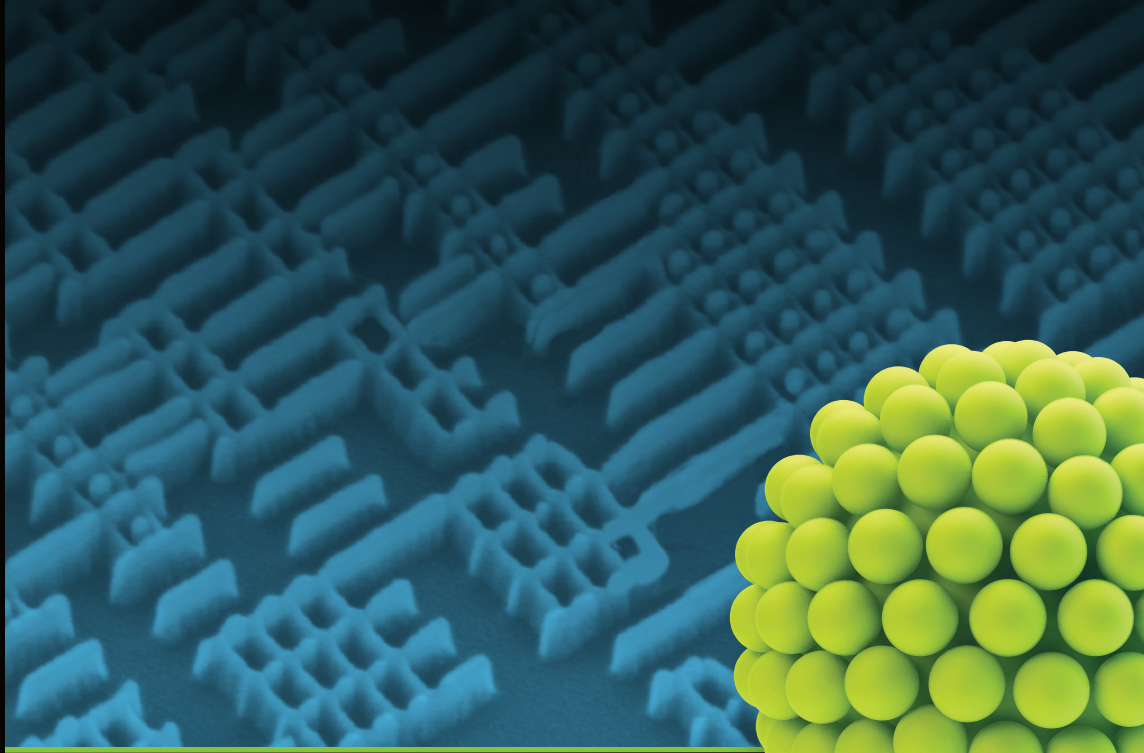


Micro- and Nanoelectronics

Emerging Device Challenges and Solutions

EDITED BY **Tomasz Brozek**

CRC Press
Taylor & Francis Group

Micro- and Nanoelectronics

Emerging Device Challenges and Solutions

Devices, Circuits, and Systems

Series Editor
Krzysztof Iniewski
CMOS Emerging Technologies Research Inc.,
Vancouver, British Columbia, Canada

PUBLISHED TITLES:

Atomic Nanoscale Technology in the Nuclear Industry
Taeho Woo

Biological and Medical Sensor Technologies
Krzysztof Iniewski

Building Sensor Networks: From Design to Applications
Ioanis Nikolaidis and Krzysztof Iniewski

Circuits at the Nanoscale: Communications, Imaging, and Sensing
Krzysztof Iniewski

Design of 3D Integrated Circuits and Systems
Rohit Sharma

Electrical Solitons: Theory, Design, and Applications
David Ricketts and Donhee Ham

Electronics for Radiation Detection
Krzysztof Iniewski

Embedded and Networking Systems:
Design, Software, and Implementation
Gul N. Khan and Krzysztof Iniewski

Energy Harvesting with Functional Materials and Microsystems
Madhu Bhaskaran, Sharath Sriram, and Krzysztof Iniewski

Graphene, Carbon Nanotubes, and Nanostuctures:
Techniques and Applications
James E. Morris and Krzysztof Iniewski

High-Speed Devices and Circuits with THz Applications
Jung Han Choi

High-Speed Photonics Interconnects
Lukas Chrostowski and Krzysztof Iniewski

High Frequency Communication and Sensing:
Traveling-Wave Techniques
Ahmet Tekin and Ahmed Emira

Integrated Microsystems: Electronics, Photonics, and Biotechnology
Krzysztof Iniewski

PUBLISHED TITLES:

Integrated Power Devices and TCAD Simulation
Yue Fu, Zhanming Li, Wai Tung Ng, and Johnny K.O. Sin

Internet Networks: Wired, Wireless, and Optical Technologies
Krzysztof Iniewski

Labs on Chip: Principles, Design, and Technology
Eugenio Iannone

Low Power Emerging Wireless Technologies
Reza Mahmoudi and Krzysztof Iniewski

Medical Imaging: Technology and Applications
Troy Farncombe and Krzysztof Iniewski

Metallic Spintronic Devices
Xiaobin Wang

MEMS: Fundamental Technology and Applications
Vikas Choudhary and Krzysztof Iniewski

Micro- and Nanoelectronics: Emerging Device Challenges and Solutions
Tomasz Brozek

Microfluidics and Nanotechnology: Biosensing to the Single Molecule Limit
Eric Lagally

MIMO Power Line Communications: Narrow and Broadband Standards, EMC, and Advanced Processing
Lars Torsten Berger, Andreas Schwager, Pascal Pagani, and Daniel Schneider

Mobile Point-of-Care Monitors and Diagnostic Device Design
Walter Karlen

Nano-Semiconductors: Devices and Technology
Krzysztof Iniewski

Nanoelectronic Device Applications Handbook
James E. Morris and Krzysztof Iniewski

Nanopatterning and Nanoscale Devices for Biological Applications
Šeila Selimović

Nanoplasmonics: Advanced Device Applications
James W. M. Chon and Krzysztof Iniewski

Nanoscale Semiconductor Memories: Technology and Applications
Santosh K. Kurinec and Krzysztof Iniewski

Novel Advances in Microsystems Technologies and Their Applications
Laurent A. Francis and Krzysztof Iniewski

Optical, Acoustic, Magnetic, and Mechanical Sensor Technologies
Krzysztof Iniewski

Optical Fiber Sensors: Advanced Techniques and Applications
Ginu Rajan

PUBLISHED TITLES:

Organic Solar Cells: Materials, Devices, Interfaces, and Modeling
Qiquan Qiao

Radiation Effects in Semiconductors
Krzysztof Iniewski

Semiconductor Radiation Detection Systems
Krzysztof Iniewski

Smart Grids: Clouds, Communications, Open Source, and Automation
David Bakken

Smart Sensors for Industrial Applications
Krzysztof Iniewski

Solid-State Radiation Detectors: Technology and Applications
Salah Awadalla

Technologies for Smart Sensors and Sensor Fusion
Kevin Yallup and Krzysztof Iniewski

Telecommunication Networks
Eugenio Iannone

Testing for Small-Delay Defects in Nanoscale CMOS Integrated Circuits
Sandeep K. Goel and Krishnendu Chakrabarty

VLSI: Circuits for Emerging Applications
Tomasz Wojcicki

Wireless Technologies: Circuits, Systems, and Devices
Krzysztof Iniewski

FORTHCOMING TITLES:

Analog Electronics for Radiation Detection
Renato Turchetta

Cell and Material Interface: Advances in Tissue Engineering, Biosensor, Implant, and Imaging Technologies
Nihal Engin Vrana

