Silicon-Germanium Heterojunction Bipolar Transistors For Extremely Low-Noise Applications

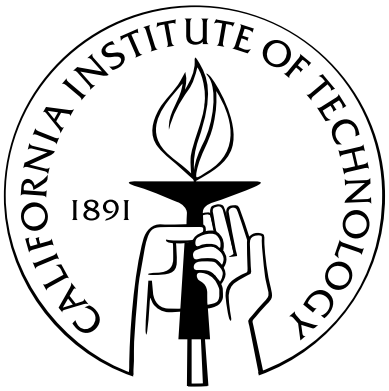
Thesis by

Joseph Cheney Bardin

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2009

(Defended May 21, 2009)

Acknowledgments

I am deeply indebted to Dr. Sander Weinreb, with whom I have worked for the past seven years, beginning as an undergraduate and continuing through my masters and doctoral work. I am extremely fortunate to have had the pleasure to work with such an insightful advisor, who not only served as an excellent mentor, but has never hesitated to share his wisdom regarding both technical and non-technical matters.

For serving as my academic advisor and offering invaluable career advice as well as feedback on my research, I would like to thank Prof. D. Rutledge. I would also like to express my gratitude to the other members of my thesis committee, Prof. A. Emami, Dr. T. Gaier, Prof. A. Hajimiri, and Prof. J. Zmuidzinas, who all took the time out of their busy schedules to review my work and offer helpful feedback.

Throughout my stay at Caltech, I have learned a great deal through my interactions with fellow graduate students, colleagues, and collaborators. For this I am grateful to G. Jones, H. Mani, M. Arii, B. Farr, J. Stern, P. Pütz, J. Yuan, T. Thrivikraman, C. Walker, C. Groppi, C. Kulesa, J.-L. Medina-Monroy, S. Shirinfar, S. Jeon, F. Wang, A. Babakhani, F. Böhn, H. Wang, Y. Wang, E. Keehr, J. Buckwalter, S. Bowers, K. Dasgupta, J. Yoo, and K. Sengupta. For assistance with travel, purchasing, computer upkeep and other day to day events, I am grateful to S. Slattery, D. Yee, M. Chen, L. Dozsa, L. Hein, J. Lilley, C. Sosnowski, and G. Waters.

The cryogenic measurements reported in this dissertation would not have been possible without access to test equipment at the Jet Propulsion Laboratory; for providing this access as well as technical support, I owe a debt of gratitude to E. Long and J. Bautista of JPL. Furthermore, without the access to IBM wafer runs free of cost through the trusted foundry program, it is unlikely that the

majority of the integrated circuits reported in this work and elsewhere would have been fabricated; for arranging this access, I am grateful to Dr. K. Johnson of the USNO.

A large portion of the work reported in this dissertation would not have been possible without the generous donation of sample transistors: I would like to express my gratitude to Dr. P. Chevalier of ST Microelectronics for providing many sample devices, detailed documentation, as well as for his generosity in providing thorough answers to my questions regarding the sample transistors. For additional sample devices, I would like to thank Dr. M. Racanelli of JAZZ semiconductor, B. Smolders of NXP, and R. Scholz and W. Kissinger of IHP.

For financial support, I would like to thank the California Institute of Technology and the IEEE Microwave Theory and Techniques Society.

Finally, I would like to express my deepest gratitude to my parents, brother, and S. Nichols, all of whom have supported me throughout this experience.

This work was partially supported by a President's Fund grant from the Jet Propulsion Laboratory.

Abstract

Historically speaking, the world of extremely low-noise solid-state amplification has been dominated by exotic technologies such as InP and GaAs HEMTs. By cryogenically cooling these devices, it is possible to realize microwave amplifiers with noise temperatures as low as 5K over decades of bandwidth. Although HEMTs can provide very low-noise amplification when cooled to cryogenic temperatures, their radiometer performance is limited by intrinsic transconductance fluctuations. It is believed that bipolar devices do not suffer from this problem. As industry has invested more and more money into silicon based technologies, silicon-germanium (SiGe) heterojunction bipolar transistors (HBTs) have continued to improve and are now at the point where they are beginning to become competitive with InP HEMTs for microwave cryogenic low-noise amplifiers. Although extremely high frequency device operation has been observed at cryogenic temperatures, little work has been done on modeling the noise of cooled SiGe HBTs.

In this report, a thorough investigation into the theoretical and practical aspects of using silicon-germanium (SiGe) heterojunction bipolar transistors (HBTs) for extremely low-noise applications is presented. The dissertation is broken up into three sections:

1) Background information: The fundamentals of SiGe HBTs are presented along with a discussion of how the properties of semiconductors change at cryogenic temperatures, as well the impact that these changes have on the performance of the devices.

2) Modeling: A comprehensive study of seven state-of-the-art HBTs at temperatures ranging from 18 K to 300 K is presented. The devices are compared in terms of dc, small-signal, and noise performance, and small-signal noise models are extracted. The section concludes with a brief summary of the important conclusions regarding the performance of SiGe devices at

cryogenic temperatures.

3) Applications: The models developed previously are applied to the design of several state-of-the-art LNAs in both MMIC and discrete form. Noise performance better than 2 K is achieved in the low-GHz range, which is comparable to the best InP results. The section concludes with a discussion of some high-impedance differential amplifiers which have recently been fabricated.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background Material

This dissertation is about the broad-band noise, dc, and RF performance of state-of-the-art silicon-germanium (SiGe) heterojunction bipolar transistors (HBTs) operating at cryogenic temperatures and their use in extremely low-noise applications. It is divided into three parts:

1) **Theory** (Chapters 2–3). The framework necessary to understand the rest of the thesis is presented. This includes an overview of silicon germanium transistors and a review of the properties of semiconductors at cryogenic temperatures. Finally, this part of the thesis concludes with a theoretical investigation as to what one would expect to happen to the operating characteristics of SiGe devices at cryogenic temperatures.

2) **Modeling** (Chapters 4–8). The temperature dependent properties of a wide variety of state-of-the-art SiGe transistors have been investigated and are reported in terms of dc, microwave, and noise-performance. The results are presented in a manner so as to highlight the differences between the various devices, while also pointing out general trends that are believed to apply to all modern SiGe HBTs. In Chapter 4, the devices under study are described and an explanation of the cryogenic test-setup used for device characterization is given. The experimental results of this comparative study are presented in Chapters 5–7. Finally, this section of the thesis is drawn to a close in Chapter 8, in which a concise summary of how SiGe devices change as a function of ambient temperature is provided.

3) **Applications** (Chapter 9). Several applications of SiGe HBTs at cryogenic temperatures are presented. In Chapter 9, the design and measurement of a variety of low-noise amplifiers is reported. Through these "proof-of-concept" amplifiers, it is shown that the performance predicted in Part II of the dissertation can be realized, thereby verifying the theory on which

the modeling was based. Furthermore, the excellent results which are achieved show that SiGe devices are well suited for low-noise cyrogenically cooled applications. The chapter concludes with a discussion of some differential amplifiers which have been developed for integration with high impedance feeds.

The remainder of this chapter is devoted to the motivation for this work, a brief history of the field, and fundamental background material.

1.1 Solid-State Technology for Extremely Low-Noise Amplifiers

The sensitivity of a receiving system is limited by its system noise temperature, which is a combination of the input-referred receiver noise and the background noise. For terrestrial communication systems, where the background noise is on the order of 300 K, reducing the noise of the LNA below 77 K, or a 1dB noise figure, provides diminishing returns as the dominant source of noise is actually the background noise. However, for some applications, such as radio astronomy, deep-space communications, and low-temperature physics research, the background noise is just a few Kelvin[1]. For example, in the case of radio astronomy, the background noise is comprised of 2.7 K that is attributed to the big bang and an additional few Kelvin due to atmospheric attenuation. For applications such as this, the noise requirement is quite stringent and it has become common practice to cryogenically cool the receiver front-end in order to reduce the system noise to within a factor of two of the background noise.

Since the early 1980s, the vast majority of solid-state amplifiers used in extremely low-noise applications have relied upon III-V technologies, with the pioneering work in the field spear-headed by Weinreb in his ground-breaking paper entitled "Low-noise Cooled GASFET Amplifiers" [1]. To understand why III-V devices have dominated the field of extremely low-noise amplifiers, it is useful to consider the RF performance of different solid-state technologies over the years. A plot comparing the maximum frequency of oscillation (f_{max}) as a function of year for state-of-the-art devices across various technologies is shown in Fig. 1.1. The data clearly shows that prior to 1994, the III-V devices were an order of magnitude better than their silicon counterparts in terms of high-frequency operation[2]. While these data are not directly applicable to the cryogenic noise performance, the high-frequency operating characteristics would certainly limit the room temperature noise performance

[1] Several applications are detailed in Table 1.1.

[2] The superior RF performance is largely due to the higher mobility afforded by the use of exotic III-V compounds

Table 1.1: Summary of key applications requiring very low-noise amplifiers

Application	Freq. Range	Comments
Radio Astronomy	0.1-100GHz	Microwave Telescopes. 1-100,000,000 elements. Cryogenic cooling not always economic.
	0-5GHz	SIS mixer IF amplifiers
	0-3GHz	HEB mixer IF amplifiers. Requires very good input match to achieve good HEB stability. Bandwidth limited by thermal time constants.
Deep Space Comm.	8GHz,32GHz	Current DSN link. Lower ground terminal system noise can be leveraged into either higher data rates or lower power transmitters/smaller transmit antennas
	0-10GHz	For use with optical communications system. Large signals. Jitter is critical.
RFSQ Logic		Superconducting computers. Interfacing circuitry to transition from low voltage Josephson-junction logic to CMOS memory.
Low Temp. Phys.	Varies	Sensor readout. Required sensitivity can be on the order of the quantum limit.

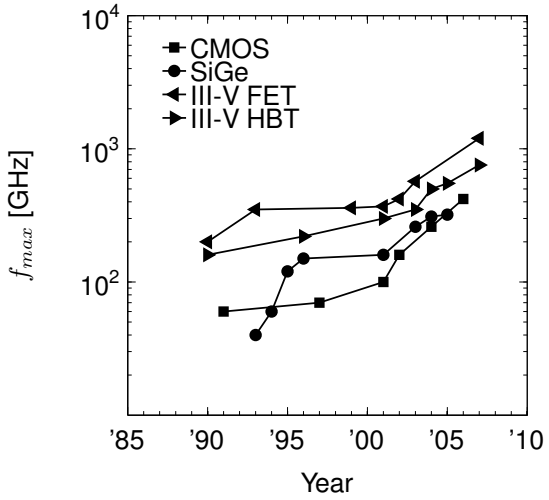


Figure 1.1: Historical f_{max} for different technologies. Data taken from [2, 3, 4, 5].

of the silicon devices and are indicative of inferior cryogenic performance as compared to the III-V devices.

Much of the research in cryogenic LNA development has been geared towards radio astronomy applications, where the systems have typically been centered around large single-pixel[3] antennas. Thus, capital investment in steel has often been the major force in determining the overall system cost, and the extra investment required to use the absolute best available devices has easily paid for

[3]i.e., one receiver per dish

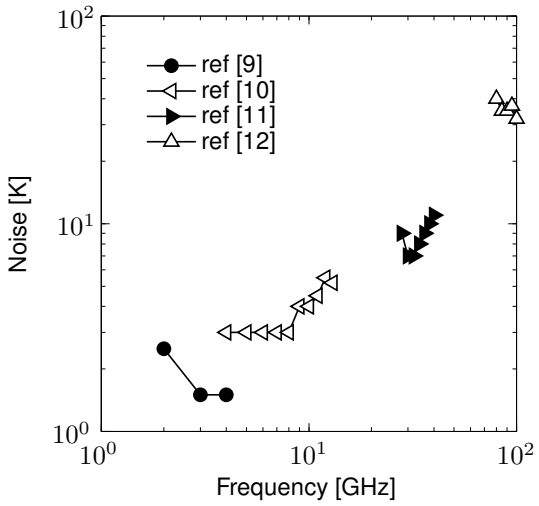


Figure 1.2: Examples of state-of-the-art cryogenic noise performance achieved using III-V HEMTs at 15 K physical temperature

itself in terms of an improvement in system sensitivity[4]. Since achieving the state-of-the-art in noise performance meant a reduction in steel and ultimately cost, III-V HEMTs have dominated the world of extremely low-noise amplification since the early 1980s with the first-generation devices using AlGaAs/GaAs materials, the second generation devices using AlGaAs/InGaAs/GaAs materials, and the third generation devices using AlInAs/InGaAs/InP materials [6]. Using these technologies, truly phenomenal noise performance has been obtained at cryogenic temperatures. For reference some of the better published cryogenic noise results for III-V HEMT base amplifiers have been plotted in Fig. 1.2.

Although their broadband noise performance is excellent, a limiting problem with HEMT devices is inherent gain fluctuations that are linked to trapping phenomena associated with surface states [7]. These gain instabilities ultimately limit the radiometer performance of III-V HEMT LNAs. As it is assumed that this problem does not exist in SiGe devices, which are vertical devices and buried in the substrate, SiGe heterojunction bipolar transistors (HBTs) would be preferential if their noise performance were on par with III-V HEMTs [8]. In addition to the lack of gain fluctuation issues, SiGe HBTs are usually fabricated as an extra step in a standard CMOS processes, meaning that they have the added advantages of eight-inch wafers, high yield, and full CMOS digital capabilities. Thus, if the cryogenic noise performance of SiGe devices were comparable to that of III-V HEMTs, it would open a whole new application space.

[4]Sensitivity is defined as A_e/T_{SYS}, where A_e is the effective aperture and T_{SYS} is the system temperature including the background noise. Thus, sensitivity can be improved by increasing the effective collecting area or decreasing the system noise temperature.

Referring back to Fig. 1.1, it is apparent that the advantage in high-frequency performance that the III-V devices held over silicon devices in 1995 has grown considerably smaller in the past ten years. For reasons explained in Chapters 2 and 3, the noise performance of SiGe devices is expected to improve substantially as these devices are cooled. Therefore, a proper understanding of their cryogenic performance is extremely important in terms of enabling future breakthroughs in the implementation of very sensitive cryogenic systems. Prior to moving on to the study of SiGe devices for this application, a brief introduction to the dominant sources of noise in bipolar devices, as well as an introduction to noise measurement techniques at cryogenic temperatures, is presented.

1.1.1 Physical Sources of Noise in BJTs

The broadband noise performance that can be achieved using a given transistor is determined by the noise sources internal to the device as well as its RF terminal characteristics. Thus, a clear understanding of the physical sources of noise in bipolar devices as well as the assumptions made regarding their associated noise spectra is very important in developing an understanding of the noise properties of SiGe HBTs. The two dominant broadband noise sources in microwave BJTs are Johnson and shot noise. Johnson noise is the by-product of the random motion of carriers in a conductor due to thermal excitation [13, 14]. It can be shown that the available power in a bandwidth Δf due to this random thermal agitation is given as [15]

$$P_{av} = \frac{hf}{e^{hf/kT_a} - 1}\Delta f = \frac{kT_a}{1 + \frac{1}{2!}\frac{hf}{kT_a} + \frac{1}{3!}\left(\frac{hf}{kT_a}\right)^2 + \frac{1}{4!}\left(\frac{hf}{kT_a}\right)^3 + \ldots}\Delta f, \tag{1.1}$$

where h is Plank's constant, k is Boltzmann's constant, f is center frequency, Δf is the bandwidth, and T_a is the ambient temperature. Thus, the power spectral density is white at low frequencies and rolls off at higher frequencies with a knee frequency of $f_{3\text{dB}} \approx T_a \cdot 26.2$ GHz. Furthermore, the frequency at which the available noise power has dropped by 10% is given as $f_{0.5\text{dB}} \approx T_a \cdot 4.3$ GHz. Thus, at temperatures above 10 K and for frequencies below 40 GHz, it is reasonable to treat Johnson noise sources as ideal white noise generators[5] with available power of kT_a. Therefore, physical resistances can be modeled in Thevenin or Norton form as shown in Fig. 1.3.

The second important noise mechanism in bipolar devices is shot noise and occurs whenever charge carriers cross a potential barrier[6] due to the discrete nature of current flow. At frequencies

[5] In this work this assumption will be made, however it is important to realize that the assumptions break down under extreme operating conditions (e.g., operation in the hundred GHz range at very low temperatures).

[6] e.g., diffusion current in a pn junction

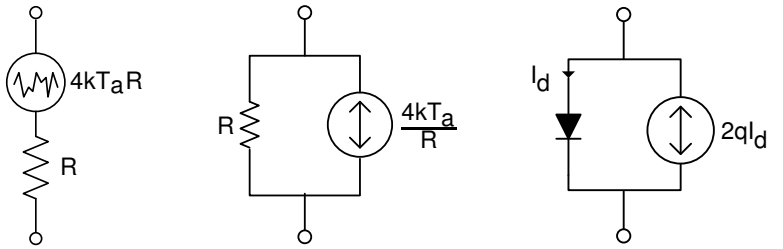


Figure 1.3: Circuit representations of Johnson noise in (a) the Thevenin representation and (b) the Norton representation. (c) Circuit representation of the shot noise associated with a forward biased diode

less than $1/\pi\tau_b$, where τ_b is the time over which the carriers transit across the barrier, the spectral density of the shot-noise current is given as $\overline{|i_n|^2} = 2qI_d$, where I_d is the dc current flowing across the barrier. At higher frequencies, the spectrum rolls off as $1/f^2$ [16]. In this work, it is assumed that transit time across the regions in which shot noise is generated[7] only makes up a small component of the total transit time of the device, and thus the spectral densities of the shot noise can be considered to be white.

