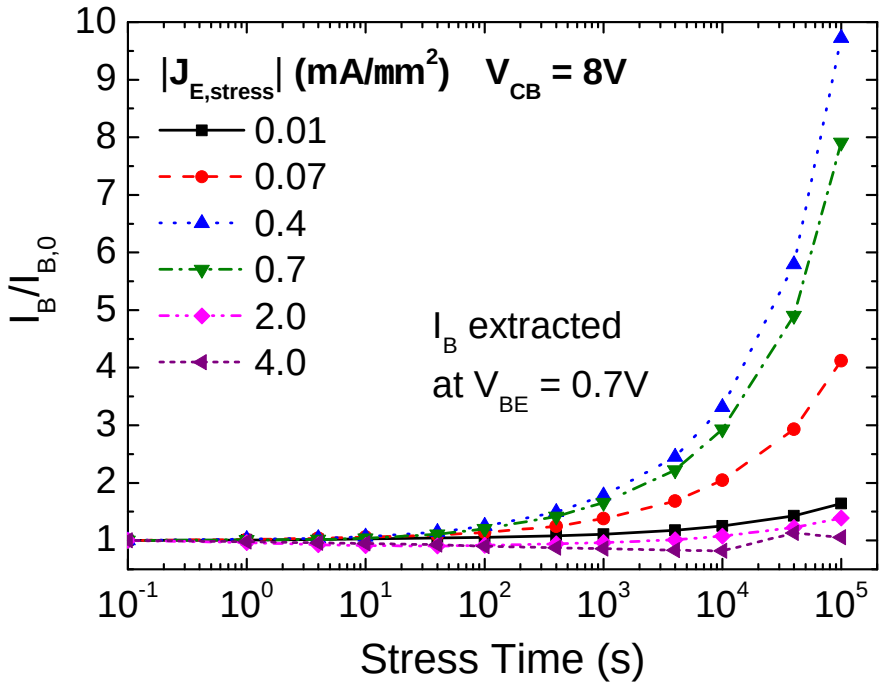

Figure 96: Normalized increase in base current as a function of increasing stress time, for a range of emitter current stress conditions with $V_{CB} = 8\,V$.

the same distance from the STI, so hot carriers will continue to damage the STI at an increasing rate as the current density rises. The change in the slope of the STI curves of Figure 94 is still expected, however, since the peak magnitude of the electric field is reduced due to the Kirk effect.

The increase in forward-mode base current, normalized to its pre-stress value, is given in Figure 96 as a function of stress time for a variety of current density stress conditions. The normalized I_B is extracted from post-stress Gummel characteristics at a fixed $V_{BE} = 0.7\,V$. Since the maximum interface trap density occurs for emitter current densities between 0.1 to $1\,mA/\mu m^2$, the greatest change in non-ideal base current is also observed for these stress conditions. At stress currents greater than $-1\,mA/\mu m^2$, the base current change is dramatically reduced, with little to no damage observed. Figure 97 highlights the effect of stress condition with much greater clarity. The normalized base current change after $10^3\,s$ stress is shown as a function of the current density stress condition for three different collector-base reverse-biases. These

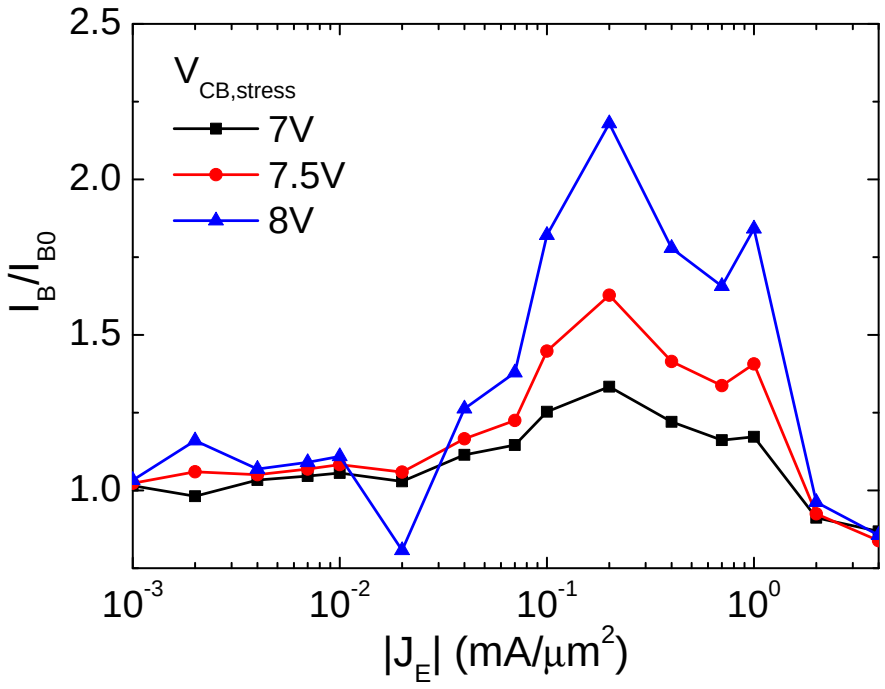

Figure 97: Normalized increase in base current after 10^3 s stress as a function of stress condition.

results show that the amount of damage is proportional to the magnitude of the collector-base electric field (i.e. V_{CB} reverse bias), whereas the damage increases with the emitter current density only up to the onset of high-injection effects.

To validate this first-order model for mixed-mode degradation in SiGe HBTs, measurements were performed on the devices to which the 2-D TCAD model was calibrated. The stress conditions used for this test held the emitter current density at $-1\,\mathrm{mA/\mu m^2}$, with the collector-base reverse-bias varied from 7–8 V. Excellent correlation between the simulated and measured base current change is shown in Figure 98. To achieve this fit, only a single free parameter was tuned, the mean free path, while all other model parameters were held constant at the values specified in Section 6.1, each of which is well in line with published values. The electron and hole capture cross-sections were set to be $10^{-14}\,\mathrm{cm^2}$ [14]. A mean free path value of 6.2 nm proved to give good agreement to the data across the different stress conditions and is a reasonable value compared to others used in the literature [15, 38]. In reality,