Circuits and Systems for Security and Privacy
Farhana Sheikh and Leonel Sousa

CMOS: Front-End Electronics for Radiation Sensors
Angelo Rivetti

CMOS Time-Mode Circuits and Systems: Fundamentals and Applications
Fei Yuan

Electrostatic Discharge Protection of Semiconductor Devices and Integrated Circuits
Juin J. Liou

FORTHCOMING TITLES:

Gallium Nitride (GaN): Physics, Devices, and Technology
Farid Medjdoub and Krzysztof Iniewski

Implantable Wireless Medical Devices: Design and Applications
Pietro Salvo

Laser-Based Optical Detection of Explosives
Paul M. Pellegrino, Ellen L. Holthoff, and Mikella E. Farrell

Mixed-Signal Circuits
Thomas Noulis and Mani Soma

Magnetic Sensors: Technologies and Applications
Simone Gambini and Kirill Poletkin

MRI: Physics, Image Reconstruction, and Analysis
Angshul Majumdar and Rabab Ward

Multisensor Data Fusion: From Algorithm and Architecture Design to Applications
Hassen Fourati

Nanoelectronics: Devices, Circuits, and Systems
Nikos Konofaos

Nanomaterials: A Guide to Fabrication and Applications
Gordon Harling, Krzysztof Iniewski, and Sivashankar Krishnamoorthy

Optical Imaging and Sensing: Technology, Devices, and Applications
Dongsoo Kim and Ajit Khosla

Physical Design for 3D Integrated Circuits
Aida Todri-Sanial and Chuan Seng Tan

Power Management Integrated Circuits and Technologies
Mona M. Hella and Patrick Mercier

Radiation Detectors for Medical Imaging
Jan S. Iwanczyk and Polad M. Shikhaliev

Radio Frequency Integrated Circuit Design
Sebastian Magierowski

Reconfigurable Logic: Architecture, Tools, and Applications
Pierre-Emmanuel Gaillardon

Soft Errors: From Particles to Circuits
Jean-Luc Autran and Daniela Munteanu

Wireless Transceiver Circuits: System Perspectives and Design Aspects
Woogeun Rhee

Micro- and Nanoelectronics

Emerging Device Challenges and Solutions

EDITED BY **Tomasz Brozek**
PDF Solutions Inc., San Jose, California, USA

Krzysztof Iniewski MANAGING EDITOR
CMOS Emerging Technologies Research Inc.
Vancouver, British Columbia, Canada

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20140904

International Standard Book Number-13: 978-1-4822-1491-8 (eBook - PDF)

Contents

Preface

The progress made in microelectronics in recent decades is unimaginable. We have been witnessing unbelievable, and so far, undisturbed advances in device scaling and growth of integrated circuits in both functionality and complexity. Myths about the limits of miniaturization and scalability of transistors have been shattered one after another; roadblocks have been removed by power of human invention, fueled by market growth. The cost of a single transistor on a silicon chip during the last 40 years has decreased more than a millionfold, as the number of devices on a single chip increased more than a millionfold, following the exponential dependence commonly known as Moore's law. Many people question whether the semiconductor industry can achieve, or even afford, further scaling, especially because numerous challenges arise from both device physics and manufacturing capabilities perspectives.

Traditional scaling based on reduction of physical dimensions of metal–oxide–semiconductor transistors, with simultaneous reduction of supply voltages and dissipated power, is reaching its limits. It has been extended by innovations in device architecture, introduction of new materials, and inventive techniques of patterning features of nanometer size. This decreasing size comes at an increasing cost, so most recent new technology does not offer much advantage in terms of cost per transistor, or cost per function, when compared to previous nodes.

Paralleling the efforts to sustain the scaling from micrometer to nanometer dimensions, the industry has developed a new paradigm, known as *More-than-Moore*, where the increased value to devices comes not from rigorous execution of Moore's law, but from added functionalities that are not necessarily scalable in the same way. This segment covers devices for analog/radio frequency functionalities, passives, high-voltage operation, microelectromechanical sensors, other sensors, and so on. They are most commonly supported by silicon technology platforms compatible with complementary metal–oxide–semiconductor (CMOS), often combining new materials, device architectures, or both.

Taking a longer view over the time horizon, researchers explore possible solutions to enable the next technology nodes, making transistors smaller and faster. This will require further innovations in device architectures and new materials. Still, CMOS scaling will reach inevitable limits, and new concepts will drive progress into the nanoelectronic era. What this new, post-CMOS world will be is difficult to answer today, but many ideas have already been developed and are being investigated in universities and research organizations across the world.

This book attempts to address some of the above topics. It looks at the challenges of today's microelectronics and discusses paths into the nanoelectronic world. This is achieved through a collection of works by renowned scholars and industry gurus in the field of micro- and nanoelectronics, all of whom contributed to this volume. The book is intended for researchers working on emerging scaling topics, for engineers resolving issues of advanced technologies, and for graduate students improving their understanding of barriers and opportunities in future

progress in micro- and nanoelectronics. We hope that all of them find it useful and inspiring.

In preparing this book, we focused on solutions developed in response to scaling problems of current silicon technologies, as well as those that may pave the way for future progress. We reached out to leading experts, researchers, and specialists from academia, research institutes, and industry. Their contributions are presented in 16 dedicated chapters, covering a wide range of topics. Many of the chapters have been written as an original contribution specifically for this book, and some of them are based on ideas most recently presented by their authors at leading conferences, like the IEEE International Electron Devices Meeting (IEDM), European Solid-State Device Research Conference (ESSDERC), International Reliability Physics Symposium (IRPS), Very-Large-Scale Integration (VLSI) Technology Symposium, and CMOS Emerging Technologies Symposium.

The book is organized into four sections, each covering a specific area with several dedicated chapters. Section I, Silicon Compound Devices, focuses on some of the key topics of technologies for More-than-Moore solutions with silicon–germanium (SiGe) and silicon carbide (SiC) materials. It contains contributions from TowerJazz on SiGe BiCMOS technology, research from the University of British Columbia on SiGe and SiGe:C device issues, and a study of SiC device problems from Simon Fraser University.

Section II addresses today's most advanced silicon devices. It contains an opening chapter from IBM researchers, introducing the concept of fully depleted devices: FinFETs and planar fully depleted silicon on insulator (FDSOI) transistors. The following chapters further explain the benefits and challenges of new types of CMOS devices that have made inroads into leading manufacturing facilities across the world. Experts from SOITEC explain the benefits of planar FDSOI devices; the in-depth study from GLOBALFOUNDRIES describes bulk-FinFET and its maturity, while research from IBM addresses the reliability consequences of changing device architecture. The last two chapters of this part look at future devices. The work from IMEC presents results of research on high-mobility channels for CMOS devices, while a study from the University of California investigates transistors built on a compound semiconductor on insulator platform.

Section III, Post-CMOS Device Concepts, explores options for future directions in micro- and nanoelectronics. The introductory chapter, from GLOBALFOUNDRIES, gives a good review of possible paths and choices regarding where microelectronic devices may be headed. The next chapters, by researchers from the Vienna University of Technology on spin-transfer torque magnetoresistive random-access memory (STT-MRAM)-based logic, from Virginia Commonwealth University on spintronics–straintronics, and from Tohoku University on nanoionic switches present in-depth studies pertaining to possible future solutions.

The final three chapters, in Section IV, Elements of Carbon Electronics, offer an insight into possible solutions based on graphene and carbon nanotubes. The chapter by researchers from the University of Texas summarizes their work on graphene device modeling, while the contribution from the Delft University of Technology presents the results of a study on carbon nanotubes for future interconnects. The final

chapter, from Hong Kong Polytechnic University, investigates material properties of ultrathin graphene layers.