Thus far, the dominant physical sources of noise in microwave bipolar-junction transistors have been presented. It is important to note that these noise sources are associated with specific physical processes which are localized to well defined regions internal to the device. Therefore, the actual physical noise sources are uncorrelated with one another. Furthermore, under the constraints discussed above, the power spectral density associated with each of the intrinsic broadband noise mechanisms can be considered frequency independent. In later sections, we will see that the noise parameters of the device have frequency dependence due to the fact that the intrinsic noise sources are imbedded in a frequency dependent network. Nonetheless, the fact that the physical intrinsic noise sources are uncorrelated and frequency independent under the assumptions stated above is important to keep in mind as the specifics of the noise performance of SiGe HBTs are discussed in the coming chapters.

1.2 Characterization of Noise at Cryogenic Temperatures

In order to specify the microwave small-signal performance of any two-port at a single frequency, twelve numbers are required; eight numbers are required to define the terminal current-voltage characteristics and four noise parameters are needed to determine the noise performance. Standard

[7] i.e., depletion regions

equivalent network representations used to describe the terminal characteristics include admittance- (Y-), impedance- (Z-), scattering- (S-), and chain-parameter (ABCD-parameter) formulations. The conversion between sets of network parameters is trivial and each particular representation proves advantageous in the analysis of a certain class of problems. However, in the measurement of network parameters, it has become standard practice to measure S-parameters using a vector network analyzer (VNA). Using a VNA, it is possible to accurately determine the S-parameters of a network from low-frequencies all the way up to the sub-millimeter wave regime. Furthermore, accurate calibration and de-embedding routines have been developed, allowing one to measure devices directly on-wafer as well as in coaxial and waveguide fixtures. Thus, S-parameter measurement techniques are easily applied in on-wafer cryogenic test setups.

As a major focus of this work is the noise performance of transistors at cryogenic temperatures, it may seem logical that a good portion of the work would be in measuring on wafer noise parameters at cryogenic temperatures. However, *in this work, the measurement of noise parameters has been explicitly avoided* for reasons explained in the following review of common noise parameter measurements techniques.

1.2.1 Noise Parameter Measurements

Unlike the measurement of S-parameters, which, in theory, is a rather straightforward procedure, the measurement of noise parameters is quite involved. Along with each of the network representations, there is an accompanying equivalent noise representation. For instance, to account for noise in the admittance representation, equivalent shunt noise-current sources across the input and output of the two-port are required along with knowledge of the complex correlation coefficient between the noise sources. These current-noise sources are not physical sources, but mathematical constructs defined in order to push the internal noise sources to the terminals of the two-port without changing its terminal noise characteristics. Thus, as the equivalent sources are not physical, they cannot be measured directly. The same is true for each of the noise representations[8], and an indirect measurement method has become the standard technique for determining the noise parameters. This technique relies on the fact that the input-referred noise temperature depends on the source

[8] An exception to this statement is the case in which the noise waves, which actually emanate from the ports, are measured [17]. However, despite its conception nearly 20 years ago, this technique is still not used in commercial systems.

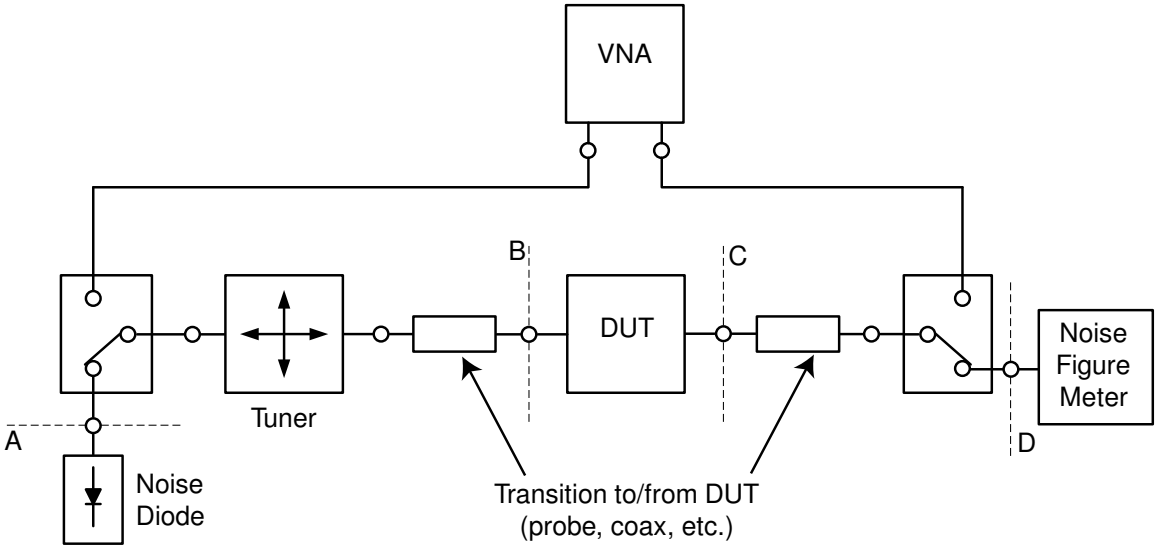


Figure 1.4: Typical noise-parameter test set

reflection coefficient (Γ_s) as [18]

$$T_e = T_{min} + 4T_0 \frac{R_n}{Z_0} \frac{|\Gamma_s - \Gamma_{opt}|^2}{|1 + \Gamma_{opt}|^2 \left(1 - |\Gamma_s|^2\right)}, \tag{1.2}$$

where T_{min} is the minimum noise temperature of the two-port taken over all source impedances, Γ_{opt} is the complex generator reflection coefficient required to achieve T_{min}, and R_n is known as the noise resistance and determines the sensitivity of T_e to Γ_s.

As T_{min}, R_n, and Γ_{opt} constitute a set of noise parameters that can easily be transformed to any of the other sets of noise parameters, the determination of these four numbers is sufficient to completely specify the noise properties of the network. Thus, the standard method for measuring noise parameters at ambient temperatures is to measure the noise temperature of the device under test (DUT) while varying the generator impedance using a tuner. A typical setup that would be used to perform this task appears in Fig. 1.4. By measuring the generator impedance presented to the DUT and knowing the exchangeable noise power at the input of the DUT[9] for each tuner position, it is possible to determine several values of T_e as a function of Γ_s. Once at least four of the values are known, the noise parameters can be determined using methods provided by Lane, Mitama, or others [19, 20].

In order for the tuner noise parameter measurement to work, it is usually assumed that all passive losses are isothermal, meaning that the noise parameters of the lossy component can be obtained

[9] i.e., knowing the loss in the mechanical tuner very accurately in order to account for its noise contribution

from Bosma's theorem as $\overline{\mathbf{C}} = kT_a\left(1 - \mathbf{S}\mathbf{S}^{\dagger}\right)$, where $\overline{\mathbf{C}}$ is the noise parameter matrix in noise wave format and the † symbol is the hermitian transpose operator [21, 22]. Thus, the noise parameters of the input network, including the tuner, as well as the output network can be determined directly through S-parameter measurements and the noise presented to the DUT at reference plane B can determined accurately from knowledge of the noise at reference plane A. Similarly, the noise at reference plane C can be determined from a noise measurement made at reference plane D.

While this procedure works well at 300 K, the assumption that all passive losses are isothermal is clearly violated when the measurements are made at cryogenic temperatures due to a temperature gradient along the coaxial cables connecting the cooled DUT to the outside world. To complicate matters, in order to reduce the heat load on the cooler, these coaxial cable are usually of the stainless-steel variety and tend to be lossier than standard copper coaxial cables. Furthermore, the center conductor of each cable, where most of the loss occurs, is heatsunk though the dielectric of the coaxial cable and the DUT, making its physical temperature profile impossible to measure. While there has been some work done on the computer aided modeling of the temperature distribution along the input line [23], it is believed that the level of uncertainty in this type of measurement is still prohibitively high to allow for the accurate measurement of state-of-the-art devices with minimum noise temperatures in the low K range.

1.2.2 Noise Figure Measurements

To avoid the measurement of noise parameters, Dambrine has proposed the 50 Ω noise measurement method, which is applicable to the class of devices in which a general small-signal equivalent circuit noise model exists and contains only one parameter that cannot be measured directly [24]. In this technique, a small-signal model is extracted and then a single 50 Ω noise measurement is made using a noise figure meter. Using the frequency dependence of the noise figure data in conjunction with the small-signal model, the unknown parameter can then be determined. For example, in the FET noise model proposed by Pospieszalski, all component values can be extracted through S-parameter and dc measurements with the exception of T_d, which is the effective temperature of the drain-source conductance and plays a major role in the noise properties [25]. Thus, the noise model can be determined after extracting the small-signal model by minimizing the difference between simulated and measured 50 Ω noise as a function of the unknown parameter, T_d.

While the 50 Ω noise measurement method is more suitable for determining noise parameters at cryogenic temperatures, there are still technical difficulties which arise due to the unknown noise

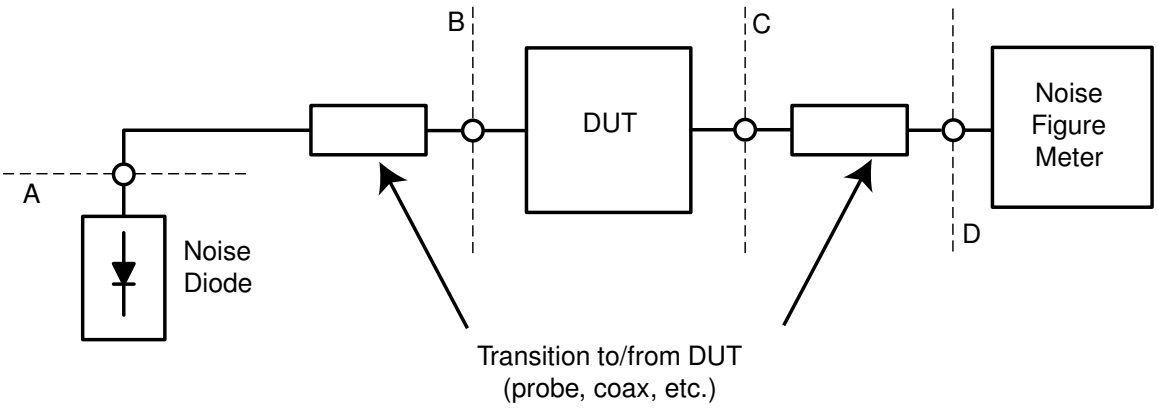


Figure 1.5: Typical 50 Ω noise figure test setup

contribution on the lines connecting to the DUT as well as the uncertainty in the calibration of the noise diode, both of which make the noise power at the reference plane of the DUT difficult to ascertain. To understand this issue, it is helpful to review the details of a noise temperature measurement. The noise temperature is usually measured using the Y-factor method in which "hot" and "cold" source termination are presented sequentially to the DUT and the total output power is measured in each state. If we denote the total measured output power under hot and cold excitation as $P_{hot,m}$ and $P_{cold,m}$, then the measured *Y-factor* is given by

$$Y = \frac{P_{hot,m}}{P_{cold,m}} = \frac{T_{hot} + T_e}{T_{cold} + T_e}, \tag{1.3}$$

where T_{hot} and T_{cold} are the effective noise temperature of the noise reference in the on and off state as referenced to the plane of the DUT. Thus, the effective noise temperature of the DUT is calculated as

$$T_e = \frac{T_{hot} - YT_{cold}}{Y - 1}. \tag{1.4}$$

In the derivation of equations (1.3) and (1.4), it is implicitly assumed that the exact value of the noise source ENR is known. However, typical state-of-the-art noise diodes have specified ENR uncertainties of greater than 0.1 dB [26]. In the case of an ENR uncertainty, it can be shown that the resulting uncertainty in the noise temperature measurement result is equal to

$$\Delta T_e = \pm \frac{(T_{hot} - T_{cold})\,(T_e + T_{cold})}{T_{hot} + \Delta T_{hot} - T_{cold}} \left(10^{\Delta ENR/10} - 1\right) \approx \pm\, T_{cold} \left(10^{\Delta ENR/10} - 1\right), \tag{1.5}$$

where ΔENR is the uncertainty of the noise source in dB. Evaluating equation (1.5) for the opti-

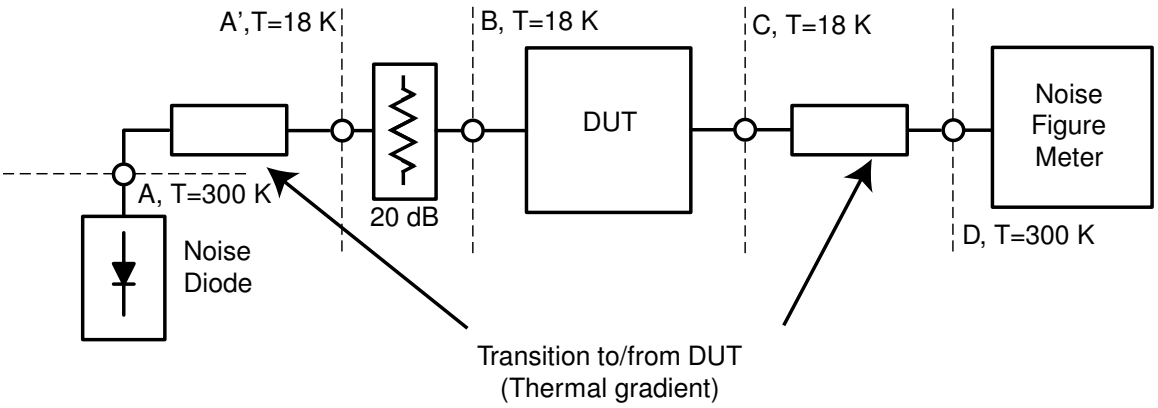


Figure 1.6: 50 Ω noise figure test setup employing a cooled attenuator

mistic case in which $\Delta ENR = \pm 0.1$ dB, we find that the uncertainty arising from the noise diode in the measured value of T_e is ± 6.8 K. Clearly this is unacceptable, as the devices we would like to measure have several times less noise. Fortunately, a method has been developed to accurately measure the noise temperature of extremely low-noise devices in the presence of noise diode ENR uncertainties. This method is the topic of the following section.

1.2.3 Cryogenic Noise Temperature Measurement: The Cooled Attenuator Method

While the issue of noise diode ENR uncertainty may seem like an insurmountable problem, or at least one in which any tractable solution must be extremely complex, the reality of the issue is that a very elegant solution to this problem exists; putting a cooled 20 dB attenuator[10] at the input of the amplifier is all that is needed to solve the problem. A block diagram of such a setup appears in Fig. 1.6. In addition to solving the ENR uncertainty issue, placing an attenuator in series just before the DUT has the additional benefit that it helps ensure that the source impedance remains 50 Ω. To understand why the cooled attenuator helps, we can begin by writing the new measured Y-factor as

$$Y = \frac{T_{hot}/L + T_l(L-1)/L + T_e}{T_{cold}/L + T_l(L-1)/L + T_e}, \tag{1.6}$$

where L is the linear value of the attenuator and T_l is the physical temperature of the attenuator. Using equations (1.4) and (1.6), the uncertainty that results from the noise diode specification when

[10]The value of attenuation can be changed based upon the application.

Table 1.2: Noise temperature uncertainty as a function of ENR uncertainty

Setup	ΔENR						
	0.001dB	0.003dB	0.01dB	0.03dB	0.10dB	0.15dB	0.20dB
Standard method	0.068 K	0.200 K	0.70 K	2.0 K	7.0 K	10.2 K	13.7 K
Cooled 20dB attenuator	0.005 K	0.014 K	0.048 K	0.14 K	0.48 K	0.73 K	0.98 K

using the cold attenuator method can be derived and is given as

$$\Delta T_{e,attn} = \frac{(T_{hot}/L - T_{cold}/L)\,(T_{cold} + T_l L/(L-1) + T_e)}{T_{hot}/L + \Delta T_{hot}/L - T_{cold}/L}\left(10^{\Delta ENR/10} - 1\right)$$

$$\approx \left(T_{cold}\frac{1}{L} + T_l\frac{L}{L-1}\right)\left(10^{\Delta ENR/10} - 1\right). \tag{1.7}$$

While the cooled attenuator method can give excellent results, there is one major caveat; in order for the results to be accurate and reproducible, the loss before the attenuator must be time invariant and the aggregate loss prior to the DUT must be known very accurately. If this is the case, the system can be calibrated by comparing results with national laboratories such as NIST. Generally, the calibration required is related to temperature loss along the input line and constitutes a 1–2 K offset in the effective temperature of the cooled attenuator[11].