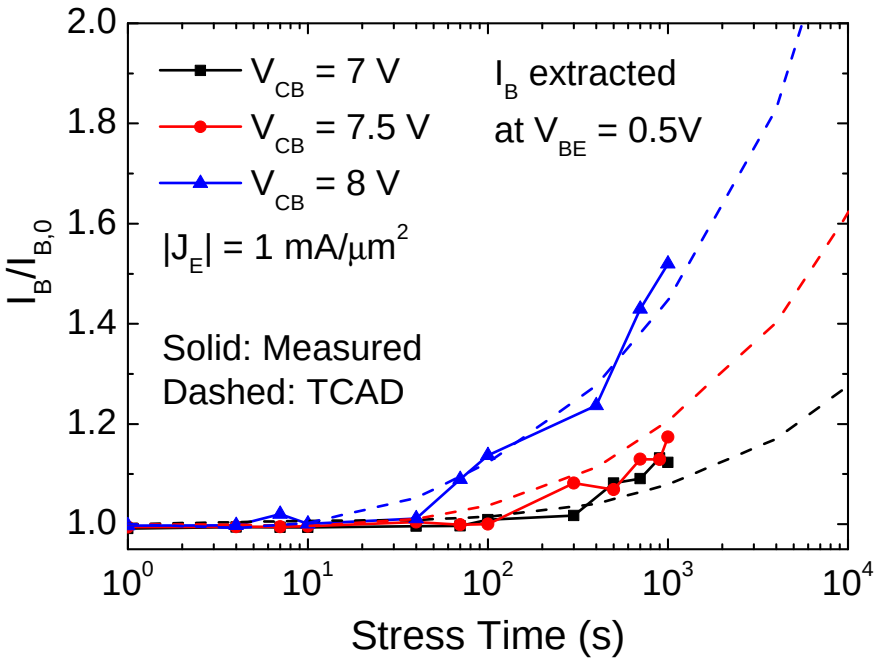

Figure 98: Comparison between simulation and measurement of normalized base current vs. stress time for a range of V_{CB} stress conditions with $J_E = -1\,mA/\mu m^2$.

the mean free path is an energy-dependent parameter that is neither constant nor universal across all technologies and devices, but for the purposes of this first-order degradation model, it is sufficient to use it as a free parameter with a value comparable to those previously reported in the literature.

6.4 Application to Mixed-Mode TCAD

Since the bias-dependence of the mixed-mode degradation mechanism can be accurately captured by the model developed in this chapter, this model is well-suited for transient degradation simulations of not just single devices, but also circuits. This is specifically enabled by the coupling that is built in to the model, where the impact of increasing trap concentrations at the oxide interfaces is accounted for at each new time step of the transient degradation simulation. Sentaurus Device provides a mixed-mode TCAD capability in which compact models and TCAD models are solved simultaneously. With this in mind, the degradation model presented here can be enabled within a

mixed-mode TCAD simulation of a mixed-signal circuit, and in such a simulation, the TCAD device will be exposed to the time-varying stress conditions imposed by the external circuit. Since the degradation model physically accounts for the bias-dependence of trap formation, this will provide a much more predictive assessment of the degradation and time-to-failure of that specific circuit.

Several areas remain to enhance this model, specifically the inclusion of annealing that occurs at higher temperatures and large current densities. Presently, the solution for the interface trap density given in (17) was arrived at by several assumptions, one of which is that the forward trap reaction rate is much larger than the reverse trap reaction rate. In reality, when the stress condition is such that the rate of hot carriers reaching the interface is negligible, hydrogen molecules will diffuse back to the interface and repassivate the interface traps, resulting in a net decrease in traps. Moreover, the influence of self-heating cannot be neglected, as this not only affects carrier transport and impact ionization, but will also change the diffusion constant of hydrogen; at higher temperatures, hydrogen will diffuse more quickly, enhancing both the trap formation and annealing rates, depending on the stress condition.

CHAPTER VII

CONCLUSIONS

This dissertation has focused on improving predictive modeling techniques to address challenges faced during the development of new silicon-based mixed-signal technologies, recognizing that device and circuit reliability is a key concern for all operating environments. A diverse array of individual topics has been addressed in this research, spanning the material, device, and circuit layers of abstraction, along with both conventional and radiation-related reliability. Nevertheless, the topics are unified by a single theme, which is to apply predictive TCAD to address circuit-level reliability issues, rooted in a fundamentally physical understanding of the underlying reliability mechanisms. Thus, while one of the initial topics of this research encompasses the basic physics of charge transport, it then feeds forward to improve predictive modeling of device operation and reliability, which then in turn feeds forward to enable better predictive modeling of circuit-level reliability. One of the greatest contributions of this work has been to apply novel mixed-mode TCAD simulations to explain with unprecedented detail how reliability effects within a specific circuit can be traced back to the underlying physical mechanisms within the transistors. It is exactly this type of approach that is needed to understand and address reliability challenges in the most effective and efficient way, taking a more holistic approach to reliability, in which device-level mechanisms are considered within the context of real-life circuits, since the dynamic circuit conditions necessarily couple into and influence the physics of carrier transport in the device. With this type of contextual understanding, device-level reliability problems can be mitigated in the most effective way to influence the overall circuit and system reliability response.

In addition to taking a vertical view of reliability issues and tracing individual effects from their physical origins to their circuit-level manifestation, much can be gained by applying lessons learned in one reliability area to adjacent areas. In my research I have done this by applying knowledge gained from radiation-related reliability studies to address conventional degradation modes of SiGe BiCMOS circuits. This type of approach can bear much fruit, and this has been demonstrated previously in reliability publications. One specific example is how an understanding of conventional hot-carrier degradation of the SiGe HBT was applied to explain the total-ionizing-dose degradation of the same devices [16]. Here, in the context of predictive modeling of reliability, my extensive research concerning how best to model circuit-level single-event transients sets the foundation for the development of a novel mixed-mode TCAD capability for modeling conventional hot-carrier-induced degradation of SiGe BiCMOS circuits.

A number of specific contributions can be counted from this dissertation, and their significance to the greater semiconductor device community is evidenced through many peer-reviewed publications. Briefly, the main contributions of this work are summarized here:

- A novel 2-D quasi-static regional transit time analysis [69]. This work contributes toward optimization and analysis of the SiGe HBT in several significant ways. First, this analysis can be used in place of traditional frequency-domain simulations to estimate the cutoff frequency across bias conditions, greatly reducing the simulation time needed for iterative optimization of doping and Ge profiles. Moreover, this analysis provides a unique view of the 2-D contributions to the total device delay, enabling more intelligent optimization of device profiles based on the limiting regions. Finally, the streamline-based regional transit-time analysis further allows device designers to identify trends and challenges posed to continued performance-scaling of the SiGe HBT.