Working on this book was an exciting adventure, and we would like to thank the authors and our collaborators for contributing to this collection and sharing the results of their knowledge and their work with all of us. Micro- and nanoelectronics are fast-paced areas, so any effort to capture them in a steady state or slow-motion mode is an almost impossible task—but even attempting to achieve that was a truly rewarding experience. We hope to continue this work and keep capturing such "snapshots" from the research as we struggle to move further up the steep slopes of Moore's law.

Tomasz Brozek
PDF Solutions

Editor

Tomasz Brozek is a technical fellow and director at PDF Solutions, San Jose, California. He was born and grew up in Poland. He graduated with a master's degree in electrical engineering from Lvov Polytechnic Institute and received his PhD in physics from the Institute of Semiconductors in Kiev, Ukraine. Dr. Brozek was an assistant professor at Warsaw University of Technology, Poland. He taught academic courses and conducted research first in Warsaw and then at the University of California, Los Angeles, in the area of metal–oxide–semiconductor device physics, gate dielectric characterization, radiation effects, and plasma-induced damage. During his industrial tenure, he worked at Motorola R&D organizations in Texas and Arizona, focusing on technology characterization, with emphasis on process-induced damage, and reliability.

Dr. Brozek joined PDF Solutions in 2000. He is responsible for advanced node technology characterization, methods development, and early yield ramp. Since joining the company, he has led multiple projects focusing on advanced technology nodes of logic, memory (DRAM, Flash), and image sensor processes, ranging from 0.25 μm down to 10 nm technology nodes. Dr. Brozek has had more than 50 papers and conference presentations published and holds several patents.

Krzysztof (Kris) Iniewski is manager of R&D at Redlen Technologies Inc., a start-up company in Vancouver, Canada. Redlen's revolutionary production process for advanced semiconductor materials enables a new generation of more accurate, all-digital, radiation-based imaging solutions. Kris is also a president of CMOS Emerging Technologies Research Inc. (www.cmosetr.com), an organization of high-tech events covering communications, microsystems, optoelectronics, and sensors. During his career, Dr. Iniewski has held numerous faculty and management positions at the University of Toronto, University of Alberta, Simon Fraser University, and PMC-Sierra Inc. He has published over 100 research papers in international journals and conferences. He holds 18 international patents granted in the United States, Canada, France, Germany, and Japan. He is a frequent invited speaker and has consulted for multiple organizations internationally. He has written and edited several books for CRC Press, Cambridge University Press, IEEE Press, Wiley, McGraw-Hill, Artech House, and Springer. His personal goal is to contribute to healthy living and sustainability through innovative engineering solutions. In his leisure time, Kris can be found hiking, sailing, skiing, or biking in beautiful British Columbia. He can be reached at kris.iniewski@gmail.com.

Contributors

Deji Akinwande
Department of Electrical and Computer Engineering
Texas Materials Institute
University of Texas
Austin, Texas

Frederic Allibert
Soitec
Bernin, France

Jayasimha Atulasimha
Department of Mechanical and Nuclear Engineering
Virginia Commonwealth University School of Engineering
Richmond, Virginia

Supriyo Bandyopadhyay
Department of Electrical and Computer Engineering
Virginia Commonwealth University
Richmond, Virginia

A.F. Basile
Department of Physics
University of Bologna
Bologna, Italy

C.I.M. Beenakker
Delft Institute of Microsystems and Nanotechnology
Delft University of Technology
Delft, The Netherlands

Eduard A. Cartier
IBM
Thomas J. Watson Research Center
Yorktown Heights, New York

Yang Chai
Department of Applied Physics
The Hong Kong Polytechnic University
Hung Hom, Hong Kong

An Chen
GLOBALFOUNDRIES Inc.
Santa Clara, California

Kangguo Cheng
IBM SDRC
Albany, New York

Nadine Collaert
IMEC
Heverlee, Belgium

Yuanwei Dong
Department of Materials Engineering
University of British Columbia
Vancouver, Canada

Bruce Doris
IBM Reserch
Albany, New York

Noel D'Souza
Department of Mechanical and Nuclear Engineering
Virginia Commonwealth University
Richmond, Virginia

Hui Fang
Department of Materials Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, Illinois

Terence Hook
IBM SDRC
Essex Junction, Vermont

R. Ishihara
Delft Institute of Microsystems and Nanotechnology
Delft University of Technology
Delft, The Netherlands

Rehan Kapadia
Ming Hsieh
Department of Electrical Engineering
University of Southern California
Los Angeles, California

Ali Khakifirooz
Spansion
Sunnyvale, California

Siddarth A. Krishnan
IBM
Semiconductor Research and Development Center
Hopewell Junction, New York

Barry P. Linder
IBM
Thomas J. Watson Research Center
Yorktown Heights, New York

Hiwa Mahmoudi
Institute for Microelectronics
Technical University Vienna
Vienna, Austria

Christophe Marville
Soitec
Bernin, France

Witek P. Maszara
GLOBALFOUNDRIES Inc.
Santa Clara, California

Carlos Mazure
Soitec
Bernin, France

Patricia M. Mooney
Department of Physics
Simon Fraser University
Burnaby, Canada

Vijay Narayanan
IBM
Thomas J. Watson Research Center
Yorktown Heights, New York

Bich-Yen Nguyen
Soitec
Bernin, France

Takeo Ohno
WPI Advanced Institute for Materials Research
Tohoku University
Sendai, Japan

Kristen Parrish
Kilby Labs
Texas Instruments
Dallas, Texas

Edward Preisler
TowerJazz
Newport Beach, California

Marco Racanelli
TowerJazz
Newport Beach, California

Siegfried Selberherr
Institute for Microelectronics
Technical University Vienna
Vienna, Austria

Viktor Sverdlov
Institute for Microelectronics
Technical University Vienna
Vienna, Austria

Kuniharu Takei
Department of Physics and Electronics
Osaka Prefecture University
Osaka, Japan

Chunyan E. Tian
IBM
Semiconductor Research and
Development Center
Hopewell Junction, New York

S. Vollebregt
Delft Institute of Microsystems and
Nanotechnology
Delft University of Technology
Delft, The Netherlands

Thomas Windbacher
Institute for Microelectronics
Technical University Vienna
Vienna, Austria

Guangrui (Maggie) Xia
Department of Materials
Engineering
University of British Columbia
Vancouver, British Columbia

Yuda Zhao
Department of Applied Physics
The Hong Kong Polytechnic
University
Hung Hom, Hong Kong

Section I

Silicon Compound Devices

1 SiGe BiCMOS Technology and Devices

Edward Preisler and Marco Racanelli

CONTENTS

1.1 INTRODUCTION

Over the past decade, the silicon–germanium bicomplementary metal–oxide semiconductor (SiGe BiCMOS) has evolved into a dominant technology for the implementation of radio-frequency (RF) circuits. By providing performance, power consumption, and noise advantages over standard complementary metal–oxide–semiconductor (CMOS) transistor technology while leveraging the same manufacturing infrastructure, SiGe BiCMOS technologies can offer a cost-effective solution for challenging RF and analog circuit applications. Today, many cell phones, wireless local area network (WLAN) devices, global positioning system (GPS) receivers, and digital TV tuners employ some SiGe BiCMOS circuitry for either RF receive or transmit functions because of these advantages. Recently, advanced-node RF CMOS has achieved performance levels that enable some of these applications to be realized in CMOS alone, but SiGe BiCMOS continues to provide advantages for many leading-edge products. These existing markets, as well as emerging applications in the use of SiGe for power amplifiers and millimeter-wave products, continue to drive SiGe technology development.

In this chapter, we review SiGe BiCMOS technology and its most significant applications. First, we provide a basic understanding of how SiGe devices achieve a performance advantage over traditional bipolar and CMOS devices. Next, we review historical application drivers for SiGe technology and project a roadmap for SiGe applications well into the future. Then, we discuss RF performance metrics

for SiGe heterostructure bipolar transistor (HBT) devices, followed by a discussion of how the devices can be optimized to maximize these performance metrics. Finally, we discuss some of the components built around SiGe devices that are part of modern SiGe BiCMOS technologies and make them suitable for advanced RF product design.