1.3 Current Status of Research in the RF and Noise Performance SiGe HBTs at Cryogenic Temperatures

In this work, we are interested in understanding the cryogenic performance of SiGe HBTs both in order to determine their ultimate performance limitations, and to apply them to extremely low-noise cryogenically cooled amplifiers. Thus we are interested not only in how the dc and RF terminal characteristics of the devices change as a function of temperature, but also in how the small-signal parameters and noise characteristics of the device change with cooling.

While there is quite an active community engaged in the study and optimization of silicon-germanium HBT technology, the body of research on the properties of these devices at cryogenic temperatures is quite small, with interested parties for the most part focusing of operation at 77 K. Cressler and his colleagues are responsible for much of the pioneering work in the field including the design of a SiGe HBT designed especially for operation at 77 K [27]. In addition, his research group

[11]The correction is usually made as a correction due to a loss at a given physical temperature. Explicitly, the correction is given as $-T_l L/L - 1$.

has made significant contributions in the area of device modeling and circuit design at cryogenic temperatures [28, 29, 30]. In this work, Cressler has shown that the transport properties of SiGe HBTs are enhanced at cryogenic temperatures, thereby making them extremely promising in terms of cryogenic applications.

Despite their potential for cryogenic applications, the extent to which the cryogenic RF and noise properties of SiGe HBTs have been investigated is quite limited. The majority of the reported cryogenic RF measurements are in terms of f_t and f_{max}, rather than the intrinsic small-signal parameters [31, 32, 33, 34, 35]. Furthermore, in the cases in which small-signal model parameters were extracted at cryogenic temperatures, in one case data were only provided at 77 K and in another case, data were only provided at a single bias point at 78, 123, and 300 K [36, 37]. Thus, the information needed to understand how the small-signal behavior of a SiGe HBT changes as the device is cooled does not exist in the literature, making such a study quite valuable.

The amount of research on the noise performance of SiGe HBTs at cryogenic temperatures is even more limited. The noise parameters of single devices have been measured by Provost et al. at 78 K using the 50 Ω method and at 85 K by Banerjee et al. using a source pull method [37, 38]. In both cases, the noise performance of the device is not studied as a function of bias and no fundamental limits are placed upon the noise performance.

The major contributions presented in this thesis are 1) a thorough study of the dc and RF characteristics of SiGe transistors at cryogenic temperatures in terms of equivalent circuit models, 2) a study of the ultimate cryogenic noise performance limitations of SiGe HBTs in terms of a figure of merit which takes gain into account, and can therefore be used to accurately predict achievable system performance, 3) the development of accurate cryogenic noise models for SiGe devices, and 4) the demonstration of state-of-the-art SiGe low-noise amplifiers in both discrete and MMIC form.

The results described in this dissertation have been presented in the *2007 IEEE Transactions on Microwave Theory and Techniques* [39], in the *Proceedings of the 2008 IEEE International Microwave Symposium* [8], at the *2008 URSI General Assembly* [40], in the *Proceedings of the 2008 IEEE Bipolar Circuits Technology Meeting* [41], in the *2009 IEEE Microwave and Wireless Component Letters* [42], in the *2009 Review of Scientific Instrumentation* [43], and in the *Proceedings of 20th International Symposium on Space Terahertz Technology* [44]. Additional publications are in preparation [45, 46].

Part I

Theory

Chapter 2

A Primer on SiGe Transistors

This chapter is dedicated to the operating principles and performance of bipolar transistors and to the features that make SiGe heterojunction bipolar transistors special. The discussion begins with an analysis of bipolar devices in terms of energy band diagrams; an exercise that leads to a physical understanding of the terminal currents, high-frequency performance, and fundamental limitations of these devices. This is followed by a similar discussion of the physics of SiGe HBTs in which it is shown that the bandgap engineering in SiGe transistors is responsible for considerable performance enhancements. Next, the equivalent small-signal circuit model of a bipolar transistor is presented and explained. With the small-signal model and a basic understanding of the device physics in place, the noise performance of bipolar transistors is discussed and limitations are presented. Finally, the chapter ends with a discussion of higher-order effects that are typically neglected in the analysis of bipolar devices, but nonetheless prove to be critical in the actual operation of state-of-the-art devices. In this chapter, it will be assumed that the devices are operating at room temperature. Cryogenic operation will be discussed in Chapter 3.

2.1 The Difference Between Si and SiGe Transistors: Device Physics

In 1948, William Shockley filed the initial patent for the bipolar junction transistor (BJT) [47] and nine years later, Herbert Kroemer introduced the idea of a heterojunction bipolar transistor (HBT) in his seminal paper, "Theory of a Wide-Gap Emitter for Transistors" [48]. Kroemer's work was motivated by his desire to identify a way to decouple a bipolar transistor's dc current gain (β_{DC}) from the ratio of doping in its base to that of its emitter, and it earned him the 2000 Nobel

Prize in Physics[1] [50]. During his Nobel lecture, Kroemer emphasized the significance of the unique bandgap engineered nature of heterojunction based devices [51]:

> Whenever I teach my semiconductor device physics course, one of the central messages I try to get across early is the importance of energy-band diagrams. I often put this in the form of "Kroemer's Lemma of Proven Ignorance":
>
> *If, in discussing a semiconductor problem, you cannot draw an Energy-Band-Diagram, this shows that you don't know what you are talking about,*
>
> with the corollary:
>
> *If you can draw one, but don't, then your audience won't know what you are talking about.*
>
> Nowhere is this more true than in the discussion of heterostructures, and much of the understanding of the latter is based on one's ability to draw their band diagrams—and knowing what they mean.

As Kroemer implies, a strong grasp of the band diagrams of SiGe HBTs and how they differ from those of Si devices is critical in understanding the operation of SiGe devices. Thus, the discussion of SiGe devices will begin with a comparison of Si and SiGe band diagrams.

2.1.1 Fundamentals of Bipolar Transistor Physics

In this section, an introduction to the basic operating principle of the bipolar transistor is given[2]. A schematic representation of a typical silicon npn BJT appears in Fig. 2.1(a). The device consists of a pair of pn junctions that have been butted together such that they are sharing a single p-doped region. The doping of the devices is such that one of the n-doped regions, called the emitter, is very heavily doped whereas the other n-doped region, called the collector, is only moderately doped. Finally, the p-doped region in the center of the device is called the base, and is doped at an intermediate level. In Fig. 2.1(a), the voltages have also been drawn as required to bias the device in the forward active region[3].

[1] Although the prize was awarded 43 years after Kroemer published the basis for HBTs, it only 4 years earlier that these devices had been first commercially produced in silicon. [49]!

[2] In this section, only final results will be given. For derivations of key results, see Appendix A

[3] The base voltage must be positive with respect to the emitter. However, it is acceptable for the collector voltage to be slightly below that of the base for some devices.

2.1.1.1 Terminal Currents

The operation of a bipolar device under forward active operation can be understood conceptually by studying the energy band diagram shown in Fig. 2.1(b). The emitter region is heavily doped meaning that there are a large number of ionized impurities, leading to a large number of electrons in the conduction band. Thus, there will be a diffusion current of electrons injected from the emitter to the base with magnitude equal to the number of electrons that have enough thermal energy to overcome the base–emitter electrostatic-potential-barrier, which has height equal to $q\left(V_{0,BE} - V_{BE}\right)$, where $V_{0,BE} = kT_a/q \cdot \ln\left(N_{AB}^{-}N_{DE}^{+}/n_{io}^{2}\right)$, N_{AB}^{-} is the ionized acceptor concentration[4] in the base, N_{DE}^{+} is the ionized donor concentration in the emitter, q is the charge of an electron, and n_{io} is the intrinsic carrier concentration [52, 53]. If the base is sufficiently short[5], we can neglect recombination in the base and assume that all of the electrons that diffuse into the base are swept into the collector via the electric field across the collector-base junction. Thus, as the distribution of thermal energy among the electrons in the conduction band is approximately Boltzmann distributed, the collector current density is exponentially dependent on the barrier height [30, 52]:

$$J_C \approx \frac{kT_a\mu_{nb}n_{io}^{2}}{W_B N_{AB}^{-}} e^{qV_{BE}/kT_a} = N_{DE}^{+}\frac{kT_a\mu_{nb}}{W_B} e^{-q(V_{0,BE}-V_{BE})/kT_a}, \tag{2.1}$$

where μ_{nb} is the minority carrier mobility in the base. Similarly, due to the base doping level, there will be a large number of ionized acceptor impurities in the base valence band leading to a diffusion current of holes from the base to the emitter. Once again ignoring recombination current in the

[4] In this work, the dopant concentration in a region Y is assigned the variable N_{XY}, where X is an indicator as to the type of dopant (i.e., acceptor or donor). Furthermore, if we are discussing the ionized impurity concentration as opposed to the net impurity concentration, a superscript is used to indicate the charge of the ionized dopants.

[5] i.e., much shorter than a diffusion length

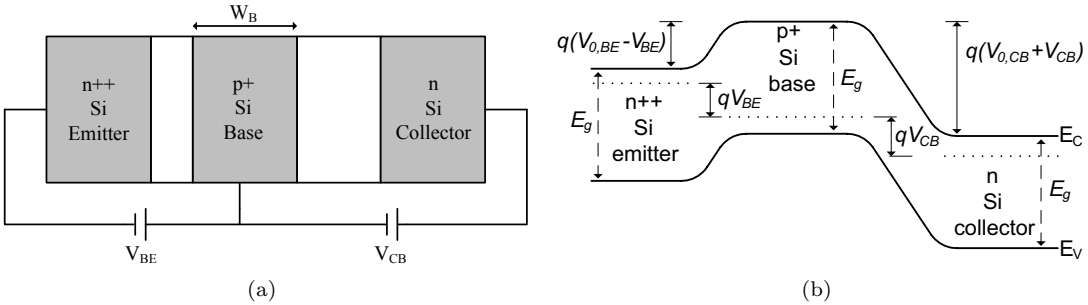


Figure 2.1: (a) Basic BJT structure. The white areas indicate the base–emitter and base–collector depletion regions. (b) Energy-band-diagram for a standard bipolar transistor under forward active bias. The Fermi levels are indicated by dotted lines in each region and would line up under zero bias. Note that the bandgap, E_g, is the same in all regions of the device.

base[6], the base current density can be written [30, 52]:

$$J_B \approx \frac{kT_a\mu_{pe}n_{io}^2}{L_{PE}N_{DE}^+}e^{qV_{BE}/kT_a} = N_{AB}^-\frac{kT_a\mu_{pe}}{L_{PE}}e^{-q(V_{0,BE}-V_{BE})/kT_a}, \tag{2.2}$$

where μ_{pe} and L_{PE} are the hole mobility and diffusion length in the emitter. Thus, for a standard npn bipolar transistor, the dc current gain is approximated as:

$$\beta_{DC} \equiv \frac{J_C}{J_B} \approx \frac{\mu_{nb}}{\mu_{pe}}\frac{L_{PE}}{W_B}\frac{N_{DE}^+}{N_{AB}^-}, \tag{2.3}$$

where W_B is the base width and L_{PE} is the diffusion length for holes injected into the emitter[7]. As explained in Section 2.2, the RF noise properties of SiGe bipolar devices in the low-GHz frequency range are largely determined by β_{DC}, with higher dc current gain corresponding to lower noise. Thus, it makes sense to investigate what parameters can be used in order to optimize β_{DC}. Referring to equation (2.3), β_{DC} is determined by three ratios:

1) μ_n/μ_p. As mobility is a material property, it is assumed that this is not a tunable parameter for standard bipolar devices. From Table 2.1, $\mu_n/\mu_p \approx 2.8$, in the low-doping limit.

2) L_{PE}/W_B. This ratio tends to increase with technology node, but otherwise cannot be easily engineered.

3) N_{DE}/N_{AB}. The emitters of modern silicon bipolar transistors are formed by depositing a layer of n-doped polysilicon layer on top of the base layer and then performing diffusion through an annealing step to form an n-doped single crystal layer between the deposited polysilicon layer and the base [54]. Thus, the doping level in the emitter is not easily tuned. On the other hand, controlling the base doping is possible. Therefore, this ratio can be controlled.

Clearly, the strongest knob we have to adjust β_{DC} is the ratio of doping in the emitter to that in the collector. However, there are serious physical limitations as to how large this ratio can be made as the base sheet resistance[8],

$$R_{B,sheet} = \frac{1}{q\mu_{pb}N_{AB}^-W_B} \quad (\Omega/\square)\,, \tag{2.4}$$

[6] i.e., attributing all of the base current to back injected holes from the base to the emitter

[7] For a short emitter, L_{PE} should be replaced by W_E.

[8] This is the sheet resistance seen flowing into the base from the base terminal, not the sheet resistance seen flowing across the base from the emitter to the collector.

Table 2.1: Properties of semiconductor materials at 300K [52]

	E_g (eV)	μ_n (cm^2/V·s)	μ_p (cm^2/V·s)	m_n^*/m_0 (m_l, m_t)	m_p^*/m_0 (m_{lh}, m_{hh})	a Å	ϵ_r -	Density (g/cm^3)	Melting Pt. °C
Si	1.11	1350	480	0.98,0.190	0.160,0.49	5.43	11.8	2.33	1415
Ge	0.67	3900	1900	1.64,0.082	0.040,0.28	5.65	16.0	5.32	936
GaAs	1.43	8500	400	0.067	0.074,0.50	5.65	13.2	5.31	1238
InP	1.35	4000	100	0.077	0.089,0.85	5.87	12.4	4.79	1070
InSb	0.18	100000	1700	0.014	0.015,0.40	6.48	17.7	5.78	525

is inversely proportional to the base doping [55, 56]. Thus, there is a direct trade-off between base resistance and β_{DC} due to the fact that both of these parameters are highly dependent on N_{AB}.

2.1.1.2 Transit Times

Bipolar junction transistors are often referred to as charge control devices. This is due to the fact that internal to the device, the transfer current is being controlled by the minority carrier charge distribution in the neutral base region. In the dc limit, the collector current is determined by the static minority carrier charge in the base. However as the frequency deviates from dc, a majority carrier must be supplied to the base for every minority carrier stored in order to maintain charge neutrality in the base. The ratio of base charge modulation to collector current modulation will increase with frequency, and is an important parameter in evaluating the high-frequency limitations of the device. The total forward transit time, that is the delay from when a voltage change occurs at the base terminal to when an electron emerges from the collector terminal, is given as [30, 54]

$$\tau_{ec} \approx \frac{\partial I_C}{\partial Q_n} = \tau_e + \tau_b + \tau_{cbd} + \frac{kT_a}{qI_C}\left(C_{jbe} + C_{jcb}\right) + r_c C_{jcb}, \tag{2.5}$$

where τ_e is the emitter charge storage time, τ_b is the base transit time, τ_{cbd} is the base–collector depletion region transit time, C_{jbe} and C_{jcb} are junction capacitances associated with the base–emitter and base–collector space charge regions, and r_c is the collector resistance. For a silicon BJT with a short emitter, the time constants are given as [30, 54],

$$\tau_e \approx \frac{q}{2kT_a}\frac{W_E^2}{\mu_{pe}\beta_{DC}} \tag{2.6}$$

$$\tau_b \approx \frac{qW_B^2}{2\mu_{nb}kT_a}, \tag{2.7}$$

and

$$\tau_{cbd} \approx \frac{W_{CBD}}{2v_{sat}}, \tag{2.8}$$

where W_{CBD} is the width of the depletion region and v_{sat} is the electron saturation velocity.

A very common figure of merit for a bipolar device is the unity-current-gain cut-off frequency, $f_t = 1/\tau_{ec}$, which is the frequency at which the short circuit ac-current gain is equal to one. Generally, foundries try to maximize f_t while maintaining reasonable breakdown voltages[9]. Referring to equations (2.5)-(2.8), it is apparent that minimization of τ_{ec} requires that 1) β_{DC} be maximized, 2) W_E and W_B be minimized, 3) W_{CBD} be minimized, 4) C_{jbe} and C_{jcb} be minimized, and 5) r_c be minimized. Thus, minimization of τ_{ec} requires a careful tradeoff between the dopant concentration in each region of the device.

2.1.2 SiGe Heterojunction Bipolar Transistors

As discussed above, a fundamental shortcoming of Si bipolar transistors is the inherent tradeoff that must be made between the dc current gain and the base resistance. This tradeoff comes about primarily because the holes being back injected into the emitter and the electrons being injected into the base each see a potential barrier of the same height. Therefore, for a given thermal excitation, the electron and hole currents both feel the same thermal push (i.e., $\exp\{(V_{BE} - V_0)/V_T\}$) and β_{DC} has to be optimized by making the supply of mobile electrons in the emitter higher than that of mobile holes in the base. Hence, increasing β_{DC} requires reducing N_{AB}, which in turn increases the base resistance. As it turns out, one can circumvent this limitation by introducing Ge into the base material.