- Measurement and modeling of charge carrier transport parameters (mobility,

incomplete ionization, and Shockley-Read-Hall recombination) as a function of radiation dose and temperatures spanning from the deep cryogenic regime up to 300 °C [67]. The primary contributions of this work include a large body of temperature-dependent recombination lifetime data, something that did not exist hitherto for such a wide temperature range or for a commercial IC technology. Furthermore, this data and the equally large body of n-type and p-type resistance data were applied to develop calibrated mobility, incomplete-ionization, and recombination models, which together provide a basis for accurate modeling of charge transport in silicon-based devices and can be applied to predictive TCAD modeling of single-event transients.

- The first direct measurements of single-event transients in 45 nm RF-CMOS on SOI [66]. The primary contribution of this work is the illumination of the influence of different body-tie geometries on the single-event transient sensitivity. This data is necessary to identify the optimal device layout to simultaneously balance RF performance, total-ionizing dose sensitivity, and single-event effects sensitivity.

- Modeling of single-event transients in an analog circuit building block, a SiGe precision voltage reference [65]. The main contributions of this study include the first comparison of circuit-level SET modeling approaches with measured data, which highlights the fundamental difference between coupled and decoupled modeling approaches. Moreover, the new mixed-mode TCAD approach was used to identify for the first time the underlying mechanism that drives the SET of this voltage reference topology.

- Modeling of single-event transients in a high-speed SiGe digital latch [68]. This study builds heavily upon the knowledge gained from the voltage reference study and provides much needed insight that explains single-event upsets in

SiGe shift registers. One key contribution is the identification of two distinct SEU mechanisms that are linked to their physical origins, one caused by high forward currents induced by emitter strikes, and the other caused by substrate-subcollector charge collection. The other significant contribution is direct proof and analysis of the inaccuracies of conventional decoupled circuit SET simulations.

- A physics-based model for the mixed-mode degradation mechanism in SiGe HBTs, suitable for mixed-mode TCAD simulations of the degradation of dynamically-biased SiGe circuits. The greatest contribution of this work is the first implementation of a fully-coupled TCAD-based model that can accurately capture mixed-mode degradation of the SiGe HBT. This model was validated directly against mixed-mode stress measurements

7.1 Future Work

Although this dissertation has answered many questions about best modeling practices for reliability, it has also opened up several new areas that should prove fruitful for further research. From the knowledge I have gained throughout this research, I will attempt to outline here the main topics that need further attention to better address both conventional and radiation-related circuit reliability challenges.

7.1.1 Single-Event Transient Modeling

In Chapter 5, I have discussed at length how the coupled mixed-mode TCAD approach is unique in its ability to link circuit-level single-event effects to their physical origins within the transistor. This was shown for a basic analog building block, the voltage reference, as well as a basic digital building block, the flip flop. Those results showed conclusively how decoupled current-injection simulations are fundamentally flawed and can provide completely inaccurate predictions of the circuit SET response. Future work

needs to extend this type of study to other basic building blocks, including RF circuits such as voltage-controlled oscillators and low noise amplifiers. In dynamic circuits such as those, decoupled simulation results are likely to be even further from the true circuit response; in those cases, coupled mixed-mode TCAD simulations are uniquely situated to provide a detailed look at the interaction between the device SET mechanism and the circuit response. Furthermore, there is room for additional research along the lines of the flip flop study of Section 5.3, where the SEE response of additional latch topologies such as the RHBD latch designs of [49] and [50] can be modeled in mixed-mode TCAD to better explain existing experimental data. Ultimately, there are two main aspects to this research that should be built upon: (1) establishing guidelines for circuit designers on how to best model circuit-level SEE based on different circuit types and technologies, and (2) identifying the physical origin of circuit-level SEE sensitivities so that better hardening approaches can be developed.

7.1.2 Conventional Reliability Modeling

Chapter 6 branches off from the application of mixed-mode TCAD to radiation effects in order to take a similar approach for conventional reliability issues in the SiGe HBT. The mechanisms behind the long-term degradation of semiconductor devices are enormously complex, so a strategic approach is necessary. Before mixed-mode TCAD simulations of circuit degradation can be performed, predictive models of the physical device degradation mechanisms are needed. For mixed-signal SiGe circuits, the mixed-mode degradation mechanism is of chief concern, since it is triggered at the typical bias conditions present in those circuits (high J_C and large reverse collector-base bias). This dissertation provides a huge step toward coupled modeling of this degradation mode by developing a physics-based model that not only is validated against measured device stress data, but is also designed for coupled mixed-mode simulations. Several enhancements were already pointed out in Chapter 6 as to how

the physics model can be enhanced to account for trap annealing and its temperature dependence. In addition, suitable mixed-signal circuits need to be fabricated and tested, then compared directly against mixed-mode TCAD degradation simulations of the same circuit using calibrated TCAD models of the transistors. At that point, optimization of the transistors themselves can be performed and the impact on circuit reliability can be directly quantified.

REFERENCES

[1] Adell, P., Schrimpf, R., Barnaby, H., Marec, R., Chatry, C., Calvel, P., Barillot, C., and Mion, O., "Analysis of single-event transients in analog circuits," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pp. 2616–2623, 2000.

[2] Agarwal, P., Hurkx, G., Donkers, J., and Slotboom, J., "On the delay times in vertically scaled SiGe HBTs," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 264–267, 2005.

[3] Altermatt, P. P., Schenk, a., and Heiser, G., "A simulation model for the density of states and for incomplete ionization in crystalline silicon. I. Establishing the model in Si:P," *J. Appl. Phys.*, vol. 100, no. 11, p. 113714, 2006.

[4] Altermatt, P. P., Schenk, A., Schmithusen, B., and Heiser, G., "A simulation model for the density of states and for incomplete ionization in crystalline silicon. II. Investigation of Si:As and Si:B and usage in device simulation," *J. Appl. Phys.*, vol. 100, no. 11, p. 113715, 2006.

[5] Appaswamy, A., Jun, B., Diestelhorst, R., Espinel, G., Prakash, A., Cressler, J., Marshall, P., Marshall, C., Liang, Q., Freeman, G., and Others, "The effects of proton irradiation on 90 nm strained Si CMOS on SOI devices," in *Proc. IEEE Radiation Effects Data Workshop*, pp. 62–65, 2007.