1.2 SiGe HBT DEVICE PHYSICS

SiGe HBT devices are bipolar junction transistors (BJTs), which are created using a thin epitaxial base incorporating roughly 8%–30% atomic Ge content. These devices are fabricated alongside CMOS devices with the addition of four to seven masking layers relative to a core CMOS process. SiGe HBTs derive part of their performance benefits from heterojunction effects and part from their epitaxial-base architecture. The heterojunction effects of HBTs were first described in the 1950s by Kroemer (eventually earning him a Nobel Prize) and were summarized by Kroemer in 1982 [1]. These effects arise from a combination of different materials (in this case an $Si_{1-x}Ge_x$ alloy and Si) to create a variation in the bandgap throughout the device that can be manipulated to improve performance.

Two common techniques for using heterojunction effects to improve performance are depicted in Figure 1.1, where typical doping and Ge profiles are shown along with the resulting conduction band energy profile. The first technique (see Figure 1.1a) uses a box-shaped Ge profile. This creates an offset in the conduction band energy level at the emitter-base junction (due to the lower bandgap of SiGe relative to Si), lowering the barrier for electron current flow into the base, and increasing the efficiency of electron injection into the base. The band offset in the valence band is relatively unchanged compared with a Si homojunction and thus holes in the base are injected back into the emitter at roughly the same rate as they would be injected without Ge. Greater electron injection efficiency without also increasing the

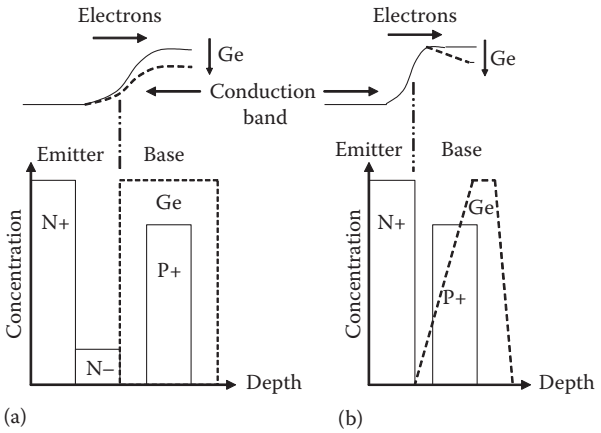


FIGURE 1.1 Common SiGe HBT doping and Ge profiles shown along with the resulting band diagrams for (a) a box Ge profile and (b) a graded Ge profile.

efficiency of the back injection of holes from the base results in higher current gain (collector current divided by base current, denoted as β for bipolar transistors). For a homojunction device, an increase in gain can only be realized by either thinning the metallurgical base width or increasing the doping in the emitter. The higher current gain in an SiGe HBT can then be traded off for increased base doping or lower emitter doping to improve the base resistance and the emitter-base capacitance, resulting in greater RF performance.

The second technique for utilizing heterojunction effects in an HBT (see Figure 1.1b) employs a *graded* Ge profile to create a built-in (exists at zero bias) electric field in the base that accelerates electrons, reducing the base transit time and improving high-frequency performance. This second technique somewhat offsets the effects of the first technique [2] because using the graded profile necessarily means a reduction in the Ge content at the emitter-base junction, thus reducing the conduction band lowering effect discussed above. Thus, careful design of the Ge profile throughout the device is a key factor in achieving optimal device performance. Today's SiGe bipolar HBTs make use of these two techniques to varying degrees to create a performance advantage over conventional bipolar devices.

The use of an epitaxially grown base rather than one that is formed by ion implantation is another reason why SiGe HBTs exhibit better performance than conventional bipolar devices. The base of a conventional bipolar device is formed by implanting base dopant into Si, which results in a relatively broad base after subsequent thermal processing. Epitaxy allows one to "grow-in" the base doping profile through the deposition of doped and undoped Si and SiGe layers controlled to nearly atomic dimensions. This allows the device designer to create an arbitrary base profile. An implanted device is limited to skewed Gaussian dopant profiles whose widths are a function of implantation energy. Usually, the epitaxy technique is used to distribute the same base dose in a narrower base width, improving the transit time through the base and resulting in better high-frequency performance.

Despite the advantages introduced by the epitaxial growth of the base layer, the final dopant profile in the device is largely determined by the subsequent thermal processing of the wafers after the base growth. Due to the large diffusion coefficient of boron (typically used as the base dopant) in Si, a narrow as-grown base profile might be dramatically diffused by the time the processing is completed. Ge itself actually serves to arrest the diffusion of boron somewhat, but in modern SiGe HBTs another atomic species, carbon, is added in the epitaxial base of SiGe devices to further arrest the diffusion of boron [3]. A small amount of carbon is added (typically <1% atomic concentration of carbon is used) in the SiGe base during epitaxial deposition such that the electrical behavior is not significantly altered but the material properties are altered to reduce boron diffusion. Carbon helps to maintain a final boron profile closer to the as-deposited profile than it would be without carbon. The electrical effect is a faster transit time due to a narrower base width, improving high-frequency performance. It should be noted that the introduction of carbon does reduce some of the beneficial band offsets that are introduced by Ge.

A final note about how the design of the epitaxial base growth affects HBT performance concerns strain. All SiGe HBTs are grown pseudomorphically on a Si substrate, meaning that the SiGe (or in modern devices SiGe:C) is strained to take on

the lattice constant of bulk Si in the plane of the wafer. Any relaxation of the SiGe layers would generate dislocation-type defects that would short-circuit the emitter through the base to the collector of the device. Thus, all SiGe base layers must necessarily be grown pseudomorphically. Since bulk SiGe has a lattice constant that is larger than that of Si, the SiGe is always under compressive strain in the plane of the wafer and the lattice stretches out in the direction perpendicular to the surface of the wafer according to the material's Poisson ratio. This strain can actually serve to enhance both the mobility of electrons traveling vertically through the device and the mobility of holes traveling horizontally from the extrinsic base to the intrinsic base. In bulk $Si_{1-x}Ge_x$, the mobility of electrons is actually *lower* than that of bulk Si until the Ge concentration nears 100%. However, strained SiGe can have electron mobilities that are equal to or even superior to those of bulk Si [4], thus enhancing the transit of the minority carrier electrons through the base. The introduction of carbon mitigates some of this strain by pushing back the SiGe:C layer's lattice constant closer to bulk Si. So, again, trade-offs exist in the introduction of carbon in various locations of the epitaxial base growth.

In summary, modern SiGe HBT devices make use of the introduction of both Ge and C into the base of the transistor in order to manipulate both the electronic structure and the metallurgical structure of the device to achieve a performance that is not otherwise obtainable in bulk Si devices.

1.3 APPLICATIONS DRIVING SiGe DEVELOPMENT

Several applications have driven advances in SiGe technology since the first high-speed SiGe bipolar devices were demonstrated in the late 1980s [5]. Initially, SiGe devices were conceived as a replacement for the Si bipolar device for emitter-coupled logic (ECL); high-speed digital integrated circuits (ICs) where SiGe transistors promised higher f_T, improving gate delay relative to their Si bipolar or CMOS counterparts. However, advancements in the density, performance, and power consumption of CMOS technology quickly made it the logical choice for all but a few of these applications. Therefore, in the mid-1990s, SiGe technology appeared to have a limited application base in only specialized very-high-speed digital functions.