In his seminal 1957 paper, "Theory of a Wide-Gap Emitter" Herbert Kroemer showed that if the emitter material were to have a wider bandgap than the base material, the result would be that minority carriers injected from the emitter to the base would see a smaller barrier than the minority carriers back injected from the base to the emitter, resulting in an exponential increase in β_{DC} [48]. For instance, if the difference in the emitter and base bandgaps is ΔE_g eV, then the value of β_{DC} for a device with a wide-bandgap emitter will be a factor of $e^{\Delta E_g/kT_a}$ times larger than that of a identically doped device without a wide-bandgap emitter. Furthermore in a second paper published in 1957, Kroemer postulated that, by engineering a decreasing bandgap across the base with the maxima on the emitter side, it would be possible to set up a quasi electric field, thereby reducing the base transit time significantly [58].

[9] According to the Johnson limit, the f_t-breakdown voltage product is a constant related to material properties [57].

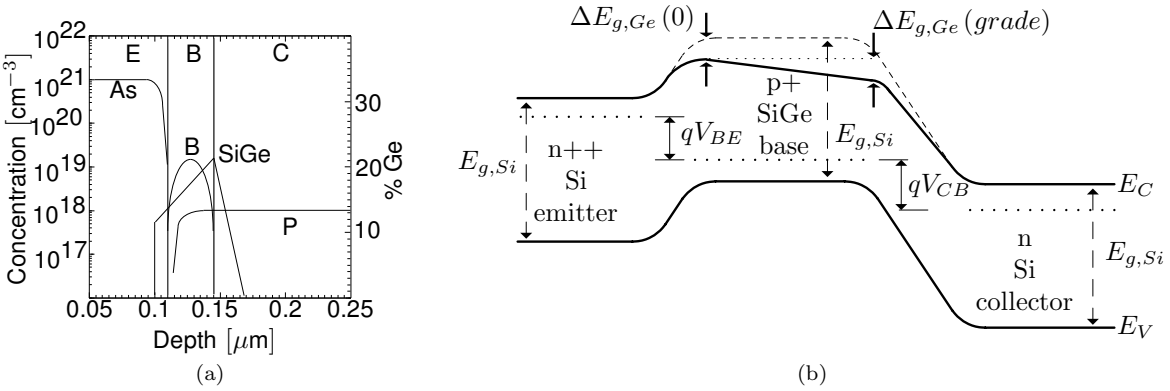


Figure 2.2: (a) A typical doping and Ge profile for a state-of-the-art SiGe HBT [59]. (b) Band Diagram for a SiGe HBT indicating deviation from that of a pure silicon transistor. Apparent bandgap narrowing effects that are discussed below have not been included in the band diagram.

Although it took thirty years for the materials processing technology to progress to the point at which Kroemer's ideas could be applied to transistors fabricated in silicon materials systems, his work has been well rewarded, as the field of SiGe HBTs would be non-existent without his theory [49]. A doping profile and band diagram for a typical state-of-the-art SiGe HBT appears in Fig. 2.2. Referring to the doping profile, we see that there is a position dependent Ge content in the base. From Table 2.1 it can be seen that the bandgap of Ge is 0.67 eV, which is significantly less than 1.11 eV (the bandgap of silicon). Thus, by introducing a small amount of Ge to the base, it is possible to reduce the bandgap in the alloy considerably from that of pure silicon. Furthermore, by grating the Ge content as a function of depth into the base, the bandgap can be reduced along the base, resulting in the reduction of transit time that was predicted by Kroemer. The resulting band structure appears in Fig. 2.2(b) along with the band structure of an identically doped Si device. In the following section, the benefits of introducing the Ge in the base will be looked at quantitatively.

2.1.3 Terminal Currents

The collector current density of a SiGe HBT is derived in Appendix A.3 and is given as

$$J_C \approx n_{io,Si}^2 \tilde{\gamma}\tilde{\eta}\frac{\mu_{nb,Si}}{N_{AB}^- W_b}\Delta E_g\,(grade) e^{\Delta E_g^{app}/kT_a} e^{\Delta E_{g,Ge}(0)/kT_a}\left(e^{qV_{BE}/kT_a}-1\right) \tag{2.9}$$

where $\tilde{\eta} = (\mu_{nb})_{SiGe}/(\mu_{nb})_{Si} > 1$, $\tilde{\gamma} = (N_C N_V)_{SiGe}/(N_C N_V)_{Si} < 1$, $\mu_{nb,Si}$ and $n_{io,Si}$ are the electron mobility and intrinsic carrier concentration in silicon, and $\Delta E_{g,app}$ is an apparent bandgap

reduction related to heavy doping effects. To gain insight, equation (2.9) can be simplified into the standard form,

$$J_C \approx J_{C0}\left(e^{V_{BE}/V_T}-1\right), \tag{2.10}$$

where

$$J_{C0} = n_{io,Si}^2\tilde{\gamma}\tilde{\eta}\frac{\mu_{nb,Si}}{N_{AB}^- W_b}\Delta E_g\,(grade)e^{\Delta E_g^{app}/kT_a}e^{\Delta E_{g,Ge}(0)/kT_a} \tag{2.11}$$

is the collector-current saturation current and is exponentially enhanced due to germanium induced bandgap reduction. The base current of a SiGe HBT is the same as that of an identically doped silicon BJT and can be written as

$$J_B \approx \frac{kT_a\mu_{pe}n_{io}^2}{L_{PE}N_{DE}^+}e^{qV_{BE}/kT_a} = \frac{q}{G_e}e^{qV_{BE}/kT_a}, \tag{2.12}$$

where $G_e = N_{DE}^+ L_{PE}/D_{PE}n_{io,e}^2$ is the emitter Gummel number [60]. Thus, the dc current gain of a SiGe HBT can be written as

$$\begin{aligned}\beta_{DC} &\approx \frac{\mu_{nb,Si}L_{pe}N_{DE}^+}{\mu_{pe}W_B N_{AB}^-}\tilde{\gamma}\tilde{\eta}\frac{\Delta E_g, Ge\,(grade)}{kT_a}e^{\Delta E_{g,app}/kT_a}e^{\Delta E_g(0)/kT_a}\\ &= \beta_{DC,Si}\left(\tilde{\gamma}\tilde{\eta}\frac{\Delta E_g, Ge\,(grade)}{kT_a}e^{\Delta E_{g,app}/kT_a}e^{\Delta E_g(0)/kT_a}\right).\end{aligned} \tag{2.13}$$

In order to evaluate equation (2.13), it is necessary to know the appropriate expressions for the apparent- and Ge-induced bandgap narrowing. The Ge induced bandgap reduction for a compressively strained SiGe film at room temperature can be estimated as a function of the Ge content, x, as [54]

$$\Delta E_{g,Ge} \approx 0.96x - 0.43x^2 + 0.17x^3, \tag{2.14}$$

and the room temperature value of $\Delta E_{g,app}$ can be estimated as a function of dopant concentration as [61]

$$\Delta E_{g,app} \approx 18\times 10^{-3}\ln\left\{\frac{N_{AB}^-}{N_{DE}^+}\right\}. \tag{2.15}$$

Equation (2.15) is valid for the case in which both the emitter and base dopant concentrations are greater than 7×10^{17} cm^{-3}. For the case in which the base is doped below this level, the doping induced apparent bandgap narrowing in the base is negligible and the expression for $\Delta E_{g,app}$ must be written as

$$\Delta E_{g,app} \approx -18\times 10^{-3}\ln\left\{\frac{N_{DE}^+}{7\times 10^{17}\text{ cm}^{-3}}\frac{300\text{ K}}{T_a}\right\} \tag{2.16}$$

Using equations (2.13)–(2.15) and assuming that $\widetilde{\gamma}$ and $\widetilde{\eta}$ are close to unity, we can quickly estimate the effect of the Ge content on β. For instance, for the case in which $N_{AB}^{-} = 5 \times 10^{18}$ cm^{-3}, $N_{DE}^{+} = 10^{20}$ cm^{-3}, the Ge content is 20% at the emitter side of the base, and there is a 10% Ge grating, then an improvement of over 400 in the dc current gain is obtained. It is clear that introduction of Ge into the base of a bipolar transistor affects the dc current gain quite favorably. In the next section, the effect that the Ge content has on ac performance will be discussed.

2.1.3.1 Transit-Times for SiGe HBTs

As a result of the quasi-electric field due to the grated Ge content in the base, it is expected that the forward transit time will be greatly reduced. For instance, a modest Ge ramp of 5% across a 30nm SiGe base results in an effective electric field of over 15 kV/cm, which is enough to cause the minority carriers to reach close to the saturation velocity. In addition, we expect that the emitter charging time constant should be reduced drastically due to the lower barrier which carriers must overcome before being injected into the base. On the other hand, we expect little change in the collector-base SCR transit time. Expressions for each of the time constants have been derived and the resulting time constants are expressed as

$$\tau_{e,SiGe} \approx \frac{q}{2kT_a} \frac{W_E^2}{\mu_{pe,Si} \cdot \beta_{DC,SiGe}} = \tau_{e,Si} \frac{\beta_{DC,Si}}{\beta_{DC,SiGe}}, \tag{2.17}$$

$$\tau_{b,SiGe} \approx \frac{qW_b^2}{\widetilde{\eta}\mu_{nb,Si}} \frac{1}{\Delta E_{g,Ge}\,(grade)} = \frac{\tau_{b,Si}}{\widetilde{\eta}} \frac{kT_a}{\Delta E_{g,Ge}\,(grade)}, \tag{2.18}$$

and

$$\tau_{cbd,SiGe} = \frac{W_{CBD}}{2v_{sat}} = \tau_{cbd,Si}. \tag{2.19}$$

As expected, both the emitter charging time constant and the base transit time are greatly reduced due to the addition of a Ge grating to the base layer.

2.1.4 BJT Small Signal Model

In order to determine the noise performance of a BJT, one needs information about the noise sources as well as the small-signal equivalent circuit. In this section, the small-signal model is developed. A drawing of a typical SiGe HBT structure appears in Fig. 2.3. The device is a vertical structure with the collector being closest to the back side of substrate and the emitter closest to the surface. The fabrication process begins with a p-doped substrate. A low resistance (5–10$\Omega/\square$)

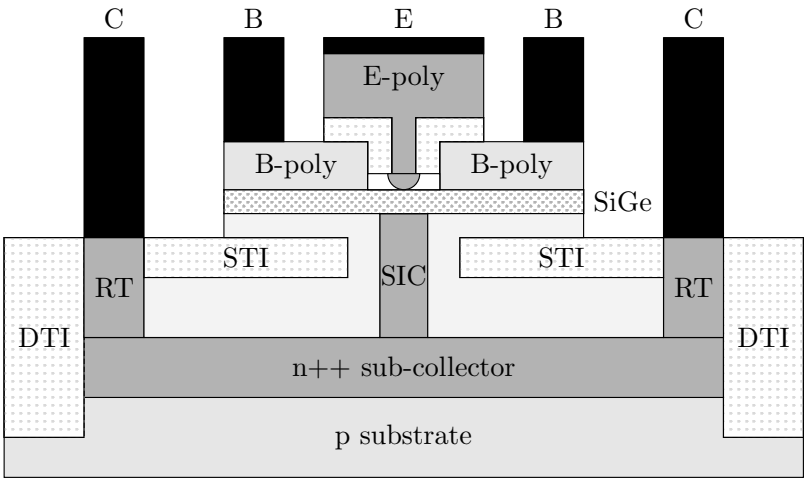


Figure 2.3: Typical state-of-the-art HBT structure

sub-collector is then formed on the surface of the wafer. Next, a lightly doped n-region is epitaxially grown providing the surface to grow the rest of the HBT. However, before continuing with the growth, deep and shallow trench isolation structures (DTI and STI) are formed, collector reach through sinkers (5–10Ω each) are implanted to allow for a metallurgical contact to the sub-collector, and a selectively implanted collector is formed by a high energy implant implantation [54, 62]. Once the collector and isolation structures are defined, the $Si_{1-x}Ge_x$ base is grown and contacted via a polysilicon extrinsic base. Finally, an in-situ doped poly-emitter is formed.

Equipped with an understanding of the physical structure of a SiGe HBT, it is a rather straightforward task to devise an equivalent circuit model for the device. The equivalent circuit model of the device is given in Fig. 2.4. The intrinsic transistor is represented in terms of the standard hybrid-π model and, as will become apparent shortly, the component values can be modeled using the information described above. In addition to the intrinsic HBT, there are four additional components: 1) r_b, which is required to model the resistance of the polysilicon extrinsic base as well as that of the SiGe intrinsic base, 2) r_e, which is needed to model the polysilicon emitter, 3) r_c which is needed to model the parasitic collector resistances, and 4) C_{CS} which is a depletion capacitance between the collector and ground that arises due to the fact that the collector was deposited on a semiconducting substrate.

We will begin the discussion with the intrinsic circuit, ignoring the delay term τ_d. Referring to equation 2.9, it can be shown that the transconductance, g_m, is given as:

$$g_m \equiv \frac{\partial I_C}{\partial V_{BE}} = \frac{I_C}{V_T}, \tag{2.20}$$

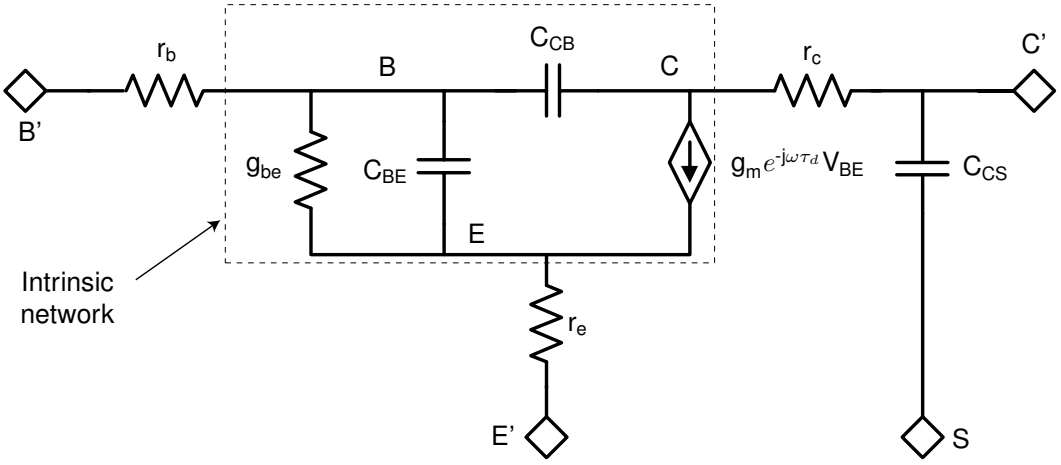


Figure 2.4: Small-signal equivalent circuit for SiGe HBT

where the thermal voltage is defined as $V_T \equiv kT_a/q$. Similarly, through inspection of equation (2.12) is can be seen that the small-signal conductance between the base and emitter is given as:

$$g_{be} \equiv \frac{\partial I_B}{\partial V_{BE}} = \frac{\partial I_B}{\partial I_C}\frac{\partial I_C}{\partial V_{BE}} = \frac{g_m}{\beta_{AC}}, \tag{2.21}$$

where β_{AC} is the ac current gain of the device. The intrinsic capacitances were indirectly discussed in Section 2.1.3.1 and are repeated here for completeness:

$$\begin{aligned} C_{BE} &= g_m\left(\tau_{b,SiGe} + \tau_{e,SiGe} + \tau_{cbd,SiGe}\right) + C_{jb} \\ &\approx g_m\left(\frac{qW_b^2}{\widetilde{\eta}\mu_{nb,Si}}\frac{1}{\Delta E_{g,Ge}\left(grade\right)} + \frac{W_{CBD}}{2v_{sat}}\right) + C_{jb} \end{aligned} \tag{2.22}$$

and

$$C_{CB} \approx C_{jc}. \tag{2.23}$$

It is possible to write equations for the extrinsic components based on materials properties and device structure [54]. However, as these components are highly dependent on geometry and composition, this will not be attempted here. Prior to continuing on to the noise performance of bipolar devices, we will mention that simple expressions for the f_t and f_{max} of the device can be written in terms of the small-signal model parameters as [63]

$$f_t = \left(\frac{g_m}{2\pi\left(C_{BE}+C_{CB}\right)}\right) || \left(\frac{1}{C_{CB}\left(r_e+r_c\right)}\right) \approx \frac{g_m}{2\pi\left(C_{BE}+C_{CB}\right)} \tag{2.24}$$

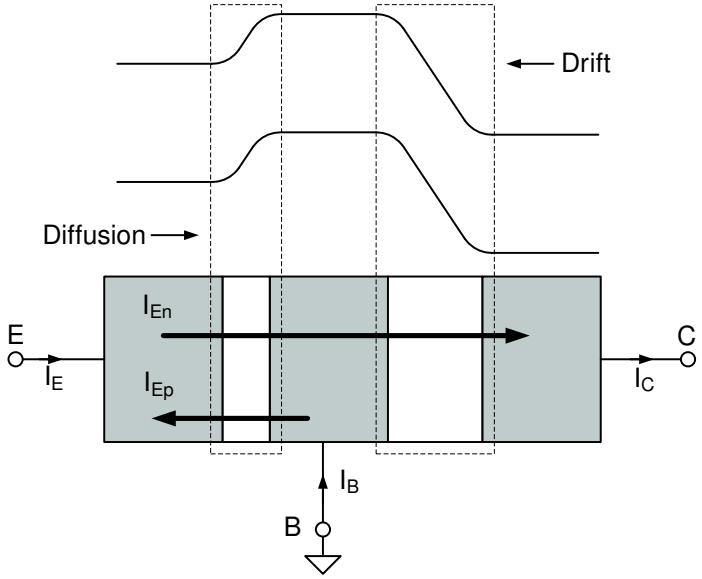


Figure 2.5: Energy and schematic diagrams for a bipolar transistor in the common-base configuration. Areas in which the drift and diffusion currents occur are highlighted. The effect of recombination in the neutral base is neglected.

and

$$f_{max} \approx \sqrt{\frac{f_t}{8\pi r_b C_{CB}}}. \tag{2.25}$$

2.2 Noise Performance of Bipolar Devices

The physical sources of noise in a bipolar device are the parasitic resistances and the diffusion currents. Modeling of the Johnson noise due to the resistances is straightforward. However, modeling the shot-noise sources resulting from the diffusion currents requires some thought. Shot noise occurs whenever dc current flows across a potential barrier as the result of the random distribution of electrons having enough energy to cross [64]. Thus shot noise is present in all forward biased diode structures as well as in any thermionic emission device, such as a vacuum tube. For noise modeling, it is helpful if we first ignore the parasitics and develop a noise model in the admittance representation. Although the model is desired for the common emitter configuration, the analysis is simplified by beginning with a representation in the common base configuration and then performing a transformation to end up with the common emitter representation.