[6] Bellini, M., Cressler, J. D., Turowski, M., Avenier, G., Chantre, A., and Chevalier, P., "3-D regional transit time analysis of SiGe HBTs on thin-film SOI," *ECS Trans.*, vol. 16, no. 10, pp. 1079–1088, 2008.

[7] Birkholz, J. E., Bothe, K., Macdonald, D., and Schmidt, J., "Electronic properties of iron-boron pairs in crystalline silicon by temperature- and injection-level-dependent lifetime measurements," *J. Appl. Phys.*, vol. 97, no. 10, p. 103708, 2005.

[8] Bleichner, H., Jonsson, P., Keskitalo, N., and Nordlander, E., "Temperature and injection dependence of the Shockley-Read-Hall lifetime in electron irradiated n-type silicon," *J. Appl. Phys.*, vol. 79, no. 12, pp. 9142–9148, 1996.

[9] Calin, T., Nicolaidis, M., and Velazco, R., "Upset hardened memory design for submicron CMOS technology," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pp. 2874–2878, 1996.

[10] CFD Research Corporation, "NanoTCAD, Version 2011."

[11] Chakraborty, P. S., Appaswamy, A. C., Saha, P. K., Jha, N. K., Cressler, J. D., Yasuda, H., Eklund, B., and Wise, R., "Mixed-mode stress degradation mechanisms in pnp SiGe HBTs," in *Proc. Int. Reliability Phys. Symposium*, pp. 83–88, 2009.

[12] Chakraborty, P. S., Moen, K. A., and Cressler, J. D., "An Investigation into the Challenges and Issues Associated with Device Design, Optimization and Scaling of Complementary SiGe HBTs," *IEEE Trans. Electron Devices*, submitted.

[13] Chakraborty, P., Moen, K., and Bellini, M., "Investigation of the device design challenges and optimization issues associated with complementary SiGe HBT scaling," in *Proc. Int. Semiconductor Device Research Symposium*, vol. 4, pp. 2009–2010, 2010.

[14] Chang, M., Zhang, J., and Zhang, W., "Assessment of capture cross sections and effective density of electron traps generated in silicon dioxides," *IEEE Transactions on Electron Devices*, vol. 53, no. 6, pp. 1347–1354, 2006.

[15] Cheng, P., Grens, C., and Cressler, J., "Reliability of SiGe HBTs for power amplifiers—Part II: Underlying physics and damage modeling," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 3, pp. 440–448, 2009.

[16] Cheng, P., Jun, B., Sutton, A., Appaswamy, A., Zhu, C., Cressler, J. D., Schrimpf, R. D., and Fleetwood, D. M., "Understanding Radiation- and Hot Carrier-Induced Damage Processes in SiGe HBTs Using Mixed-Mode Electrical Stress," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 6, pp. 1938–1945, 2007.

[17] Cheng, P., Seth, S., Grens, C., Thrivikraman, T. K., Bellini, M., Cressler, J. D., Babcock, J., Chen, T., Kim, J., and Buchholz, A., "An experimental investigation of RF safe-operating-area (SOA) in SiGe HBTs on SOI," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 17–20, 2009.

[18] Cheng, P., Zhu, C., Appaswamy, A., and Cressler, J., "A New Current-Sweep Method for Assessing the Mixed-Mode Damage Spectrum of SiGe HBTs," *IEEE Trans. Device Mater. Rel.*, vol. 7, no. 3, pp. 479–487, 2007.

[19] Cheng, P., Zhu, C., Cressler, J. D., and Joseph, A., "The Mixed-Mode Damage Spectrum of Sige HBTs," in *Proc. Int. Reliability Phys. Symposium*, pp. 566–567, 2007.

[20] Chevalier, P., Pourchon, F., Lacave, T., Avenier, G., Campidelli, Y., Depoyan, L., Troillard, G., Buczko, M., Gloria, D., Celi, D., Gaquiere, C., and Chantre, A., "A conventional double-polysilicon FSA-SEG Si/SiGe:C HBT reaching 400 GHz fMAX," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 2008–2011, 2009.

[21] Choi, L. and Sibaja-Hernandez, A., "SiGe HBTs with normal high-speed emitter-up and reverse low-power collector-up operation," *IEEE Trans. Electron Devices*, pp. 358–364, 2008.

[22] Cressler, J., "Emerging SiGe HBT reliability issues for mixed-signal circuit applications," *IEEE Trans. Device Mater. Rel.*, vol. 4, no. 2, pp. 222–236, 2004.

[23] Cressler, J., "On the potential of SiGe HBTs for extreme environment electronics," *Proc. IEEE*, vol. 93, no. 9, pp. 1559–1582, 2005.

[24] Cressler, J., *Silicon Heterostructure Handbook: Materials, Fabrication, Devices, Circuits, and Applications of SiGe and Si Strained-Layer Epitaxy*. CRC Press, 2006.

[25] Cressler, J., Krithivasan, R., Zhang, G., Niu, G., Marshall, P., Kim, H., Reed, R., Palmer, M., and Joseph, A., "An investigation of the origins of the variable proton tolerance in multiple SiGe HBT BiCMOS technology generations," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 3203–3207, 2002.

[26] Cressler, J. and Niu, G., *Silicon-Germanium Heterojunction Bipolar Transistors*. Artech House Publishers, 2003.

[27] DiMaria, D. J. and Stasiak, J. W., "Trap creation in silicon dioxide produced by hot electrons," *J. Appl. Phys.*, vol. 65, no. 6, p. 2342, 1989.

[28] Dodd, P., Shaneyfelt, M., Felix, J., and Schwank, J., "Production and propagation of single-event transients in high-speed digital logic ICs," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp. 3278–3284, 2004.

[29] Dodd, P., Shaneyfelt, M., Schwank, J., and Felix, J., "Current and future challenges in radiation effects on CMOS electronics," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 4, pp. 1747–1763, 2010.

[30] Ferlet-Cavrois, V., Marcandella, C., Giraud, G., Gasiot, G., Colladant, I., Musseau, O., and Fenouillet, C., "Characterization of the parasitic bipolar amplification in SOI technologies submitted to transient irradiation," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 3, pp. 1456–1461, 2002.

[31] Ferlet-Cavrois, V. and Paillet, P., "Analysis of the transient response of high performance 50-nm partially depleted SOI transistors using a laser probing technique," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 4, pp. 1825–1833, 2006.