However, with the boom in wireless communications that began in the mid-1990s, a new application emerged as the primary driver for SiGe technology: the transceiver of a cellular phone. This application is tailor-made for SiGe BiCMOS. It requires a good high-frequency performance to support carrier frequencies in the 900 MHz–2.4 GHz range, a very-low-noise operation (as very small signals must be received and amplified), and a large dynamic range (as large output signals are required to drive the power amplifier and the antenna communicating with a faraway base station). In addition to wireless transceivers, high-speed fiber-optic transceivers also provided a good application for SiGe transistors as these pushed to even higher speeds, moving from 3 to 10 Gbps and eventually 40 Gbps data rates. The transition from 3 to 10 Gbps provided a strong market for SiGe devices as many of the same characteristics required for wireless transceivers are important in these transceivers (high speed, low noise, large dynamic range). But the transition from 10 Gbps to higher rates was delayed with the dot-com bust and only now are 40 Gbps and above

finally becoming mainstream technologies. Despite these higher data rate communication standards not becoming a reality, it was this expected transition in the late 1990s and early 2000s that pushed researchers to invest in creating very-high-speed SiGe transistors (with f_T and f_{MAX} of 200 GHz and above) that are now poised to take advantage of perhaps other emerging applications.

Today, deep-submicron CMOS is challenging SiGe for some of these traditional applications for two reasons: the speed of CMOS is adequate for many applications (although SiGe maintains an advantage in noise and an even wider advantage in dynamic range), and the density of CMOS is now high enough to enable new architectures that rely more heavily on digital signal processing rather than high-fidelity analog manipulation. In many cases, however, SiGe technology still offers a performance and power advantage and continues to play a strong role in both the wireless and wire line transceiver market. It should also be noted that the current cost of advanced SiGe BiCMOS wafers is significantly less than the RF performance-equivalent CMOS node since the SiGe BiCMOS devices do not rely on nearly as advanced lithography nodes as their CMOS counterparts. In addition to the incursion of CMOS devices in the traditional SiGe application space, current or even past-generation SiGe transistor performance is more than adequate to serve many of these applications. Therefore, these markets are becoming less important as drivers for future technology advancements.

Looking forward, however, two applications are primarily driving SiGe performance advancements today: higher-frequency millimeter-wave communications and higher-power but lower-frequency products. Millimeter-wave applications include, for example, proposed ~60 GHz WLAN standards [6], 77 GHz automotive collision avoidance systems, 94 GHz and above "terahertz" passive imaging, and 40–100 Gbps optical networking communications. These applications will serve to drive the speed of the SiGe transistor to higher and higher levels. At the other end of the performance versus breakdown spectrum, high-power applications include, for example, the power amplifier for wireless devices and laser/optoelectronic modulator drivers for wire line transceivers. These applications will drive improved trade-offs between speed and breakdown voltage in SiGe transistors. In the next section, we will review in more detail the design of SiGe transistors and see how improved speed and improved high-power performance are being realized.

1.4 SiGe PERFORMANCE METRICS

Two figures of merit are typically used to benchmark high-frequency device performance: (a) the cutoff frequency (f_T), which, for a bipolar device, is defined as the frequency at which the ac gain is unity, and (b) the maximum frequency of oscillation (f_{MAX}), which, for a bipolar device, is defined as the frequency at which the power gain is unity (usually the unilateral power gain).

For a bipolar device, f_T and f_{MAX} are related to basic device parameters by the commonly used equations:

$$f_T = \frac{1}{2\pi \cdot \tau_F}, \tag{1.1}$$

$$\tau_F = \left(C_{\mathrm{BC}} + C_{\mathrm{BE}}\right)\cdot\left(R_E + \frac{kT}{qI_C}\right) + \frac{W_B^2}{2D_B} + \frac{W_C}{2\nu_S} + R_C \cdot C_{\mathrm{BC}}, \tag{1.2}$$

$$f_{\mathrm{MAX}} = \sqrt{\frac{f_T}{8\pi \cdot R_B \cdot C_{\mathrm{BC}}}}, \tag{1.3}$$

where:

τ_F = forward transit time
C_{BE} = emitter-base capacitance
C_{BC} = base-collector capacitance
R_E = emitter series resistance
I_C = collector current
W_B = vertical base width
D_B = electron diffusion length in the base
ν_S = electron saturation velocity
W_C = vertical collector-base depletion width
R_C = collector resistance

At low collector current, f_T is dominated by the first term in Equation 1.2 where the junction capacitances combine with internal resistances to create an $R \times C$ time constant delay that is significantly longer than the other time constants in Equation 1.2 (see Figure 1.2). At high current, the term W_B becomes a function if I_C. When the charge associated with the current through the collector-base depletion region becomes comparable to the intrinsic doping level on either side, the edges of the depletion region collapse and thus "push" the depletion region away from the base, effectively widening the base. Mathematically, this occurs when

$$J_C \approx qN_C v_{\mathrm{SAT}}, \tag{1.4}$$

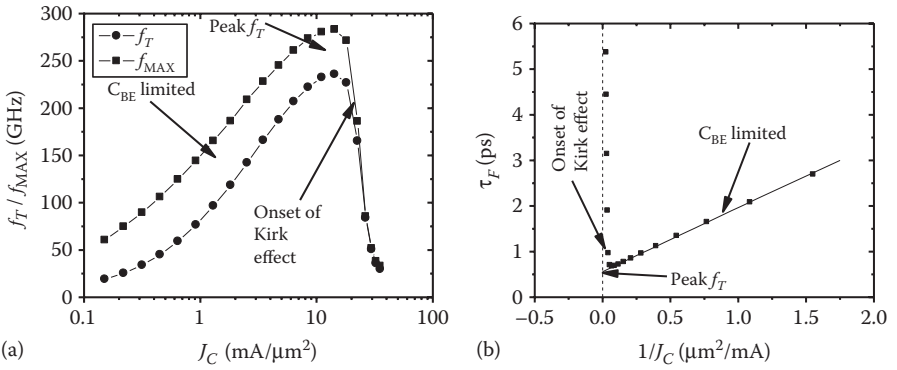


FIGURE 1.2 (a) Typical f_T versus J_C plot for a state-of-the-art, volume-manufactured SiGe HBT showing the various regions indicated by the terms in Equation 1.2. (b) Typical τ_F ($1/(2\pi f_T)$) versus $1/J_C$ plot indicating the various regions indicated by the terms in Equation 1.2.

where N_C is the nominal doping in the collector and v_{SAT} is the electron saturation velocity in the collector. This base push-out, known as the Kirk effect [7], is responsible for f_T decreasing at high current rather than saturating as would otherwise be predicted by Equations 1.2. So both f_T and f_{MAX} peak at a specific current density (see Figure 1.2).

Obtaining ever-higher peak f_T and f_{MAX} is important because, while today's volume RF applications target modest operating frequencies relative to the peak f_T's shown in Figure 1.2, high peak f_T (and f_{MAX}) can be traded off for other benefits including reduced power consumption, higher breakdown voltage, and reduced noise.

Figure 1.3 shows an example of the power savings that can be achieved with higher f_T SiGe technology even when operating at relatively low frequencies. For instance, at 25 GHz or so there is a threefold improvement in current consumption going from a 0.3 μm technology to a 0.2 μm technology. At 50 GHz, the advantage is fourfold. Thus, the scaling of the emitter is a key factor in reducing the power consumption required for a given f_T. Alternatively, when used as a gain stage in an amplifier, one could operate the device at peak f_T and simply use fewer gain stages to achieve the same total circuit gain. For instance, in theory, one could use any of the top three technologies shown in Figure 1.3 (0.13, 0.15, and 0.2 μm emitter-width SiGe HBT technologies) to achieve gain at 100 GHz. However, the 0.13 μm technology provides approximately 7 dB of gain at 100 GHz for peak f_T conditions whereas the 0.2 μm technology provides only 3.5 dB. Thus, the number of gain stages could be halved to achieve the same total gain, which provides an advantage in terms of power consumption, circuit area, and the total noise added by the circuit.