In Fig. 2.5, the energy band diagram and schematic drawing of a npn BJT in the common base configuration are shown. Inspection shows that there are two places where diffusion current is occurring (i.e., where carriers must overcome a potential barrier): when holes are injected from the

base to the emitter, and when electrons are injected from the emitter to the base. Thus, there is an independent shot-noise source associated with both the hole and electron components of the emitter currents, and the total emitter noise current PSD can be written as [65]

$$\overline{|i_e|^2} = 2qI_{Ep} + 2qI_{En}. \tag{2.26}$$

Now, I_C is simply a delayed version of I_{En}; $I_C = I_{En}e^{j\omega\tau_n}$, where τ_n is the transit time corresponding to the transport of the emitter-base electron shot noise to the collector [66]. Thus, the collector current noise is fully correlated with the emitter electron shot-noise source and can be expressed as [66]

$$i_{n,c} = i_{n,e}e^{-j\omega\tau_n}, \tag{2.27}$$

where τ_n is a delay term associated with the transport of the shot noise from the emitter to the collector. It should be noted that τ_n is not the same as τ_{ec}; shot noise is coupled to dc currents, so the delay term does not include the time constant associated with the ac modulation of the base charge.

With the noise model in place, it is trivial to convert to the common emitter representation [66]:

$$\overline{|i_{n,b}|^2} = \overline{|i_{n,c}|^2} + \overline{|i_{n,e}|^2} - 2\,\Re\left\{\overline{i^*_{n,c}i_{n,e}}\right\}, \tag{2.28}$$

$$\overline{|i_{n,c}|^2} = 2qI_C, \tag{2.29}$$

and

$$\overline{i^*_{n,b}i_{n,c}} = \left(\overline{i^*_{n,c}i_{n,e}}\right)^* - \overline{|i_{n,c}|^2} = 2qI_C\left(e^{-j\omega\tau_n} - 1\right) \tag{2.30}$$

Finally to complete the model, thermal noise sources should be embedded to account for the resistive losses due to r_c, r_b, and r_e.

This model is interesting for several reasons. First of all, it belongs to the class of two-ports to which the 50 Ω noise characterization method can be applied as there is only one noise parameter that cannot be directly measured through dc or RF measurements. Furthermore, the unknown parameter, τ_n, only describes a correlation term and therefore cannot degrade the fundamental noise performance of the device. In other words, *it is possible to ignore the correlation term, τ_n, and put an upper boundary on what is achievable in terms of low-noise performance. Under this condition, the noise is completely determined by the dc currents, which set the magnitude of the shot-noise sources, and the physical temperature, which sets the thermal sources.* This is a very

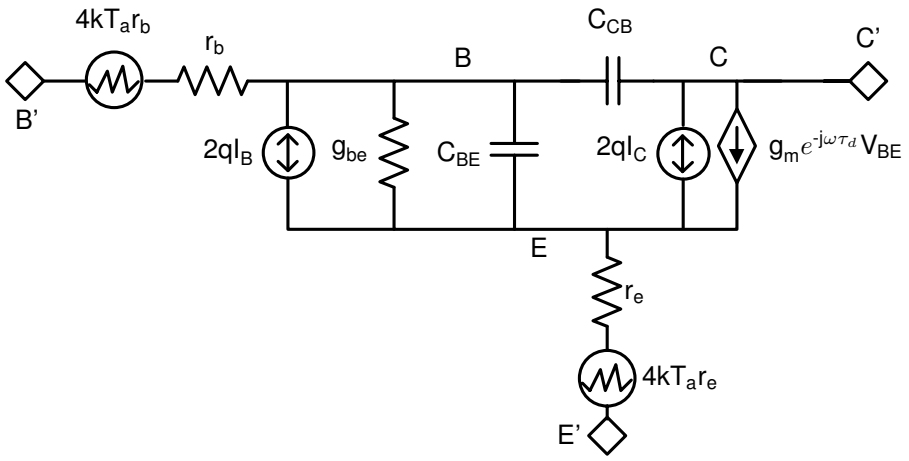


Figure 2.6: Simplified SiGe HBT noise model. The effects of the collector resistance and collector–substrate capacitance have been ignored

important as it makes it possible to estimate of the obtainable cryogenic noise performance of the devices without having to rely on error-prone on-wafer cryogenic noise measurements. In this work, τ_n is assumed to be zero in the frequencies of interest, resulting in a potential over-estimation of the noise[10].

Now that the noise model has been presented, we will proceed to discuss the noise parameters of the device. Detailed derivations appear in Appendix E and only the final results will be presented here. A schematic representation of the simplified[11] noise model appears in Fig. 2.6. The noise parameters are given as:

$$T_{min} \approx T_a n_c \sqrt{\frac{1}{\beta_{DC}}\left(1+2\frac{g_m\left(r_b+r_e\right)}{n_c}\right)+2\frac{g_m\left(r_b+r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2}, \tag{2.31}$$

$$G_{OPT} \approx \frac{g_m}{1+2g_m\left(r_b+r_e\right)/n_c}\sqrt{\frac{1}{\beta_{DC}}\left(1+2\frac{g_m\left(r_b+r_e\right)}{n_c}\right)+2\frac{g_m\left(r_b+r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2} \tag{2.32}$$

$$B_{OPT} \approx -\frac{f}{f_t}\frac{g_m}{1+2g_m\left(r_b+r_e\right)/n_c}, \tag{2.33}$$

and

$$R_n \approx \frac{n_c}{2g_m}\frac{T_a}{T_0}\left[1+2\frac{g_m\left(r_b+r_e\right)}{n_c}\right], \tag{2.34}$$

where $n_c = I_C/g_m V_T$ is the collector current ideality factor. Analysis of equations (2.31)–(2.34)

[10] As a side note, compact models such as the VBIC model do not account for the correlation of shot-noise sources [67].

[11] the collector resistance and collector–substrate capacitance have been neglected in the model

reveals several interesting features:

1) The noise resistance R_n is frequency independent. It should be noted that this is true for field-effect-transistors as well [24].

2) The frequency response of T_{min} has an identical shape as that of G_{OPT}. At low frequencies the two noise parameters are constant, and at high frequencies, their values rise proportional to f/f_t. The knee frequency which marks onset of the frequency range in which their values increase is given as $f_{knee} = f_t/\sqrt{\beta_{DC}}\sqrt{1 + n_c/2g_m\left(r_b + r_e\right)}$.

3) The optimum susceptance, B_{OPT}, is zero at dc and increases proportionally with frequency with a slope determined by f_t and $g_m/\left(1 + 2g_m\left(r_b + r_e\right)/n_c\right)$.

4) The minimum noise temperature is related to the optimum source conductance and the noise resistance as $T_{min} = 2T_0R_nG_{opt} = 2T_0N$. Thus, the sensitivity factor[12] is given as $N = T_{min}/2T_0$. This is quite interesting because it implies that for the case in which the shot-noise sources are not correlated, to first order there are only three independent noise parameters. Furthermore, Pospiezalski has shown that the sensitivity factor must fall into the range of $N \in [T_{min}/4T_0,\, T_{min}/2T_0]$, which means that the value of N is the maximum permissible [69]. As N describes the sensitivity of the device noise performance to deviations of the source impedance away from Y_{OPT}, this can be interpreted to imply that correlation can only reduce N (i.e. reduce the sensitivity of T_e to Y_S).

In terms of circuit design, it is often useful to have knowledge of the optimum source impedance, $Z_{OPT} = R_{OPT} + jX_{OPT}$, which for a HBT can be written as

$$R_{OPT} \approx \frac{\beta_{DC}}{g_m\left(1 + \beta_{DC}\left(f/f_t\right)^2\right)}\sqrt{\frac{1}{\beta_{DC}}\left(1 + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\right) + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2}, \tag{2.35}$$

$$X_{OPT} \approx \frac{\beta_{DC}}{g_m}\frac{f/f_t}{1 + \beta_{DC}\left(f/f_t\right)^2}. \tag{2.36}$$

Thus, X_{OPT} will increase from zero at low frequencies and drop at high frequencies with a maximum value of $\sqrt{\beta}/2g_m$ at $f = f_t/\sqrt{\beta_{DC}}$. Furthermore, R_{OPT} will be constant at low frequencies and roll-off at higher frequencies with a knee frequency of $f = f_t\sqrt{\beta_{DC}}$. This information is important when designing a broadband low-noise amplifier as will become apparent in Chapter 9.

[12]The sensitivity factor is an invariant parameter of device size and is therefore considered a more fundamental parameter than R_n, as the noise resistance can always be reduced by putting devices in parallel [68]. Thus, when reporting noise parameters in Chapter 7, N will be discussed as opposed to R_n

At low-frequencies, where the parasitic capacitances have very large impedances, and under the condition that $g_m\left(r_b + r_e\right)/n_c \ll 1/2$, the noise parameters simplify drastically and can be written as

$$T_{MIN,LF} \approx T_a \frac{n_{cx}}{\sqrt{\beta_{DC}}}, \tag{2.37}$$

$$R_{OPT,LF} \approx \frac{\sqrt{\beta_{DC}}}{G_m}, \tag{2.38}$$

$$X_{OPT,LF} \approx 0 \tag{2.39}$$

and

$$R_{n,LF} \approx \frac{T_a}{T_0}\frac{n_{cx}}{2G_m} \tag{2.40}$$

where $G_m = g_m/(1 + g_m r_e)$ is the extrinsic transconductance and $n_{cx} = I_C/G_m V_T$ is the extrinsic collector current ideality factor. Thus, the noise performance at low frequencies can be determined entirely from dc measurements.

2.3 High Injection Effects

Thus far, in analysis of the device operating characteristics, we have assumed low-injection conditions. However, referring to equation (2.5), it is apparent that minimization of the time constant related to the charging of the junction capacitances requires large collector current densities. For instance, given a typical base–collector capacitance of 15 fF/μm^2, the charging-time constant for the base–collector junction capacitance will be 0.4 pS/J_C, where J_C is given in mA/μm^2. Now, if we are running a device at an f_t of 200 GHz, and only budget a tenth of the total time constant to the charging and discharging of the base–collector junction capacitance, then a collector current density of 20 mA/μm^2 is required. This corresponds to a mobile charge concentration of 1.2×10^{18} cm^{-3}. As this is clearly on par with the doping concentration in the collector, it will certainly violate the low-injection assumption. Furthermore, even if we budget the entire transit time to the charging of the base–collector junction capacitance, the corresponding mobile charge concentration of 1.2×10^{17}cm^{-3} is still rather high for low-injection assumptions to be justified. Thus, it is fair to say that modern SiGe HBTs are operated in the high-injection regime in order to maximize f_t. In this section, a brief overview of some of the resulting consequences will be given.

2.3.1 Webster Effect

When computing the dc terminal currents of the SiGe and Si bipolar devices, it was assumed that the number of injected electrons in the base region was far outnumbered by the ionized acceptor impurity concentration ($n_b \ll N_A^-$). Thus, the hole concentration in the valence band, p_b was assumed to be equal to N_A^-. However, under high-injection conditions, this assumption loses validity and the correct expression for the hole concentration is $p_b = N_A^- + n_b = N_A^- + J_C/qv_d$, where v_d is the drift velocity of electrons and the collector current density is specified in A/cm^2. Considering the case of Si transistors, a first-order expression[13] for the dc current gain can be rewritten as

$$\beta_{DC,Si} \approx \frac{\mu_{nb}}{\mu_{pe}} \frac{L_{PE}}{W_B} \frac{N_{DE}^+}{p_b} \approx \frac{\mu_{nb}}{\mu_{pe}} \frac{L_{PE}}{W_B} \frac{N_{DE}^+}{N_{AB}^-} \frac{1}{1 + J_C/qN_{AB}^- v_d} = \beta_{DC0,Si} \frac{1}{1 + J_C/qN_{AB}^- v_d}. \quad (2.41)$$

Thus, at collector currents above $J_C = qN_{AB}^- v_d$, the dc current gain will roll-off as the collector current is further increased.

This effect was first described in 1954 by Webster and has since come to be known as the Webster effect [70]. Referring to equation (2.41), it can be seen that the Webster effect causes a roll-off in β_{DC} at high current levels. For example, for an electron drift velocity of 1×10^7cm·s and an ionized impurity concentration in the base of 3×10^{18}cm^{-3}, β_{DC} is degraded by a factor of two at a collector current density of 48mA/μm^2. As this doping level is typical of modern SiGe HBTs, the Webster effect is not expected to play a large role in the degradation of β_{DC}.

2.3.2 Kirk Effect

The Kirk effect is also known as base pushout and is very important in modern SiGe HBTs. In deriving the transport equations, it has been assumed that the background dopant concentration in the collector-base space charge region (SCR) is greater than the concentration of electrons being injected (i.e., $J_C/v_{sat} \ll qN_{DC}$). To consider the effect of high-injection on the collector–base SCR, we begin by writing Poisson's equation in this region, which relates the electric field to the net charge density [54]:

$$\frac{dE}{dx} = \frac{q}{\epsilon_0 \epsilon_r} \left(p - n + N_{DC}^+\right) \approx \frac{1}{\epsilon_0 \epsilon_r} \left(qN_{DC}^+ - qn\right) \approx \left(qN_{DC}^+ - J_C/v_{sat}\right), \quad (2.42)$$

[13] In order to model this effect correctly, it is required to also take the effect of the increase in p_b on the minority carrier mobility in the base into account.

where v_{sat} is the saturation velocity of electrons, p is the mobile hole charge density in the collector–base SCR, and n is the mobile electron charge density in the collector base SCR. Referring to equation 2.42, it is clear that the derivative of the electric field in the collector-base SCR depends upon the net charge in this region.

In order to satisfy charge neutrality, the net charge in the collector side of the base–collector space charge region must be equal to that in the base side of the junction such that the magnitude of the electric field is maximum at the metallurgical junction and zero at the outside edges of the SCR. Thus, under low-injection conditions, the penetration of either side of the SCR into the neutral-base and collector regions is determined by the ionized impurity concentrations in the two regions, which determines the built-in voltage, and the applied voltage which modulates the SCR width (dE/dx in each region is fixed by the background doping levels). Thus, under low-injection assumptions, the penetration of the SCR into each side of the base–collector junction is independent of the currents flowing in the device.

However, when the mobile charge density in the SCR is no longer negligible with respect to the collector background impurity concentration, the effective impurity concentration in the collector side of the base–collector SCR is reduced by J_C/qv_{sat} which, due to charge neutrality requirements, results in a widening of the SCR into the neutral collector. This will continue until the collector current density grows to be equal to $qv_{sat}N_{DC}^{+}$. At this point, the net charge density in the intrinsic collector is zero and the intrinsic collector cannot support an electric field. Thus, the collector side of the space charge region moves to the buried sub-collector. With any further increase in collector-current density, the net charge in the intrinsic collector will be positive, and the base side of the SCR will be pushed to the interface of the extrinsic and intrinsic collectors. Thus, the base will be widened by the thickness of the intrinsic collector. This widening of the base was discovered by Kirk and published in his seminal 1962 paper, "A theory of transistor cutoff frequency (f_t) falloff at high current densities" [71].

The Kirk effect is important because it increases the base transit time significantly as $\tau_b \propto W_B^2$. Referring to Fig. 2.2(a), we see that a typical doping concentration in the intrinsic collector is about 1×10^{18} cm^{-3}. Thus, at room temperature the base pushout will occur when $J_C = qv_{sat}N_{DC} = 16$ mA/μm^2. As the collector–base (Miller) capacitance is proportional $\sqrt{N_{DC}^{+}}$, C_{CB} rises as the collector doping is increased. Therefore, there is a clear tradeoff between the minimization of $\tau_{C_{jc}} = C_{jc}kT_a/qI_C$, which scales as $\sqrt{N_{DC}^{+}}/I_C$, and the onset of the Kirk effect, which moves to higher current densities as the collector doping is increased.