[32] Ferlet-Cavrois, V., Paillet, P., Gaillardin, M., Lambert, D., Baggio, J., Schwank, J., Vizkelethy, G., Shaneyfelt, M., Hirose, K., Blackmore, E., and Others, "Statistical analysis of the charge collected in SOI and bulk devices under heavy Ion and proton irradiation—Implications for digital SETs," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3242–3252, 2006.

[33] Francis, A., Turowski, M., Holmes, J., and Mantooth, H., "Efficient modeling of single event transients directly in compact device models," in *IEEE International Behavioral Modeling and Simulation Workshop*, pp. 73–77, 2007.

[34] Goebel, H. and Hoffmann, K., "Full dynamic power diode model including temperature behavior for use in circuit simulators," in *International Symposium on Power Semiconductor Devices and ICs*, pp. 130–135, 1992.

[35] Green, M., "Minority carrier lifetimes using compensated differental open circuit voltage decay," *Solid-State Electron.*, vol. 26, no. 11, pp. 1117–1122, 1983.

[36] Grens, C., Cheng, P., and Cressler, J., "Reliability of SiGe HBTs for power amplifiers—Part I: Large-signal RF performance and operating limits," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 3, pp. 431–439, 2009.

[37] Hall, R., "Electron-hole recombination in germanium," *Phys. Rev.*, vol. 87, no. 2, pp. 387–387, 1952.

[38] Hasnat, K., Yeap, C.-F., Jallepalli, S., Shih, W.-K., Hareland, S., Agostinelli, V., Tasch, A., and Maziar, C., "A pseudo-lucky electron model for simulation of electron gate current in submicron NMOSFET's," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 8, pp. 1264–1273, 1996.

[39] Hauser, J., Diehl-Nagle, S., Knudson, A., Campbell, A., Stapor, W., and Shapiro, P., "Ion track shunt effects in multi-junction structures," *IEEE Trans. Nucl. Sci.*, vol. 32, no. 6, pp. 4115–4121, 1985.

[40] Hirose, K., Saito, H., Fukuda, S., Kuroda, Y., Ishii, S., Takahashi, D., and Yamamoto, K., "Analysis of body-tie effects on SEU resistance of advanced FD-SOI SRAMs through mixed-mode 3-D simulations," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp. 3349–3353, 2004.

[41] Hu, C., "Lucky-electron model of channel hot electron emission," in *Proc. Int. Electron Device Meeting*, pp. 22–25, IRE, 1979.

[42] Jeppson, K. O. and Svensson, C. M., "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," *J. Appl. Phys.*, vol. 48, no. 5, p. 2004, 1977.

[43] Kauppila, J., Sternberg, A., Alles, M., Francis, A., Holmes, J., Amusan, O., and Massengill, L., "A bias-dependent single-event compact model implemented into BSIM4 and a 90 nm CMOS process design kit," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3152–3157, 2009.

[44] Keskitalo, N., Jonsson, P., Nordgren, K., and Bleichner, H., "Temperature and injection dependence of the Shockley-Read-Hall lifetime in electron-irradiated p-type silicon," *J. Appl. Phys.*, vol. 83, no. 8, pp. 6–12, 1998.

[45] Klaassen, D., "A unified mobility model for device simulation—I. Model equations and concentration dependence," *Solid-State Electron.*, vol. 35, no. 7, pp. 953–959, 1992.

[46] Klaassen, D., "A unified mobility model for device simulation—II. Temperature dependence of carrier mobility and lifetime," *Solid-State Electron.*, vol. 35, no. 7, pp. 961–967, 1992.

[47] Knudson, A. and Campbell, A., "Charge collection in bipolar transistors," *IEEE Trans. Nucl. Sci.*, vol. 34, no. 6, pp. 1246–1250, 1987.

[48] Knudson, A., Campbell, A., Stapor, W., Hauser, J., and Jessee, M., "Charge transport by the ion shunt effect," *IEEE Trans. Nucl. Sci.*, vol. 33, no. 6, pp. 1560–1564, 1986.

[49] Krithivasan, R., Marshall, P., Nayeem, M., Sutton, A., Kuo, W., Haugerud, B., Najafizadeh, L., Cressler, J., Carts, M., Marshall, C., and Others, "Application of RHBD techniques to SEU hardening of third-generation SiGe HBT logic circuits," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3400–3407, 2006.

[50] Krithivasan, R., Niu, G., Cressler, J., Currie, S., Fritz, K., Reed, R., Marshall, P., Riggs, P., Randall, B., and Gilbert, B., "An SEU hardening approach for high-speed SiGe HBT digital logic," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 6, pp. 2126–2134, 2003.

[51] Lacoe, R., Osborn, J., Mayer, D., Brown, S., and Hunt, D., "Total-dose radiation tolerance of a commercial 0.35 μm CMOS process," in *Proc. IEEE Radiation Effects Data Workshop*, pp. 104–110, 1998.

[52] Lee, S., Jagannathan, B., Narasimha, S., Chou, A., Zamdmer, N., Johnson, J., Williams, R., Wagner, L., Kim, J., Plouchart, J., and Others, "Record RF performance of 45-nm SOI CMOS technology," in *Proc. Int. Electron Device Meeting*, pp. 255–258, 2007.

[53] Liang, Q., Cressler, J., Niu, G., Malladi, R., Newton, K., and Harame, D., "A physics-based high-injection transit-time model applied to barrier effects in SiGe HBTs," *IEEE Trans. Electron Devices*, vol. 49, no. 10, pp. 1807–1813, 2002.

[54] Lu, Y., Cressler, J., Krithivasan, R., Li, Y., Reed, R., Marshall, P., Polar, C., Freeman, G., and Ahlgren, D., "Proton tolerance of third-generation, 0.12 μm 185 GHz SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 6, pp. 1811–1815, 2003.

[55] Macdonald, D., "Impact of nickel contamination on carrier recombination in n- and p-type crystalline silicon wafers," *Appl. Phys. A*, vol. 81, no. 8, pp. 1619–1625, 2005.

[56] Macdonald, D., Cuevas, A., Rein, S., Lichtner, P., and Glunz, S., "Temperature-and injection-dependent lifetime spectroscopy of copper-related defects in silicon," in *Proc. Photovoltaic Energy Conversion Conf.*, pp. 87–90, 2003.