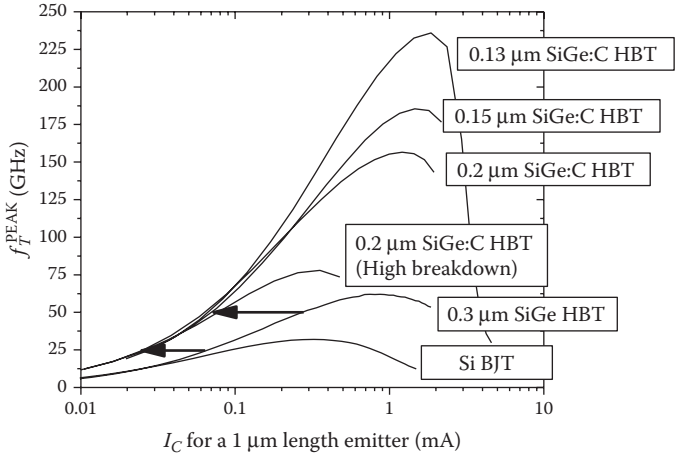


FIGURE 1.3 f_T for various TowerJazz BiCMOS technologies plotted as a function of I_C for a minimum width and unit length emitter. The dimensions in the labels refer to the minimum emitter width, not the corresponding CMOS technology node. In addition to higher peak f_T, subsequent technology nodes lower power consumption even when biasing at low f_T as indicated by the arrows.

The second advantage of higher f_T's, even in lower-frequency applications, is in the RF noise figure. The minimum noise figure can be expressed by [8]

$$NF_{\mathrm{MIN}} = 1 + \frac{n}{\beta} + \frac{f}{f_T} \cdot \sqrt{\frac{2qI_C}{kT} \cdot \left(R_E + R_B\right) \cdot \left(1 + \frac{{f_T}^2}{\beta \cdot f^2}\right) + \frac{n^2 {f_T}^2}{\beta \cdot f^2}}, \tag{1.5}$$

where n is the collector current quality factor and β is the current gain. From Equation 1.5, it is seen that with a high β, as is typically seen in SiGe devices, the noise figure reduces to

$$NF_{\mathrm{MIN}} \approx \frac{f}{f_T} \cdot \sqrt{\frac{2qI_C}{kT} \cdot \left(R_E + R_B\right)}. \tag{1.6}$$

In this limit, a higher f_T and a lower R_B result in a lower noise figure. Further, the ability to operate at a lower I_C and obtain the same f_T can lead to a reduction in the term under the radical in Equation 1.6. Thus, all the advantages enabled by using SiGe in the base of an HBT are brought to bear when one is attempting to minimize the noise figure: the high β enabled by putting SiGe at the emitter-base junction, the lower R_B for a given β enabled by increasing the base doping, and the higher f_T enabled by the Ge profile in the base.

Finally, f_T can be traded off for a higher breakdown voltage by modulating the collector doping concentration through a collector implant mask such that multiple devices spanning a range of f_T and breakdown are made available on the same wafer. Figure 1.4 shows the family of devices realized by this technique across several

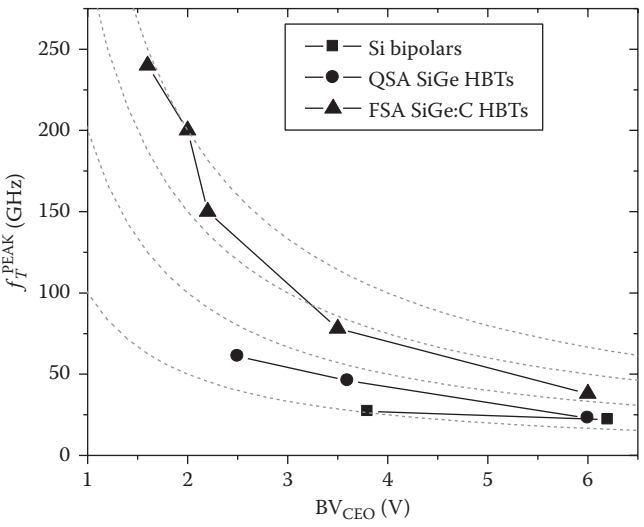


FIGURE 1.4 f_T versus $\mathrm{BV_{CEO}}$ plotted for all devices available in several generations of bipolar technologies. The dashed lines are contours of constant $f_T \times \mathrm{BV_{CEO}}$. All data are from TowerJazz electrical specifications.

generations of TowerJazz technology. Each subsequent generation not only supports devices with higher f_T but also improves the trade-off between f_T and the breakdown voltage, improving the large signal performance for applications such as integrated drivers and power amplifiers. This is in contrast with CMOS where each new generation makes the integration of power devices more difficult due to the more brittle gate oxide forcing lower voltage ratings.

1.5 DEVICE OPTIMIZATION AND ROADMAP

Vertical scaling of the HBT device enables higher f_T's. The most fundamental device enhancement with each generation of higher f_T devices is scaling of the base width; the W_B term in Equation 1.2. The fundamental limit of scaling the base width occurs when the emitter-base and the base-collector depletion regions touch and thus the device is "punched-through." It should be noted that the metallurgical base width in advanced SiGe HBTs is already many times narrower than all but the most aggressive CMOS channel lengths. The next most commonly adjusted vertical scaling parameter is the collector doping. An increase in collector doping offsets the Kirk effect as indicated by Equation 1.4, but, again, a fundamental limit is reached when the doping becomes so high that reverse bias leakage between the base and collector, either due to avalanche multiplication or tunneling, dominates the device behavior. Due to the reasons discussed in Section 1.2, the additions of Ge and C into the base of the HBT serve to delay the point at which these fundamental limits are reached and thus allow further vertical scaling of the device than would be possible for a homojunction device. It should be noted that scaling down of the base width or scaling up of the collector doping both serve to reduce f_{MAX} by increasing R_B in the first case and by increasing C_{BC} in the latter (see Equation 1.3). Thus, most techniques employed to enhance f_T trade off with a reduction in f_{MAX}, and other techniques are needed to simultaneously improve f_{MAX}.

Higher f_{MAX}'s are enabled by lateral scaling of the devices. Smaller device dimensions serve to reduce the R_B and C_{BC} terms in Equation 1.3. In fact, most of the research involved in developing a new generation of SiGe HBT devices involves creating new ways to reduce these two parasitic parameters. At the heart of the scaling of SiGe HBT devices is the emitter width, which in turn limits most of the other dimensions in the device as a whole. While the most advanced SiGe HBT devices constructed to date have emitter widths less than 100 nm [9], scaling of the emitter width is roughly 10 years behind the scaling of CMOS gate lengths. Figure 1.5 shows projection data from the International Technology Roadmaps for Semiconductors (ITRS) for CMOS and bipolar technologies [10], showing projected f_{MAX} versus the minimum feature width in the given technology node. It shows that, on average, one can achieve the same f_{MAX} in a bipolar device with a minimum feature width roughly three times larger than CMOS.

Several device architectures have been developed in the past decade to allow the scaling of SiGe HBT devices down to nanoscale dimensions. Figure 1.6 shows a generic example of the device architecture that is used to construct most modern SiGe HBTs. The first, large-scale manufactured SiGe HBT devices were built with a "quasi-self-aligned" architecture [11] where the extrinsic base is self-aligned by ion

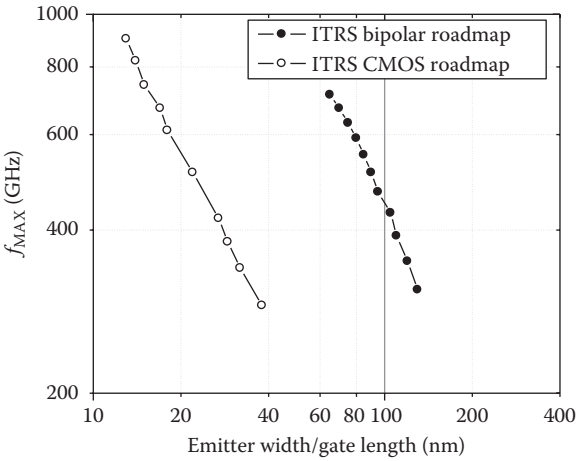


FIGURE 1.5 f_{MAX} versus minimum feature size for bipolar versus CMOS technologies. Data are from the ITRS roadmaps for CMOS and bipolar technologies. (From International Technology Roadmap for Semiconductors, International SEMATECH, Austin, TX, 2010.)