2.4 Summary

In this chapter, the basic theory needed to understand the remainder of the dissertation has been presented. In the first part of the chapter, the fundamentals of bipolar junction transistors and SiGe heterojunction bipolar transistors were presented with an emphasis placed on the shortcomings of standard bipolar transistors and how SiGe devices circumvent these issues. Next, the small-signal and noise performance of bipolar devices was presented. Finally, the chapter ended with a discussion of high-injection effects. In the next chapter, the properties of silicon materials and SiGe HBTs at cryogenic temperatures will be discussed.

Chapter 3

Expected Performance of SiGe Transistors at Cryogenic Temperatures

As explained in Chapter 2, the operating characteristics of SiGe transistors are tightly coupled to the underlying material properties. In order to understand the operation of SiGe devices at cryogenic temperatures, it is therefore necessary to study the temperature dependence of the physical properties of Si materials as well as those of SiGe alloys. In Section 2.1.2, we saw that the operating characteristics of a SiGe HBT can be described in terms of the operating characteristics of identically doped Si device with a perturbation applied in order to account for the effect of the Ge content in the base. Thus, a proper understanding of the cryogenic properties of the materials properties of silicon is a critical foundation required to understand the behavior of SiGe devices at low temperatures. This section begins with a study of the important material properties of Si as a function of temperature. Next, a look into the expected cryogenic performance of SiGe HBTs is presented in terms of the temperature dependence of various physical parameters. Finally, the chapter concludes with a summary discussion of the key trends and device properties pertaining to the cryogenic operation of SiGe transistors.

3.1 Properties of Silicon at Cryogenic Temperatures

Previously, it was shown that the terminal characteristics of SiGe HBTs can be described in terms of those of a Si device with identical doping profiles, with minor adjustments applied to account for the bandgap-engineered nature of SiGe devices. In this section, the temperature dependence of the fundamental material properties of silicon are investigated.

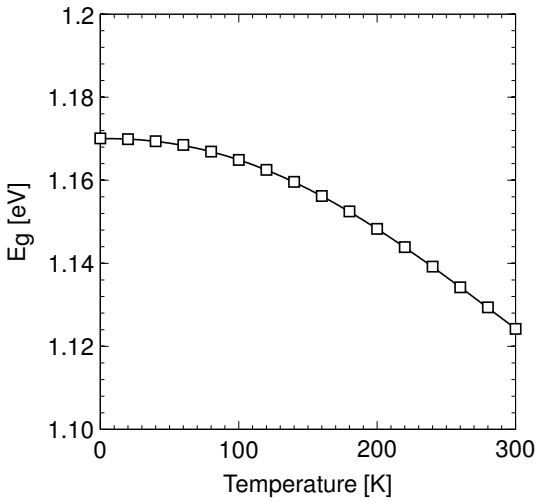


Figure 3.1: The energy bandgap versus ambient temperature for high purity silicon. Squares represent experimental data from [72] and the solid curve is a polynomial fit.

3.1.1 Bandgap: Intrinsic Si

Understanding the temperature dependence of the bandgap of intrinsic silicon material is extremely important, as the diffusion currents in a bipolar device are exponentially related to E_g. The bandgap of high purity intrinsic silicon was measured from 2–300K by Bludau, Onton, and Heinke using wavelength-modulation spectroscopy and the data were presented in tabular format in [72]. In order to predict device performance, a least-squares fourth-order polynomial fit was performed to the bandgap data, resulting in the following approximation of Bludau and his co-authors' data:

$$E_g \approx 1.17 + 5.65 \times 10^{-6} T_a - 5.11 \times 10^{-7} T_a^2 - 8.03 \times 10^{-10} T_a^3 + 2.50 \times 10^{-12} T_a^4. \tag{3.1}$$

The data from this study are plotted in Fig. 3.1.1 along with the fourth-order polynomial least squares fit. As intrinsic silicon is cooled from 300K to 0K, the bandgap changes by just over 4%. While this may not seem like a large deviation, it results in over 13 orders of magnitude change in the value of $\exp\{E_g/kT_a\}$ as the temperature is reduced from 300 to 18 K.

3.1.2 Effective Masses: Intrinsic Si

Because both the electron and hole effective masses influence the device properties, the temperature dependence of each of these quantities is important. The electron density-of-states effective mass, m^*_{de}, is a combination of the longitudinal and transverse electron effective masses and is given

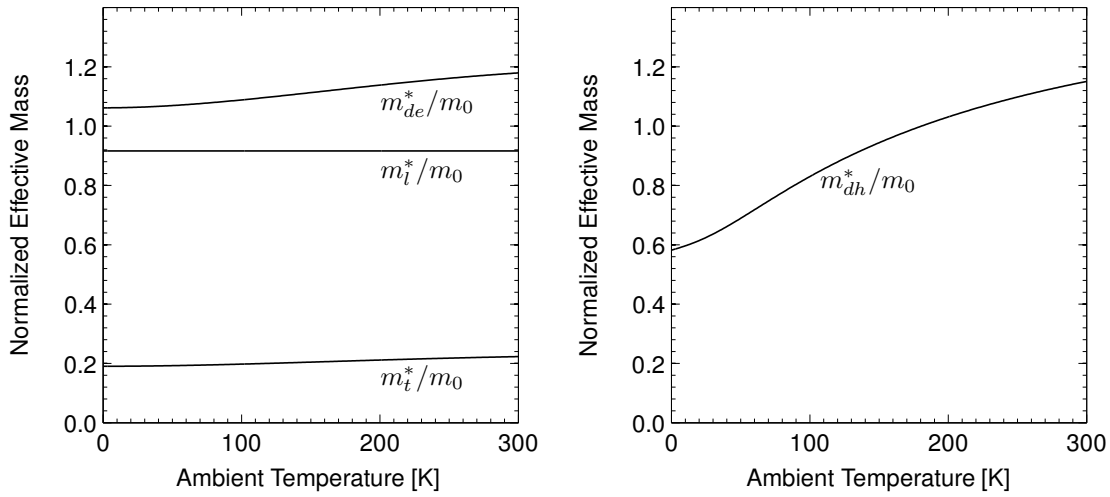


Figure 3.2: (a). The electron density of states effective mass as function of temperature. The transverse and longitudinal components are also plotted. (b). The hole density-of-states effective mass as a function of temperature.

by [52]

$$m^*_{de} = \left(6\sqrt{m^*_l m^{*2}_t}\right)^{2/3}. \tag{3.2}$$

Thus, the temperature dependence of the electron density-of-states effective mass depends upon that of the longitudinal and transverse effective masses, m^*_l and m^*_t. Studies have been carried out as to the temperature dependence of these parameters and it has been found that m^*_l is relatively temperature independent whereas m^*_t can be characterized by a fourth order polynomial [73]:

$$\frac{m^*_l}{m_0} \approx 0.9163 \tag{3.3}$$

and

$$\frac{m^*_t}{m_0} \approx 0.19049 - 2.0905\times10^{-6}T_a + 9.8985\times10^{-7}T_a^2 - 2.6789\times10^{-9}T_a^3 + 2.0270\times10^{-12}T_a^4. \tag{3.4}$$

A plot showing the temperature dependence of the electron density-of-states effective mass as well as the components m^*_l and m^*_t appears in 3.2(a). It can be seen that m^*_{de} decreases only slightly as T_a is reduced from 300 to 0 K.

The density-of-states effective mass for holes is a combination of the effective masses for holes in

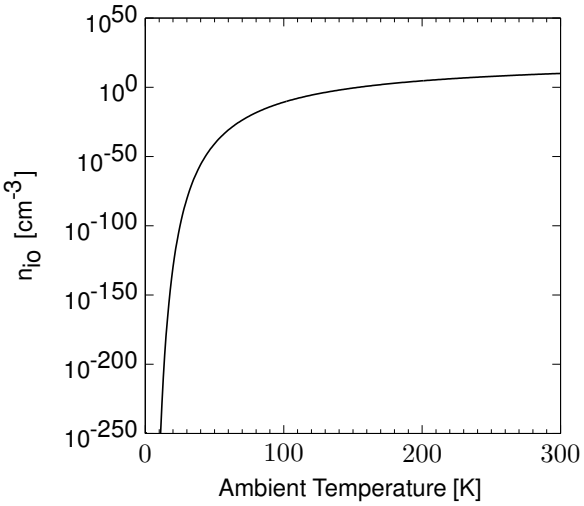


Figure 3.3: The intrinsic carrier concentration in silicon as a function of temperature. Temperature effects due to both the bandgap and density-of-states effective are accounted for.

the light hole band, the heavy hole band, and the split-off valence band[1] and is given as [73]

$$m_{dh}^{*} = \left[\left(m_{hh}^{*}\right)^{3/2} + \left(m_{lh}^{*}\right)^{3/2} + \left(m_{soh}^{*}\right)^{3/2} \right]^{2/3}. \tag{3.5}$$

As it turns out, the density of states effective mass is a complex function of temperature. However, a reasonable fit to m_{dh}^{*} has been given by Green in [74]:

$$\frac{m_{dh}^{*}}{m_0} \approx \left(\frac{0.444 + 0.361 \times 10^{-2} T_a + 0.117 \times 10^{-3} T_a^2 + 0.126 \times 10^{-5} T_a^3 + 0.303 \times 10^{-8} T_a^4}{1 + 0.468 \times 10^{-2} T_a + 0.229 \times 10^{-3} T_a^2 + 0.747 \times 10^{-6} T_a^3 + 0.173 \times 10^{-8} T_a^4} \right)^{2/3}. \tag{3.6}$$

Equation (3.6) is plotted as a function of ambient temperature in Fig. 3.2(b). Clearly, m_{dh}^{*} varies more with temperature than m_{de}^{*}. However, in terms of the overall picture, the temperature dependence displayed by m_{de}^{*} is rather benign.

3.1.3 Carrier Concentration: Intrinsic Si

The intrinsic carrier concentration is given by the well-known formula [52, 55]

$$n_{io} = \sqrt{N_C N_V} e^{-E_g/2kT_a},$$

[1] In many texts, such as [52] and [55], the effective mass contribution from the split-off valence band is neglected.

where

$$N_C \equiv 2\left(\frac{2\pi m^*_{de} kT_a}{h^2}\right)^{3/2} \tag{3.7}$$

and

$$N_V \equiv 2\left(\frac{2\pi m^*_{dh} kT_a}{h^2}\right)^{3/2}. \tag{3.8}$$

Thus,

$$n_{io} \approx 4.82e15 T_a^{3/2} \sqrt{6\frac{m^*_t}{m_0}\sqrt{\frac{m^*_l}{m_0}}\left[\left(\frac{m^*_{lh}}{m_0}\right)^{3/2} + \left(\frac{m^*_{hh}}{m_0}\right)^{3/2} + \left(\frac{m^*_{soh}}{m_0}\right)^{3/2}\right]} e^{-E_g/2kT_a}. \tag{3.9}$$

The intrinsic carrier concentration is exponentially proportional to temperature via the energy bandgap and linearly related to $T_a^{3/2}$ and the effective masses. As the effective masses were shown to be only weakly temperature dependent[2], the conclusion is that n_{io} goes as roughly $T_a^{3/2} \exp\{-E_g(T_a)/2kT_a\}$.

The intrinsic carrier concentration appears as a function of temperature in Fig. 3.3. The temperature dependence of n_{io} is quite remarkable; with cooling from 300K to 10K, the intrinsic carrier concentration changes by over 300 orders of magnitude. Furthermore, it can be seen that there is a knee around 50K below which n_{io} drops off rapidly. These trends have been confirmed experimentally down to 77K by Sproul and Green [75]. However little data exists at lower temperatures, where the theoretical intrinsic carrier concentration drops to very low levels. Nonetheless, we make the assumption that the theory is correct[3]. Thus, the intrinsic carrier concentration is expected to play a very important role in determining device performance at cryogenic temperatures.

3.1.4 Carrier Mobility

Carrier mobility is an important consideration as it determines the effective velocity of carriers under the influence of electric fields. Mobility is directly related to scattering mechanisms, which include scattering due to lattice vibrations or phonons (μ_{ps}), ionized impurities (μ_{ii}), velocity saturation (μ_{vs}), carrier-to-carrier collisions (μ_{cc}), and neutral impurities (μ_{ni}). Of these mechanisms, μ_{vs}, μ_{cc}, and μ_{ni} are independent whereas μ_{ps} and μ_{ii} are coupled to each other [73]. Thus, the

[2] i.e., the magnitude of the proportionality.

[3] It is possible that the experimental values of n_{io} would saturate at very low temperatures due to traps and other defects not included in the theoretical calculation. However, these effects should not exist in pure silicon.

total electron or hole mobility can be obtained by considering each of the effects in parallel as

$$\mu = \left(\frac{1}{\mu_{psii}} + \frac{1}{\mu_{vs}} + \frac{1}{\mu_{cc}} + \frac{1}{\mu_{ni}} \right)^{-1}, \tag{3.10}$$

where μ_{psii} is the mobility considering only phonon and ionized impurity scattering. As evidenced by equation (3.10), the mobility can be degraded by any of the scattering contributions. Therefore, it is important to understand the temperature dependence of each component.

3.1.4.1 Phonon Scattering

Phonon scattering refers to scattering due to lattice vibrations. As it turns out, the energy band gap is a function of lattice constant and is modulated by lattice vibrations[4]. This modulation of the bandgap causes scattering of free carriers. An explicit expression for the mobility due to phonon scattering has been given by Bardeen and Shockley as [76]

$$\mu_{ps} = \left[\frac{\sqrt{8\pi} h^4 c_{ii}}{3 E_1^2 m_0^{5/2} k^{3/2}} \right] T_a^{-3/2}, \tag{3.11}$$

where m_0 is the free electron mass, c_{ii} is the longitudinal elastic constant of the material, and E_1 is a parameter which has units of eV and links the vibrations to the perturbations of the energy bandgap. While equation (3.11) provides physical insight into phonon scattering[5], it depends not only on empirically determined parameters c_{ii} and E_1, but also on the effective mass, which was shown to be a function of temperature in section 3.1.2. Therefore, it is desirable to use a model of μ_{ps} that is a more clearcut function of temperature, such as that provided by Gutiérrez in [73]:

$$\mu_{ps} = \left[\frac{1}{\mu_{0a} \left(T_a / 300 \right)^{-\kappa_a}} + \frac{1}{\mu_{0b} \left(T_a / 300 \right)^{-\kappa_b}} \right]^{-1}. \tag{3.12}$$

The parameters μ_{0a}, μ_{0b}, κ_a, and κ_b required for equation (3.12) are experimentally determined coefficients and are given in Table 3.1.

[4] Phonons cause the conduction and valence bands to move in opposite directions, thus changing the bandgap, whereas electrostatic potentials cause the conduction and valence bands to move in the same direction, thereby preserving the bandgap [76].

[5] e.g., phonon scattering gets worse with increasing effective mass and scales with temperature as $T_a^{-3/2}$.

Table 3.1: Coefficients for silicon μ_{ps} calculations using equation (3.12) [73]

	μ_{0a} $\mathrm{cm}^2/\mathrm{Vs}$	μ_{0b} $\mathrm{cm}^2/\mathrm{Vs}$	κ_a -	κ_b -
Electrons	4195	2153	1.5	3.13
Holes	2502	591	1.5	3.25

Table 3.2: Coefficients for silicon μ_{psii} calculations using equation (3.14) [73]

	μ_{min} $\mathrm{cm}^2/\mathrm{Vs}$	N_{ref} cm^{-3}	κ_c -
Electrons	$197.17 - 45.505 \times \log 10\,(T_a)$	$1.12 \times 10^{17}\,(T_a/300)^{3.2}$	$0.72 \times (T_a/300)^{0.065}$
Holes	$110.90 - 25.597 \times \log 10\,(T_a)$	$2.23 \times 10^{17}\,(T_a/300)^{3.2}$	$0.72 \times (T_a/300)^{0.065}$

3.1.4.2 Ionized Impurity Scattering

The second form of scattering is that from ionized impurities (i.e., dopants) and is described by [55]

$$\mu_{ii} = \frac{64\sqrt{\pi}\epsilon_s^2\,(2kT_a)^{3/2}}{N_I q^3 m^{*1/2}} \left\{ \ln\left[1 + \left(\frac{12\pi\epsilon_s kT_a}{q^2 N_I^{1/3}}\right)^2\right]\right\}^{-1}, \tag{3.13}$$

where N_I is the concentration of ionized impurities, m^* is the conductivity effective mass, and ϵ_s is the dielectric constant of the material. However, phonon scattering and ionized impurity scattering are not fully independent, so it is not possible to simply combine equations (3.12) and (3.13) in a parallel fashion. To overcome this difficulty, an alternative model was suggested by Caughey and Thomas, who noticed that the general shape of mobility as a function of dopant concentration resembled the Fermi-Dirac function [77]. The model is given as

$$\mu_{psii} = \mu_{min} + \frac{\mu_{ps} - \mu_{min}}{1 + (N_I/N_{ref})^{\kappa_c}}, \tag{3.14}$$

where μ_{ps} is calculated using equation (3.12) and N_I is the concentration of ionized impurities in the material and is calculated using equations (3.17) and (3.18). In order to account for temperature in equation (3.14), Gutíerrez performed a fit of the coefficients as a function of temperature. The resulting coefficients appear in Table 3.2 [73].