[57] Macdonald, D., Kerr, M., and Cuevas, A., "Boron-related minority-carrier trapping centers in p-type silicon," *Appl. Phys. Lett.*, vol. 75, no. 11, p. 1571, 1999.

[58] Madan, A., Phillips, S., and Cressler, J., "Impact of Proton Irradiation on the RF Performance of 65 nm SOI CMOS Technology," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 4, pp. 1914–1919, 2009.

[59] Madan, A., Verma, R., and Arora, R., "The enhanced role of shallow-trench isolation in ionizing radiation damage of 65 nm RF-CMOS on SOI," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3256–3261, 2009.

[60] Marshall, P. W. and Marshall, C., "Proton effects and test issues for satellite designers - Part A: Ionization effects," in *1999 IEEE NSREC Short Course Notes*, pp. IV–50–IV–110, 1999.

[61] Marshall, P. W. and Marshall, C., "Proton effects and test issues for satellite designers - Part B: Displacement damage effects," *1999 IEEE NSREC Short Course Notes*, pp. IV–50–IV–110, 1999.

[62] Marshall, P., Carts, M., Campbell, A., McMorrow, D., Buchner, S., Stewart, R., Randall, B., Gilbert, B., and Reed, R., "Single event effects in circuit-hardened SiGe HBT logic at gigabit per second data rates," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pp. 2669–2674, 2000.

[63] Mavis, D. G. and Eaton, P. H., "SEU and SET Modeling and Mitigation in Deep Submicron Technologies," in *Proc. Int. Reliability Phys. Symposium*, pp. 293–305, 2007.

[64] McMorrow, D., Lotshaw, W., Melinger, J., Buchner, S., and Pease, R., "Subbandgap laser-induced single event effects: Carrier generation via two-photon absorption," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 3002–3008, 2002.

[65] Moen, K., Najafizadeh, L., Seungwoo, J., Raman, A., Turowski, M., and Cressler, J., "Accurate Modeling of Single-Event Transients in a SiGe Voltage Reference Circuit," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 3, pp. 877–884, 2011.

[66] Moen, K., Phillips, S., Wilcox, E., Cressler, J., Nayfeh, H., Sutton, A., Warner, J., Buchner, S., McMorrow, D., Vizkelethy, G., and Dodd, P., "Evaluating the Influence of Various Body-Contacting Schemes on Single Event Transients in 45-nm SOI CMOS," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 6, pp. 3366–3372, 2010.

[67] Moen, K. A. and Cressler, J. D., "Measurement and Modeling of Carrier Transport Parameters Applicable to SiGe BiCMOS Technology Operating in Extreme Environments," *IEEE Trans. Electron Devices*, vol. 57, no. 3, pp. 551–561, 2010.

[68] Moen, K. A., Phillips, S. D., Kenyon, E. W., and Cressler, J. D., "Establishing Best-Practice Modeling Approaches for Understanding Single-Event Transients in Gb/s SiGe Digital Logic," *IEEE Trans. Nucl. Sci.*, in press.

[69] Moen, K. A., Yuan, J., Chakraborty, P. S., Bellini, M., Cressler, J. D., Ho, H., Yasuda, H., and Wise, R., "Improved 2-D Regional Transit Time Analysis for Optimized Scaling of SiGe HBTs," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 257–260, 2010.

[70] Montes, E. J., Reed, R. A., Pellish, J. A., Alles, M. L., Schrimpf, R. D., Weller, R. A., Varadharajaperumal, M., Niu, G., Sutton, A. K., Diestelhorst, R., Espinel, G., Krithivasan, R., Comeau, J. P., Cressler, J. D., Marshall, P. W., and Vizkelethy, G., "Single event upset mechanisms for low-energy-deposition events in SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 3, pp. 1581–1586, 2008.

[71] Munteanu, D. and Autran, J., "Modeling and simulation of single-event effects in digital devices and ICs," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 4, pp. 1854–1878, 2008.

[72] Najafizadeh, L., Phillips, S. D., Moen, K. a., Diestelhorst, R. M., Bellini, M., Saha, P. K., Cressler, J. D., Vizkelethy, G., Turowski, M., Raman, A., and Marshall, P. W., "Single Event Transient Response of SiGe Voltage References and Its Impact on the Performance of Analog and Mixed-Signal Circuits," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3469–3476, 2009.

[73] Najafizadeh, L., Sutton, A., Diestelhorst, R., Bellini, M., Jun, B., Cressler, J., Marshall, P., and Marshall, C., "A comparison of the effects of x-ray and proton irradiation on the performance of SiGe precision voltage references," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 6, pp. 2238–2244, 2007.

[74] Narasimha, S., Onishi, K., Nayfeh, H., Waite, A., Weybright, M., Johnson, J., Fonseca, C., Corliss, D., Robinson, C., Crouse, M., and Others, "High performance 45-nm SOI technology with enhanced strain, porous low-k BEOL, and immersion lithography," in *Proc. Int. Electron Device Meeting*, pp. 1–4, 2006.

[75] Niu, G., Cressler, J., Shoga, M., Jobe, K., Chu, P., and Harame, D., "Simulation of SEE-induced charge collection in UHV/CVD SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pp. 2682–2689, 2000.

[76] Niu, G., Krithivasan, R., Cressler, J., Marshall, P., Marshall, C., Reed, R., and Harame, D., "Modeling of single-event effects in circuit-hardened high-speed SiGe HBT logic," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 6, pp. 1849–1854, 2001.

[77] Ogawa, S. and Shiono, N., "Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO2 interface," *Physical Review B*, vol. 51, no. 7, pp. 4218–4230, 1995.

[78] Pease, R., "Modeling single event transients in bipolar linear circuits," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 4, pp. 1879–1890, 2008.

[79] Pellish, J., Reed, R., McMorrow, D., Melinger, J., Jenkins, P., Sutton, A., Diestelhorst, R., Phillips, S., Cressler, J., Pouget, V., and Others, "Laser-induced current transients in silicon-germanium HBTs," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 6, pp. 2936–2942, 2008.

[80] Pellish, J., Reed, R., McMorrow, D., Vizkelethy, G., Cavrois, V., Baggio, J., Paillet, P., Duhamel, O., Moen, K., Phillips, S., and Others, "Heavy ion microbeam-and broadbeam-induced transients in SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3078–3084, 2009.