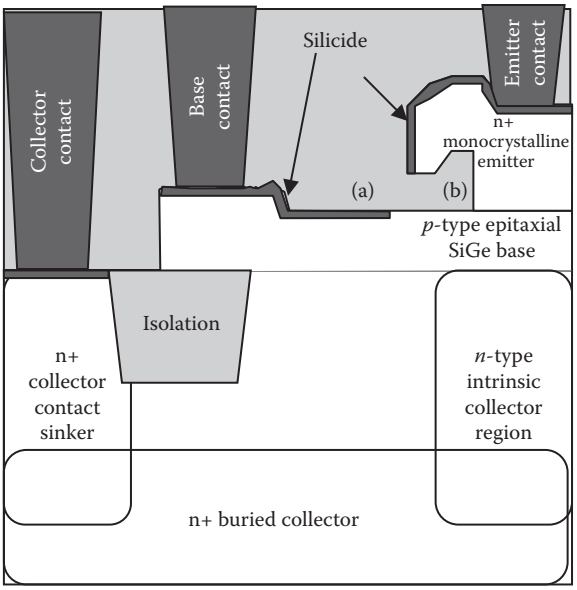


FIGURE 1.6 Example cross section of a modern SiGe HBT: (a) denotes the silicided extrinsic base region and the implanted extrinsic base region in quasi-self-aligned devices and (b) denotes the "link" or "spacer" region, which is doped (and thus implies "fully self-aligned") by the various techniques discussed in the text.

implantation to the edges of the emitter poly but not the emitter itself (see the area labeled [a] in Figure 1.6). The next generation of devices split off into several different architectures, which are "fully self-aligned," meaning that the extrinsic and intrinsic base alignment does not depend on mask alignment. One type of device uses a deposited polycrystalline extrinsic base followed by "selective epitaxy" of the intrinsic SiGe base [12]. A second method uses a sacrificial emitter post and spacer similar to the construction of a CMOS device [13]. Finally, various methods of growing a "raised extrinsic base" after the epitaxial growth of the intrinsic SiGe base have been developed [14,15]. All of these modern techniques essentially serve to dope the region denoted by (b) in Figure 1.6 at a higher *p*-type level than that of the SiGe epitaxy. This extra doping in the extrinsic base region serves to lower the total R_B of the device and thus improve f_{MAX}. While the techniques used to enhance f_T discussed at the beginning of Section 1.5 tend to reduce f_{MAX}, the scaling and architecture enhancements discussed here serve to improve f_{MAX} without any significant penalty to f_T. Thus, these innovations have allowed continuous scaling of SiGe HBT devices akin to what is done in CMOS.

Figure 1.7 shows a compilation of f_T and f_{MAX} data from over 100 SiGe HBT publications overlaid with the 2010 ITRS roadmap for bipolar devices [10]. The scatterplot shows the basic correlation of the progression of f_T and f_{MAX} despite the trade-offs mentioned above. The roadmap data predict that the same lithography advancements responsible for the CMOS roadmap will enable improved SiGe performance for the foreseeable future.

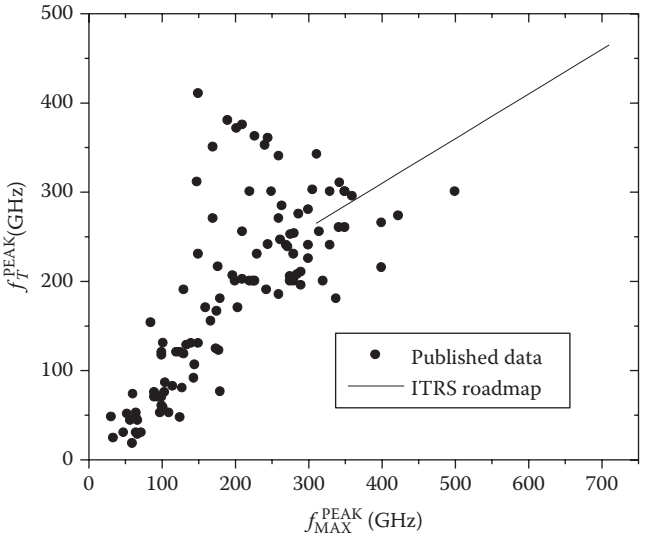


FIGURE 1.7 f_T versus f_{MAX} scatterplot for published SiGe HBT data. The line is projection data from the ITRS bipolar roadmap. (From H. Rucker, et al., *BCTM Proc*., 121–124, 2003.)

To realize useful RF and analog circuits, however, more than just high-speed SiGe devices are necessary. In the next section, we will discuss modules integrated with SiGe transistors that help create a more complete modern platform for RF and analog IC design.

1.6 MODERN SiGe BICMOS RF PLATFORM COMPONENTS

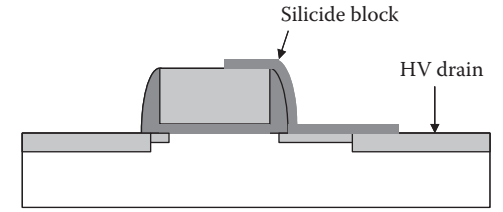
Technology features integrated with SiGe transistors that make them useful for product design include active elements such as high-density CMOS, high-voltage CMOS, and high-performance PNPs as well as passive elements such as high-density metal–insulator–metal (MIM) capacitors and high-quality inductors.

Today, most SiGe development is done in the context of a BiCMOS process in a CMOS node that typically trails the most advanced digital node by several generations. The critical hurdle to integrating advanced CMOS and SiGe devices is to marry their respective thermal budgets without degrading either device. The addition of carbon to SiGe layers, as discussed in Section 1.2, has been used as a partial solution to this problem as it helps reduce boron diffusion allowing for a higher thermal budget after SiGe deposition. This, along with careful optimization of the integration scheme, has resulted in demonstrations of SiGe integration down to the 90 nm node [16].

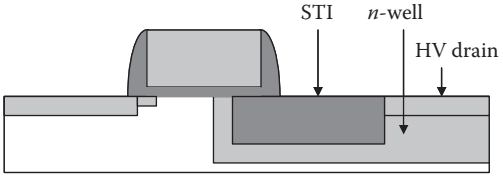
Power management circuitry can be enabled with higher-voltage CMOS devices (typically requiring tolerance of 5–8 V). In smaller geometries that support only lower core voltage levels, these are enabled by introducing drain extensions to the CMOS devices that can enable higher drain bias than is supported in the native transistor. An example of such devices is shown in Figure 1.8; these are becoming common modules in SiGe technology offerings, often not costing additional masking layers to create.

A high-speed vertical PNP (VPNP) can form a complementary pair with an SiGe NPN and is important for certain high-speed analog applications such as fast data converters, push-pull amplifiers, and output stages for hard disk drive pre-amps. Using a separate SiGe deposition step, a VPNP can be made very fast and f_T's as high as 100 GHz have been reported [17]. But the cost associated with such a VPNP is prohibitive for most applications today. A more popular approach reuses many of the steps needed to create the SiGe HBT and CMOS devices while adding specialized implants to optimize the performance of the VPNP. In this scheme, devices with f_T's of up to 30 GHz can be achieved as shown in Figure 1.9.