3.1.4.3 Neutral Impurity Scattering

In addition to scattering from the ionized impurities, carriers will also undergo scattering from neutral impurities. This effect becomes important when a material contains a large number of non-ionized impurities. The mobility due to neutral donor impurities has been modeled by Li and Thurber [78] as

$$\mu_{ni} = \frac{2\pi^3 q^3 m_{ce}^*}{5 N_N \epsilon_s h^3} \times 10^{-2} \left[\frac{2}{3} \left(\frac{kT_a}{E_N} \right)^{1/2} + \frac{1}{3} \left(\frac{E_N}{kT_a} \right)^{1/2} \right], \tag{3.15}$$

where N_N is the density of non-ionized donor atoms, $E_N = 1.136 \times 10^{-19} \left(m_{ce}^*/m_0\right) \left(\epsilon_0/\epsilon_S\right)^2$, and $m_{ce} = 1/3 \left(1/m_t^* + 2/m_l^*\right)$ is the effective conduction mass [73]. Unfortunately, a relationship for neutral impurity scattering in p-type silicon has not been identified, so this effect will be ignored for mobility in p-type silicon.

3.1.4.4 Carrier-to-Carrier Scattering

At high injection levels, the carrier concentration can be high enough that the mobility is compromised due to inter-carrier scattering. This effect has been studied by Dorkel and Leturcq [79] and can be modeled as

$$\mu_{cc} = 2 \times 10^{17} \frac{T_a^{3/2}}{\sqrt{np} \ln \left\{ 1 + 8.28 \times 10^8 T_a^2 \left(np\right)^{-1/3} \right\}}, \tag{3.16}$$

where n and p are the levels of electrons and holes in the silicon. As this is a high-injection effect, it is not of great importance to the work reported here. Therefore, this effect will be ignored in mobility calculations.

3.1.4.5 Overall Dependence of Mobility on Ambient Temperature

For the overall analysis, it is assumed that the dominant sources of mobility degradation are phonon and ionized impurity scattering. This assumption is valid for the case in which the majority of the impurities have been ionized and the operating currents are well below high-injection. The mobility due to phonon and ionized impurity scattering appears in Fig. 3.4. For low ionized dopant concentrations, the both the electron and hole mobilities increase by several orders of magnitude as the ambient temperature is reduced from 300 to 18 K. However, the increase in mobilities become much less important as the ionized dopant concentrations reach levels on the order of 10^{18} or higher. Thus, the changes in mobility with cooling would not be expected to have a large impact on cryogenic

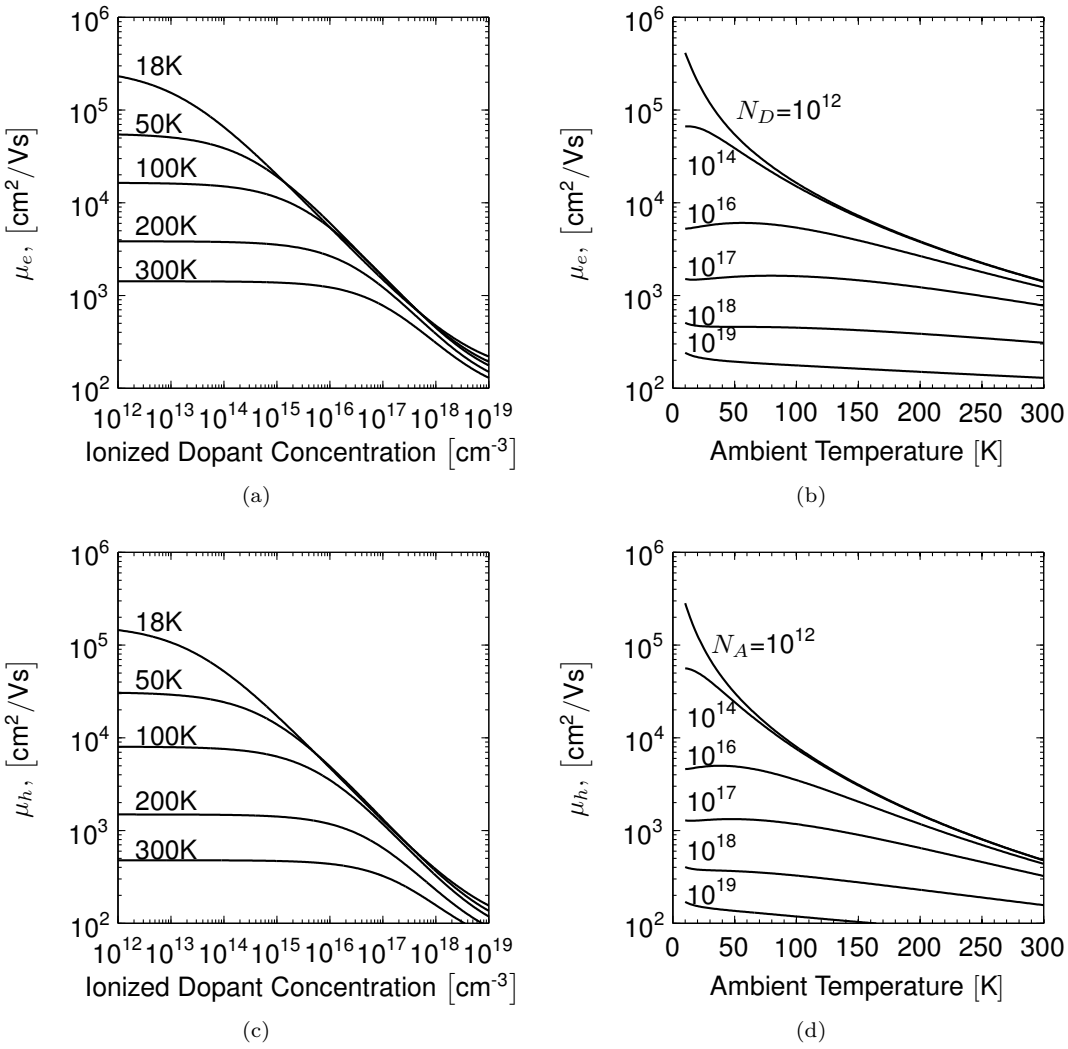


Figure 3.4: (a). Computed electron mobility and (c) hole mobility as a function of temperature. High field effects have been neglected. For doping levels above $\sim 10^{17}$, the mobilities are not strongly dependent on temperature.

device performance.

3.1.5 Carrier Freeze-out

In order for a material to conduct current, free carriers must be present in the form of holes in the valence band or electrons in the conduction band. It is well known that the conductivity of metals increases greatly with cooling due to an increase in carrier mobility. However in some semiconducting materials, the population of free carriers diminishes as the ambient temperature is lowered. This phenomena is known as *carrier freeze-out* and is of interest as the properties of

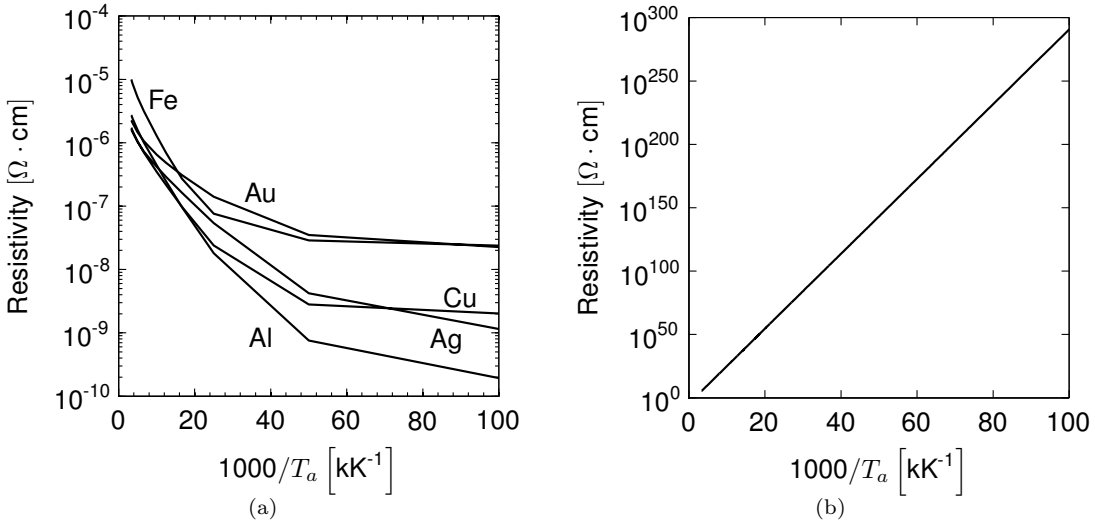


Figure 3.5: (a) Resistivity of various metals as a function of temperature [80]. (b) The resistivity of intrinsic silicon as a function of temperature

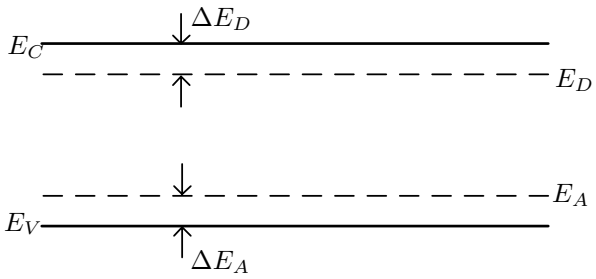


Figure 3.6: Band structure for a doped semiconductor. In order to excite carriers from the donor and acceptor states, an ionization energy of ΔE_D and ΔE_A is required.

semiconductor devices are tightly coupled to the free carrier concentration in the materials. To demonstrate the problem, in Fig. 3.5 the conductivity of several metals is plotted alongside that of intrinsic Si, both as a function of temperature. The impact of carrier freeze-out in intrinsic Si is quite clear; as the number of free carriers exponentially decreases as the ratio of the bandgap to the ambient temperature increases, intrinsic silicon becomes an excellent insulator at cryogenic temperatures.

As it turns out, the issue of carrier freeze-out is eliminated if the semiconductor is doped at a level above what is known as the Mott-transition. To understand why heavy doping is beneficial, it is instructive to examine a band diagram of a lightly doped[6] piece of silicon as shown in Fig. 3.6. Doped silicon differs from intrinsic silicon in that a much smaller thermal excitation is required to

[6] i.e., below 10^{17}cm$^{-}3$.

Table 3.3: Mott-transition (N_{crit}) and low doping ionization energy for various silicon alloys.

Material	Si:Sb	Si:P	Si:As	Si:Bi	Si:B
Type	D	D	D	D	A
$E_{A(D)}$ [meV]	42.7	45.5	53.7	71.0	44.4
N_{crit} [cm^{-3}]	3.0×10^{18}	3.7×10^{18}	8.5×10^{18}	1.8×10^{19}	4.1×10^{18}

generate carriers, as typical dopant ionization energies are about a factor of twenty lower than the bandgap of Si [81, 82, 83]. Thus, we would expect a lightly doped piece of silicon to conduct at lower temperatures than an undoped piece of silicon. However, at absolute zero temperature, we would still expect a lightly doped piece of silicon to be a perfect insulator as there is no thermal excitation to free the carriers that are needed to conduct current from the dopant atoms. However, the energy required to ionize the dopants depends upon the spacing of the impurity centers and when the dopants are close enough that their wave-functions overlap[7], impurity bands will form and no ionization energy is required to conduct current. The dopant level at which the impurity bands form is known as the *Mott-transition*, and above this level no carrier freeze-out effects will occur[8] [84]. The Mott-transition for various silicon alloys is given in Table 3.3 [85]. It is apparent the Mott-transition in B, As, and P doped silicon is high enough that there is no guarantee that the devices analyzed in this work will be doped heavily enough to completely avoid freeze-out effects. Thus, a model is needed to calculate the fraction of ionized dopants as a function of temperature and dopant concentration.

A detailed treatment of carrier freeze-out has recently been presented by Altermatt, Schenk, and Heiser with the end result being an accurate numerical model of carrier freeze-out effects [85, 86]. This model will be used here as it is accurate down to well cryogenic temperatures. Using this model, the ionized acceptor and donor concentrations can be computed as

$$\frac{N_D^+}{N_D} = \frac{1}{2}\left[1 - b - g\frac{n_1}{N_D} + \sqrt{\left(g\frac{n_1}{N_D} + b - 1\right)^2 + 4g\frac{n_1}{N_D}}\right] \tag{3.17}$$

and

$$\frac{N_A^-}{N_A} = \frac{1}{2}\left[1 - b - g\frac{p_1}{N_A} + \sqrt{\left(g\frac{p_1}{N_A} + b - 1\right)^2 + 4g\frac{p_1}{N_A}}\right]. \tag{3.18}$$

The equations and coefficients required to evaluate equations (3.17) and (3.18) can be found in Table 3.4 and plots of the ionization ratio as a function of temperature and dopant concentration for

[7] i.e., there is a finite probability that two bound carriers can both be in the same location.

[8] This is true for a doped material in which the dopants are equally spaced within the lattice. For random dopant spacing, the formation of impurity bands is gradual.

Table 3.4: Coefficients and equations required to evaluate the ionization ratio [85]

Parameter		Si:P	Si:As	Si:B
$\Delta E_{D0(A0)}$	(meV)	45.5	53.7	44.39
N_{ref}	(cm^{-3})	2.2×10^{18}	3×10^{18}	1.3×10^{18}
c		2	1.5	1.4
N_b	(cm^{-3})	6×10^{18}	9×10^{18}	4.5×10^{18}
d		2.3	1.8	2.4
g		1/2	1/2	1/4
Equation				
$\Delta E_{D(A)}$	$E_{D0(A0)} / \left(1 + N_{D(A)}/N_{ref}\right)^c$			
b	$1 / \left(1 + N_{D(A)}/N_{ref}\right)^d$			
n_1	$N_C \exp\{-\Delta E_D / kT_a\}$			
p_1	$N_V \exp\{-\Delta E_A / kT_a\}$			

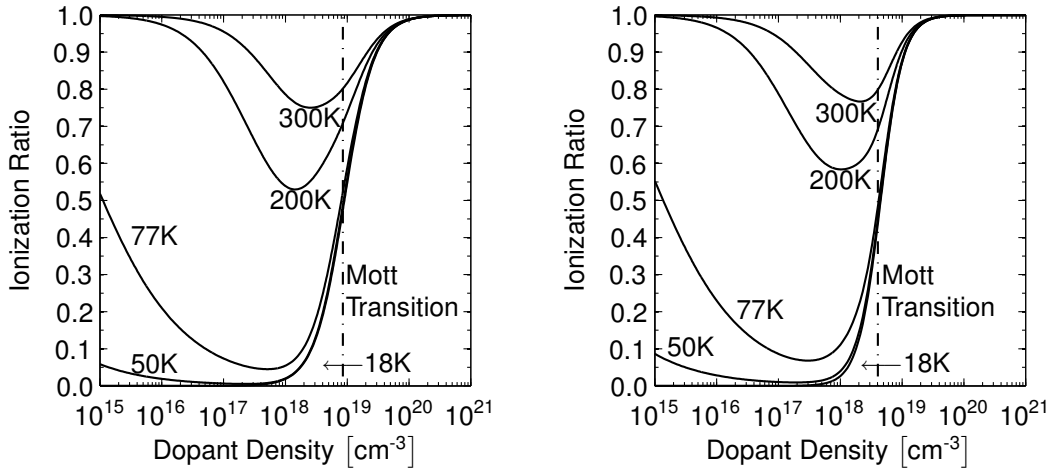


Figure 3.7: Fraction of dopants ionized as a function of dopant concentration and temperature for (a) arsenic donor impurities and (b) boron acceptor impurities. The doping threshold for the Mott transition is also indicated.

Si:As and Si:B are shown in Fig. 3.7. Carrier freeze-out effects are clearly important at cryogenic temperatures for devices in which the dopant concentration is on the order of the Mott-transition. Thus, the effects of carrier freeze-out will be considered in Chapter 6, when the temperature dependence of the resistances associated with doped semiconductor regions is discussed.

3.2 Properties of SiGe Transistors at Cryogenic Temperatures

3.2.1 Doping-Induced Apparent Bandgap Narrowing

As discussed in Section 2.1.3, the collector current in a SiGe HBT depends upon the apparent bandgap narrowing which can be written[9]

$$\Delta E_{g,app} \approx 18.7 \times 10^{-3} \ln\left\{\frac{N_{AB}^{-}}{N_{DE}^{+}}\right\}. \tag{3.19}$$

and occurs due to heavy doping in the base and emitter. As ΔE_g^{app} depends on the natural logarithm of the ratio of ionized impurities in the base to that in the emitter, the net effect is that the bandgap actually looks wider in terms of the barrier that the electrons being injected into the base must overcome. Furthermore, since this is a bandgap effect, it is exponentially enhanced with decreasing temperature. As it turns out, this, as opposed to carrier freeze-out, is actually the limiting factor in the operation of modern silicon BJTs at cryogenic temperatures [30]. Fortunately, the bandgap narrowing that the Ge induces is sufficient to compensate out the apparent bandgap narrowing. Nevertheless, it offsets the effect of the Ge and, in the case of a triangular Ge profile (i.e., no Ge content on the base side of the base–emitter depletion region), it is a limiting effect.