[81] Pellish, J. A., Reed, R. A., Schrimpf, R. D., Alles, M. L., Varadharajaperumal, M., Niu, G., Sutton, A. K., Diestelhorst, R. M., Espinel, G., Krithivasan, R., Comeau, J. P., Cressler, J. D., Vizkelethy, G., Marshall, P. W., Weller, R. A., Mendenhall, M. H., and Montes, E. J., "Substrate engineering concepts to mitigate charge collection in deep trench isolation technologies," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3298–3305, 2006.

[82] Phillips, S., Thrivikraman, T., Appaswamy, A., Sutton, A., Cressler, J., Vizkelethy, G., Dodd, P., and Reed, R., "A novel device architecture for SEU mitigation: the inverse-mode cascode SiGe HBT," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3393–3401, 2009.

[83] Rashkeev, S., Fleetwood, D., Schrimpf, R., and Pantelides, S., "Effects of hydrogen motion on interface trap formation and annealing," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp. 3158–3165, 2004.

[84] Reed, R., Marshall, P., Pickel, J., Carts, M., Fodness, B., Niu, G., Fritz, K., Vizkelethy, G., Dodd, P., Irwin, T., and Others, "Heavy-ion broad-beam and microprobe studies of single-event upsets in 0.20-μm SiGe heterojunction bipolar transistors and circuits," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 6, pp. 2184–2190, 2003.

[85] Rein, S., *Lifetime Spectroscopy: A Method of Defect Characterization in Silicon for Photovoltaic Applications.* Springer Series in Materials Science, Springer, 2005.

[86] Rein, S. and Glunz, S. W., "Electronic properties of the metastable defect in boron-doped Czochralski silicon: Unambiguous determination by advanced lifetime spectroscopy," *Appl. Phys. Lett.*, vol. 82, no. 7, p. 1054, 2003.

[87] Rein, S. and Glunz, S. W., "Electronic properties of interstitial iron and iron-boron pairs determined by means of advanced lifetime spectroscopy," *J. Appl. Phys.*, vol. 98, no. 11, p. 113711, 2005.

[88] Roche, P., Jacquet, F., Caillat, C., and Schoellkopf, J.-P., "An alpha immune and ultra low neutron SER high density SRAM," in *Proc. Int. Reliability Phys. Symposium*, pp. 671–672, 2004.

[89] Rosenits, P., Roth, T., Glunz, S. W., and Beljakowa, S., "Determining the defect parameters of the deep aluminum-related defect center in silicon," *Appl. Phys. Lett.*, vol. 91, no. 12, p. 122109, 2007.

[90] Sadovnikov, A., "Influence of the extrinsic base on the base current kink in SiGe BJTs," *Applied Surface Science*, vol. 224, no. 1-4, pp. 320–323, 2004.

[91] Schenk, A., Altermatt, P., and Schmithusen, B., "Physical model of incomplete ionization for silicon device simulation," in *Proc. Simulation Semi. Processes and Devices Conf.*, pp. 51–54, 2006.

[92] Schmidt, J., Bothe, K., and Hezel, R., "Oxygen-related minority-carrier trapping centers in p-type Czochralski silicon," *Appl. Phys. Lett.*, vol. 80, no. 23, p. 4395, 2002.

[93] Schwank, J., Ferlet-Cavrois, V., Shaneyfelt, M., Paillet, P., and Dodd, P., "Radiation effects in SOI technologies," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 3, pp. 522–538, 2003.

[94] Schwank, J., Shaneyfelt, M., McMorrow, D., Ferlet-Cavrois, V., Dodd, P., Heidel, D., Marshall, P., Pellish, J., LaBel, K., Rodbell, K., and Others, "Estimation of heavy-ion LET thresholds in advanced SOI IC technologies from two-photon absorption laser measurements," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 4, pp. 1827–1834, 2010.

[95] Seth, S., Thrivikraman, T., Cheng, P., Cressler, J. D., Babcock, J. A., and Buchholz, A., "A large-signal RF reliability study of complementary SiGe HBTs on SOI intended for use in wireless applications," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 133–136, 2010.

[96] Shi, Y. and Niu, G., "Vertical profile design and transit time analysis of nano-scale SiGe HBTs for terahertz f_T," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 213–216, 2004.

[97] Shi, Y. and Niu, G., "2-D analysis of device parasitics for 800/1000 GHz f_T/f_{max} SiGe HBT," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 252–255, 2005.

[98] Shockley, W. and Read, W., "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, 1952.

[99] Slotboom, J., *Analysis of Bipolar Transistors*. PhD thesis, Eindhoven, Technische Hogeschool, 1977.

[100] "SRIM: The Stopping and Range of Ions in Matter." [Online]. Available: `http://www.srim.org`.

[101] Stesmans, A. and Van Gorp, G., "Maximum density of Pb centers at the (111) Si/SiO2 interface after vacuum anneal," *Appl. Phys. Lett.*, vol. 57, no. 25, p. 2663, 1990.

[102] Sutton, A., Moen, K., Cressler, J., Carts, M., Marshall, P., Pellish, J., Ramachandran, V., Reed, R., Alles, M., and Niu, G., "Proton-induced SEU in SiGe digital logic at cryogenic temperatures," *Solid-State Electron.*, vol. 52, no. 10, pp. 1652–1659, 2008.

[103] Sutton, A., Bellini, M., Cressler, J., Pellish, J., Reed, R., Marshall, P., Niu, G., Vizkelethy, G., Turowski, M., and Raman, A., "An evaluation of transistor-layout RHBD techniques for SEE mitigation in SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 6, pp. 2044–2052, 2007.

[104] Sutton, A., Haugerud, B., Lu, Y., Kuo, W., Cressler, J., Marshall, P., Reed, R., Rieh, J., Freeman, G., and Ahlgren, D., "Proton tolerance of fourth-generation 350 GHz UHV/CVD SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp. 3736–3742, 2004.

[105] Synopsys Inc., "Sentaurus Workbench User Guide," 2011.

[106] Turowski, M., Raman, A., and Jablonski, G., "Mixed-mode simulation and analysis of digital single event transients in fast CMOS ICs," in *Proc. Mixed Design of Integrated Circuits & Systems Conference*, pp. 433–438, 2007.

[107] Turowski, M., Pellish, J. A., Moen, K. A., Raman, A., Cressler, J. D., Reed, R. A., and Niu, G., "Reconciling 3-D mixed-mode simulations and measured single-event transients in SiGe HBTs," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 6, pp. 3342–3348, 2010.