In addition to active components, high-quality passive components are necessary to enable advanced RF circuits. The most critical passive elements for RF design are capacitors and inductors as these can consume significant die area and, at times, limit the performance of RF and analog circuits. MIM capacitors are available in most commercial SiGe BiCMOS and RF CMOS processes as they achieve excellent linearity and matching. The density of MIM capacitors has been steadily increasing over time, helping to shrink RF and analog die. Figure 1.10 shows a timeline of capacitance density for TowerJazz integrated MIM capacitors. An initial improvement in density from <1 to 2 fF/μm^2 was enabled by a move from oxide to nitride dielectrics [18]. Then, the stacking of a 2 fF/μm^2 capacitor on two consecutive metal



(a)



(b)

	NFET	PFET	
Vds	<10	< 8	V
Vgs	< 3.6	< 3.6	V
Idsat	420	108	µA/µm
Ron	4.5	14	ohm-mm
BVdss	> 11	> 11	V

(c)

FIGURE 1.8 Sketch of two types of commonly used extended drain devices: (a) silicide block extension and (b) shallow trench isolation (STI) extension as well as (c) a table showing the characteristics of high-voltage (HV) devices available in a 0.18 µm SiGe BiCMOS technology using approach (b). Idsat is quoted for 3.3 V Vgs and 5 V Vds.

layers enabled a move from 2 to 4 fF/µm^2. Finally, a further optimization of the nitride dielectric resulted in a density of 5.6 fF/µm^2. Today, high-K dielectrics and various types of MIM trench capacitors are being investigated to enable even higher densities and it is conceivable that in the next few years, densities of 10–20 fF/µm^2 will be introduced.

Integrated inductor performance, measured as the quality factor (Q), is improved by the reduction of metal resistance, which is made possible by thicker metal layers. Inductor Q can be traded off for a reduced footprint such that a thicker metal layer can also help reduce the chip area. This concept is demonstrated in Figure 1.11, where the area required to realize an inductor with Q of 10 is compared between the use of a 6 µm and a 3 µm top metal in a four-layer Al metal process. A 6 µm metal inductor consumes half the die area of a 3 µm metal inductor while achieving the same Q in this example.

Die scaling enabled by the advanced passive elements described in this section can often more than pay for the additional processing cost. An optimized process can, in many cases, not only provide better performance than a digital CMOS process but also lower the die cost. Similarly, the integration of advanced active modules described in this section can help integrate more analog functionality on fewer die, reducing the overall system-level costs.

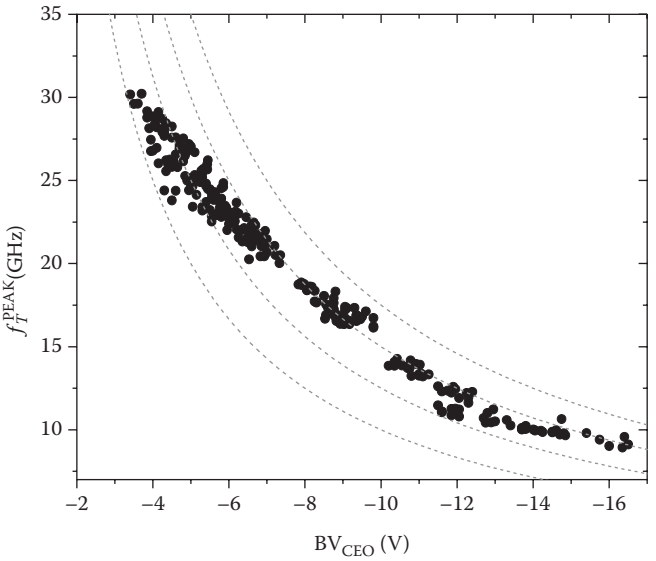


FIGURE 1.9 Performance (f_T) versus breakdown (BV_{CEO}) trade-off of vertical Si PNP devices integrated with SiGe NPNs to form a complementary pair. The dashed lines are contours of constant $f_T \times BV_{CEO}$. Data are from internal TowerJazz development wafers.

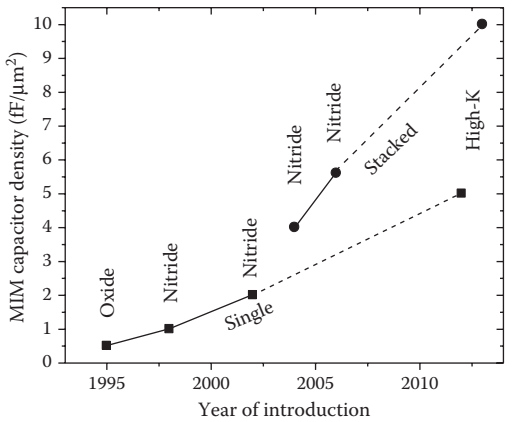


FIGURE 1.10 MIM capacitor density plotted as a function of year of first production (actual or planned) showing progression in dielectric technology (from oxide to nitride to high-K) and in integration (single to stacked capacitors).

1.7 CONCLUSIONS

In this chapter, we have reviewed SiGe BiCMOS technology and discussed how it has become important for many RF applications by providing a performance advantage over stand-alone CMOS while sharing its manufacturing infrastructure to provide integration and cost advantages over III–V technology. In addition to a

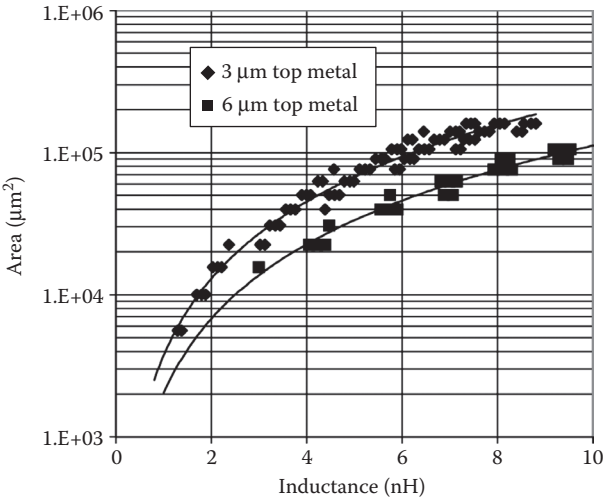


FIGURE 1.11 Inductor area as a function of inductance for a four-turn inductor with a peak Q of 10 built in 3 and 6 μm Al metal layers, respectively.

higher speed, we have seen that an intrinsic advantage of SiGe over CMOS is its ability to maintain higher breakdown voltages and therefore support applications that require a higher dynamic range. This gap will widen with more advanced generations of both CMOS and SiGe as each new generation of CMOS results in lower breakdown voltages while each new generation of SiGe results in a better trade-off between speed and breakdown. In addition, we have seen that the performance of SiGe devices can be improved with advanced lithography much in the same way as with CMOS devices such that a raw performance gap will continue to exist between SiGe and CMOS as more advanced nanometer nodes are created in the future. This will continue to enable a market for SiGe at the bleeding edge of performance, which today is translating into interest for SiGe in several millimeter-wave applications and very-high-speed networks.

The biggest threat to SiGe advancements is the failure to identify high-speed, high-volume applications that take advantage of these benefits in the future; however, much like Moore's law for CMOS, which has held true for decades and which applications have taken full advantage of, the imagination of the industry has never let us down and is not likely to do so in this case.

ACKNOWLEDGMENTS

The authors would like to acknowledge the help of current and former coworkers at TowerJazz including Volker Blaschke, Dieter Dornisch, David Howard, Chun Hu, Paul Hurwitz, Amol Kalburge, Arjun Karroy, Paul Kempf, Lynn Lao, Zachary Lee, Pingxi Ma, Greg U'Ren, Jie Zheng, and Bob Zwingman, as well as contributions from many of TowerJazz's customers and partners. A final special acknowledgment