[108] Turowski, M., Raman, A., and Fedoseyev, A., "Mixed-mode simulation of single event upsets in modern SiGe BiCMOS mixed-signal circuits," in *Proc. Mixed Design of Integrated Circuits & Systems Conference*, pp. 462–467, 2009.

[109] Van Den Biesen, J., "A simple regional analysis of transit times in bipolar transistors," *Solid-State Electron.*, vol. 29, no. 5, pp. 529–534, 1986.

[110] Vanhoucke, T., Hurkx, G., Panko, D., Campos, R., Piontek, A., Palestri, P., and Selmi, L., "Physical Description of the Mixed-Mode Degradation Mechanism for High Performance Bipolar Transistors," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 1–4, 2006.

[111] Varadharajaperumal, M., Niu, G., Wei, X., Zhang, T., Cressler, J., Reed, R., and Marshall, P., "3-D simulation of SEU hardening of SiGe HBTs using shared dummy collector," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 6, pp. 2330–2337, 2007.

[112] Wei, X., Zhang, T., Niu, G., Varadharajaperumal, M., Cressler, J., and Marshall, P., "3-D mixed-mode simulation of single event transients in SiGe HBT emitter followers and resultant hardening guidelines," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 6, pp. 3360–3366, 2008.

[113] Yuan, J., *Cryogenic Operation of Silicon-Germanium Heterojunction Bipolar Transistors and its Relation to Scaling and Optimization.* PhD thesis, Georgia Institute of Technology, Atlanta, GA, 2010.

[114] Yuan, J. and Cressler, J., "A novel superjunction collector design for improving breakdown voltage in high-speed SiGe HBTs," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 75–78, 2009.

[115] Yuan, J., Cressler, J., Krithivasan, R., Thrivikraman, T., Khater, M., Ahlgren, D., Joseph, A., and Rieh, J., "On the performance limits of cryogenically operated SiGe HBTs and its relation to scaling for terahertz speeds," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1007–1019, 2009.

[116] Yuan, J., Cressler, J. D., Moen, K. A., and Chakraborty, P. S., "An investigation of collector-base transport in SiGe HBTs designed for half-terahertz speeds," in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 157–160, 2010.

[117] Zerounian, N., Garcia, E., Aniel, F., Chevalier, P., Geynet, B., and Chantre, A., "SiGe HBT featuring f_T greater than 600GHz at cryogenic temperature," in *ECS Trans.*, vol. 153, pp. 1069–1077, 2008.

[118] Zhang, G., Cressler, J. D., Niu, G., and Joseph, A. J., "A New "Mixed-Mode" Reliability Degradation Mechanism in Advanced Si and SiGe bipolar transistors," *IEEE Trans. Electron Devices*, vol. 49, no. 12, pp. 2151–2156, 2002.

[119] Zhu, C., Liang, Q., Al-Huq, R., Cressler, J., Joseph, A., Johansen, J., Freeman, G., and Ahlgren, D., "An investigation of the damage mechanisms in impact ionization-induced "mixed-mode" reliability stressing of scaled SiGe HBTs," in *Proc. Int. Electron Device Meeting*, pp. 7.8.1–7.8.4, 2003.

[120] Zhu, C., Liang, Q., Al-Huq, R., Cressler, J., and Joseph, A., "Damage mechanisms in impact-ionization-induced mixed-mode reliability degradation of SiGe HBTs," *IEEE Trans. Device Mater. Rel.*, vol. 5, no. 1, pp. 142–149, 2005.

VITA

Kurt A. Moen was born in Long Beach, CA in 1984. He received the B.S. degree in electrical engineering and computer science in 2005 from the University of California, Berkeley. As an undergraduate student, he worked in the research group of Dr. Costas Spanos on the development of a novel photolithography metrology tool. He also worked independently under Dr. Nathan Cheung on the design and optimization of nanowire MOSFETs through device TCAD simulations.

In 2006, Kurt joined Dr. John D. Cressler's SiGe research team at the Georgia Institute of Technology. At Georgia Tech, he earned the M.S. degree in electrical and computer engineering in 2008, where his M.S. thesis addressed the measurement and modeling of semiconductor carrier transport parameters across temperature and under radiation exposure.

As a continuing Ph.D. student in Dr. Cressler's team, Kurt was awarded the Texas Instruments / Global Research Collaboration (TI/GRC) graduate fellowship by the Semiconductor Research Corporation in 2010. His Ph.D. research focused on performance-scaling and optimization of SiGe HBTs, radiation effects in SiGe HBTs and SOI MOSFETs, and particularly the application of novel TCAD techniques to understand and mitigate reliability challenges in advanced SiGe BiCMOS technologies. In 2011, he interned in the Analog Technology Development group at Texas Instruments in Dallas, TX. Following the completion of his Ph.D., he will begin employment as a Senior Device Engineer at TowerJazz Semiconductor in Newport Beach, CA.

Predictive Modeling of Device and Circuit Reliability in Highly Scaled CMOS and SiGe BiCMOS Technology

Kurt A. Moen

157 Pages

Directed by Professor John D. Cressler

The advent of high-frequency silicon-based technologies has enabled the design of mixed-signal circuits that incorporate analog, RF, and digital circuit components to build cost-effective system-on-a-chip solutions. Emerging applications provide great incentive for continued scaling of transistor performance, requiring careful attention to mismatch, noise, and reliability concerns. If these mixed-signal technologies are to be employed within space-based electronic systems, they must also demonstrate reliability in radiation-rich environments. SiGe BiCMOS technology in particular is positioned as an excellent candidate to satisfy all of these requirements. The objective of this research is to develop predictive modeling tools that can be used to design new mixed-signal technologies and assess their reliability on Earth and in extreme environments. Ultimately, the goal is to illuminate the interaction of device- and circuit-level reliability mechanisms and establish best practices for modeling these effects in modern circuits. To support this objective, several specific areas have been targeted first, including a TCAD-based approach to identify performance-limiting regions in SiGe HBTs, measurement and modeling of carrier transport parameters that are essential for predictive TCAD, and measurement of device-level single-event transients to better understand the physical origins and implications for device design. These tasks provide the foundation for the bulk of this research, which addresses circuit-level reliability challenges through the application of novel mixed-mode TCAD techniques. All of the individual tasks are tied together by a guiding theme: to develop a holistic understanding of the challenges faced by emerging broadband technologies by coordinating results from material, device, and circuit studies.