3.2.2 DC Terminal Currents

3.2.2.1 Base Current

In Section 2.1.3, expressions for the dc currents in a SiGe HBT were presented. These equations depend upon the terminal currents in an identically doped silicon device as well as the bandgap narrowing effects due to the Ge content in the base. With the tools developed in Section 3.1, it is now possible to predict the changes that occur in the base current saturation coefficient, J_{B0}, with cooling. From equation 2.12, the base current can be written as

$$J_B(T_a) \approx \frac{kT_a\mu_{pe}(T_a)\, n_{io}^2(T_a)}{L_{PE}(T_a)\, N_{DE}^{+}(T_a)} e^{qV_{BE}/kT_a} = J_{B0}(T_a)\, e^{qV_{BE}/V_T}. \tag{3.20}$$

Of particular interest is the change in base saturation current, J_{B0}, with cooling. From equation 3.20 and the information discussed in Section 3.1.3, the fractional change in base saturation current with

[9] Equation (3.19) is only valid for the case in which $N_{AB}^{-} > 7 \times 10^{17}$. For the case in which this is not true, equation (2.16) must be used instead.

cooling from 300 K can be written as

$$\frac{J_{B0}\left(T_a\right)}{J_{B0}\left(300\right)} = \left(\frac{T_a}{300}\right)^4 \left(\frac{\mu_{pe}\left(T_a\right)}{\mu_{pe}\left(300\right)}\right) \left(\frac{L_{PE}\left(300\right) N_{DE}^{+}\left(300\right)}{L_{PE}\left(T_a\right) N_{DE}^{+}\left(T_a\right)}\right) \frac{e^{-E_g(T_a)/kT_a}}{e^{-E_g(300)/(k300)}}. \tag{3.21}$$

In order to arrive at an explicit temperature dependence for J_{B0}, it is necessary to write the temperature dependences of μ_{pe}, L_{PE}, N_{DE}^{+}, and E_g. This can be done as follows:

μ_{pe}: Referring back to Fig. 3.4, it can be seen that the sensitivity of carrier mobility to temperature is strongly dependent on the dopant concentration, with higher dopant concentrations resulting in a reduced sensitivity to temperature. Furthermore, the dopant level in the intrinsic mono- or poly-emitter is on the order of 10^{19}cm^{-3} or higher at the metallurgical base–emitter junction [59, 87]. Thus, μ_{pe} can be assumed to be only weakly temperature dependent[10].

N_{DE}^{+}: As discussed in Section 3.1.5, carrier freeze-out is not an issue in heavily doped polysilicon materials such as the poly-emitter. Therefore, N_{DE}^{+} is nearly independent of ambient temperature.

L_{PE}: The fundamental definition of hole diffusion length is [55]: $L_P \equiv \sqrt{D_P \tau_p} = \sqrt{\mu_p \tau_p k T_a / q}$, where τ_p, the hole minority carrier lifetime, is the average lifetime for a hole diffusing into the n-doped region. The carrier lifetime in heavily doped materials is limited by band to band recombination with a lifetime equal to $1/C_A {N_D^{+}}^2$, where C_A known as the Auger coefficient [88]. Dziewior and Schmid have studied the temperature dependence of the Auger coefficient for holes in heavily doped silicon and found that there is only a very weak correlation between C_A and the ambient temperature [89]. Therefore, since N_{DE}^{+} is nearly independent of temperature and μ_{pe} is only weakly dependent on temperature, a reasonable approximation is that $L_{PE}\left(T_a\right) \propto \sqrt{T_a}$.

E_g: The temperature dependence of the silicon bandgap, E_g, was discussed in Section 3.1.1. Explicitly, the bandgap can be written as a function of temperature as $E_g\left(T_a\right) \approx 1.17 + 5.65 \times 10^{-6} T_a - 5.11 \times 10^{-7} T_a^2 - 8.03 \times 10^{-10} T_a^3 + 2.50 \times 10^{-12} T_a^4$.

Thus, equation (3.21) can be rewritten as:

$$\frac{J_{B0}\left(T_a\right)}{J_{B0}\left(300\right)} = 7.83 \times 10^{18} \left(\frac{T_a}{300}\right)^{7/2} e^{-E_g(T_a)/kT_a}. \tag{3.22}$$

[10] In Section 3.1.4, the majority carrier mobilities were discussed. However, here we are referring to the minority carrier mobilities. This may result in a small error, but does not change the result significantly.

From this analysis, it is apparent that J_{B0} should decrease significantly at cryogenic temperatures. Furthermore, the decrease should depend only upon the energy band gap of silicon, which is independent of device technology.

3.2.2.2 Collector Current

The same procedure can be repeated for the collector current saturation coefficient, J_{C0}, in order to determine its temperature dependency. The value of $J_{C0}(T_a)$ normalized to $J_{C0}(T_{RT})$ is given as

$$\frac{J_{C0}(T_a)}{J_{C0}(300)} \approx \left(\frac{T_a}{300\ K}\right)^3 \frac{\widetilde{\gamma}(T_a)}{\widetilde{\gamma}(300)} \frac{\widetilde{\eta}(T_a)}{\widetilde{\eta}(300)} \frac{\mu_{nb,Si}(T_a)}{\mu_{nb,Si}(300)} \frac{N_{AB}^{-}(300)}{N_{AB}^{-}(T_a)} \frac{e^{E_g(300)/k300}}{e^{E_g(T_a)/kT_a}} \frac{e^{\Delta E_{g,app}/kT_a}}{e^{\Delta E_{g,app}/k300}} \frac{e^{\Delta E_{g,Ge}(0)/kT_a}}{e^{\Delta E_{g,Ge}(0)/k300}}. \tag{3.23}$$

Since $\widetilde{\gamma} = \mu_{nb,SiGe}/\mu_{nb,Si}$, its value is assumed to be temperature independent. Furthermore, based on the analysis given by Sokolić and Amon, $\widetilde{\eta} = (N_C N_V)_{SiGe} / (N_C N_V)_{Si}$ is assumed to be only weakly temperature dependent [90]. Thus, it will be considered constant as a function of temperature. As discussed above, the mobilities are not strongly dependent upon temperature as the doping level in the base is quite high. Thus, assuming the base is doped well above the Mott transition, equation (3.23) can be simplified to

$$\frac{J_{C0}(T_a)}{J_{C0}(300)} \approx 7.83 \times 10^{18} \left(\frac{T_a}{300}\right)^3 e^{-E_g(T_a)/kT_a} \frac{e^{(\Delta E_{g,app} + \Delta E_{g,Ge}(0))/kT_a}}{e^{(\Delta E_{g,app} + \Delta E_{g,Ge}(0))/k300}}. \tag{3.24}$$

Referring to equation (3.24), it is apparent that detailed information regarding the temperature dependence of both $\Delta E_{g,app}$ and $\Delta E_{g,Ge}(0)$ is required in order to predict the temperature dependence of the collector saturation current. To complicate matters, both the exact location of the base edge of the base–emitter space charge region and the magnitude of apparent bandgap narrowing are dependent upon the doping profile in the base and emitter, thus making it impossible to determine these quantities without detailed SIMS profiles. Thus, an ambiguity exists and it is assumed that information about the collector saturation current cannot be used to determine the Ge content or the doping levels. Due to these considerations, the temperature behavior of J_{C0} is expected to vary greatly among the different SiGe devices.

3.2.2.3 DC Current Gain

As discussed in Chapter 2, β_{DC} is critical in determining the low-GHz range noise performance of modern SiGe bipolar transistors. The temperature dependence of β_{DC} can be evaluated by taking the ratio of equation (3.24) to equation (3.22):

$$\frac{\beta_{DC}\left(T_a\right)}{\beta_{DC}\left(300\right)} \approx \sqrt{\frac{300}{T_a}}\, \frac{e^{\Delta E_{g,app}+\Delta E_{g,Ge}(0)/kT_a}}{e^{\Delta E_{g,app}+\Delta E_{g,Ge}(0)/k300}}. \tag{3.25}$$

Therefore, the dc current gain is exponentially enhanced by cooling provided that $\Delta E_{g,app} + \Delta E_{g,Ge}\left(0\right) > 0$. Furthermore, as the enhancement is related to the Ge concentration at the edge of the base–emitter space charge region, this effect is expected to vary greatly among various SiGe HBT technology platforms.

3.2.2.4 Non-Equilibrium Transport

In deriving the transport equations for bipolar transistors, it is usually assumed that the carriers are at thermal equilibrium. Under this assumption, the Boltzmann approximation applies and standard drift-diffusion theory can be used. However, researchers have observed that the collector current ideality factor increases at cryogenic temperatures. Through the use of Monte-Carlo simulation tools, the non-ideal collector current slope has been traced down to quasi-ballistic transport in the base, which can be modeled as an increase in the effective temperature of the electrons [29]. In terms of measured iv curves, non-equilibrium transport will appear as a degradation in the slope of the collector current under low-injection conditions. Furthermore, this effect will limit the increase in transconductance that can be obtained with cooling.

3.3 Summary

In this chapter, the cryogenic properties of silicon materials and SiGe devices have been explored. In the first part of the chapter, it was shown that significant changes occur in the bandgap and intrinsic carrier concentration of relatively pure silicon materials as well as in the carrier mobility and ionized carrier concentration in lightly doped Si samples. In the second part of the chapter, the expected effects of cooling on the dc currents of SiGe HBTs were discussed. It was seen that the dc current gain is expected to grow exponentially with cooling, but that this enhancement effect is expected to vary significantly among different processes. This concludes the theoretical section of the thesis. In the next part of the dissertation, the cryogenic properties of a variety of SiGe devices

will be evaluated experimentally.

Part II

Modeling

Chapter 4

Introduction to Modeling and Description of Devices

In Part I, theoretical information regarding the operation of SiGe HBTs at cryogenic temperatures was presented. In this part of the thesis, the dc, RF, and noise modeling of a wide variety of SiGe HBTs at cryogenic temperatures is investigated. In this chapter, an introduction to the characterized devices and the measurement techniques will be given.

4.1 Summary of Device Technologies

In order to conduct a general investigation of the cryogenic operating characteristics of SiGe HBTs, several different devices have been studied. In this section, a description of the various transistors that have been evaluated will be given. A schematic cross section of a typical SiGe HBT appears in Fig. 4.1. As they are vertical structures, the fabrication of SiGe HBTs requires both the implantation of materials into the silicon substrate as well as the epitaxial growth of features on the surface of the wafer. Furthermore, the fabrication of SiGe HBTs has to be done in such a way that it is compatible with a standard CMOS process[1]. In order to achieve the high level of performance seen in today's state-of-the-art SiGe devices, it is critical that each of these steps be highly optimized. As it turns out, this optimization process has lead each foundry to develop a slightly different recipe and these differences will be highlighted in the discussion that follows. In particular, the following approaches have been taken to critical processing steps:

Collector module The n+ sub-collector module can be created either by ion-implantation, or by first selectively doping n+ regions on the surface of the silicon wafer and then epitaxially growing a Si layer to create buried sub-collector regions. While the former approach is less

[1] i.e., the devices must be able to withstand being heated to temperatures in the range of 1000° C

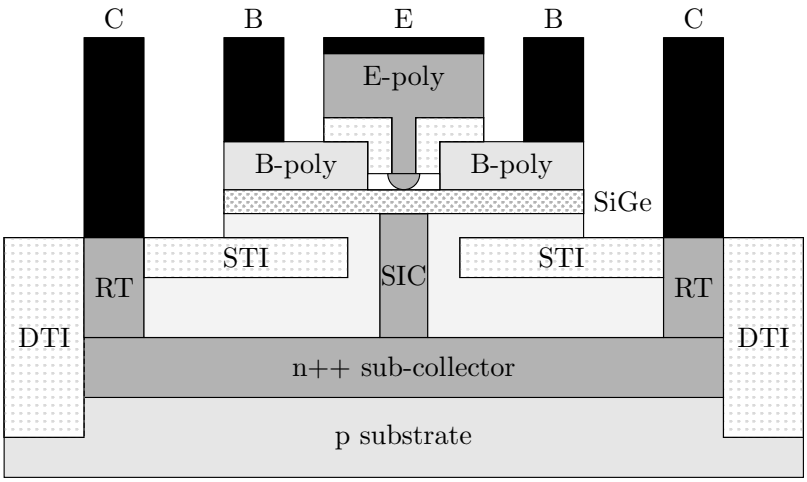


Figure 4.1: Typical state-of-the-art HBT structure

costly, the latter approach results in lower sheet resistance [91]. Thus, buried layer (BL) sub-collectors are superior to implanted sub-collectors.

Extrinsic base There are several flavors of extrinsic base commonly found in SiGe HBTs. First of all, a distinction can be drawn between a *raised extrinsic base* and an *implanted extrinsic base*. While a raised extrinsic base requires extra processing steps in comparison to an implanted extrinsic base, the extra effort is rewarded by a improvement in performance [91]. This is because the raised extrinsic base is self-aligned to the emitter, which allows for independent optimization of the base resistance and the base–collector capacitance [92]. The second way in which processes vary with respect to the extrinsic base is in the actual patterning of the extrinsic base. In general, the E-B junction of SiGe HBTs is defined in either a *fully-self-aligned* (FSA) or a *quasi-self-aligned* (QSA) manner. In the case of a FSA with non-selective epitaxy, a sacrificial emitter is used[2] in order to define the extrinsic base, allowing the extrinsic base to butt up against the edge of the active base–emitter region, resulting in a low extrinsic base contribution to the overall base resistance. On the other hand, a QSA extrinsic base is not aligned to the E-B junction but by the alignment of two different masks and, as a consequence, it is not possible to get the extrinsic base right up against the edge of the active E-B junction [60]. Therefore, a QSA base will have a larger extrinsic base resistance than a FSA extrinsic base.

Carbon doping of intrinsic base During the deposition of the SiGe base, a narrow B spike with

[2] An alternative approach to the fabrication of a FSA extrinsic base is the use selective base epitaxy [87].

Table 4.1: A summary of the devices investigated. Transistor metrics are those reported by the foundries [60, 87, 95, 96, 97, 98, 99]

Label	Foundry	Process name	Tech. node μm	f_t/f_{max} GHz	BV_{CE0}/BV_{CB0} V	Eff. device area μm^2
IBM-G4	IBM	BiCMOS8HP	0.12	200/280	1.7/5.7	0.12×18
IHP-G4	IHP	SG13	0.13	255/315	1.8/5.6	$0.17\times16\times2.09$
ST-G4	ST	BiCMOS9MW	0.13	230/280	1.6/5.8	$0.13\times9.86\mu m^2$
ST-X2	ST	BipX2	0.15	233/249	1.50/5.45	$0.15\times2\times14.77$
ST-X1	ST	BipX1	0.17	263/290	1.44/5.44	$0.17\times2\times14.79$
ST-X3	ST	BipX3	0.17	263/290	1.44/5.44	$0.17\times2\times14.79$
JAZZ-G3	JAZZ	SBC18	0.18	150/190	2.2/7.0	$0.18\times6\times60$
NXP-G3	NXP	Qubic4Xi	0.25	216/177	1.44/5.20	0.3×1

Foundry	IBM	IHP	ST	ST	ST	ST	JAZZ	NXP
Process	BiCMOS8HP	SG13	BiCMOS9MW	BiPX2	BiPX1	BiPX3	SBC18	Qubic4Xi
Date rec'd	07/2008	11/2008	09/2008	01/2008	01/2008	09/2008	04/2007	10/2008
Wafer id	A3CRQNP	-	J748CBZ25	J535RYV24	J514VBF03	J514VBF13	K58386-9	-

concentration on the order of 10^{19} cm^{-3} is typically introduced in order to dope the SiGe base. However, later in the processing a rapid thermal annealing (RTA) process is carried out in order to activate the polysilicon source, drain, gate, base, and emitter regions. Unfortunately, the RTA process causes the dopants in the base to diffuse, putting a limit on the minimum base thickness. One approach to overcoming this limitation without adversely affecting the structural or electrical properties of the SiGe layer is to lightly dope the base with carbon, which is effective in blocking the diffusion of boron during the RTA stage [93, 94].

Ge Profile As discussed at length in Chapters 2 and 3, the Ge profile across the base has serious consequences in terms of device performance. First of all, β_{DC} is determined largely by the Ge content at the base edge of the base–emitter space-charge region. Secondly, the slope of Ge across the base influences the base transit time, thereby playing a role in the overall speed of the device. Finally, the Ge slope near the base edge of the base–emitter space-charge region determines the extent to which the reverse-Early effect[3] will influence device performance. Common profiles include triangular (0% Ge on the emitter side of the base), trapezoidal, and box (constant Ge) profiles.

A summary of the devices compared in the following chapters appears in Table 4.1 and information about the processes from which each of the devices originate is given below.

[3] See [30], pg. 186.

4.1.1 IBM BiCMOS8HP

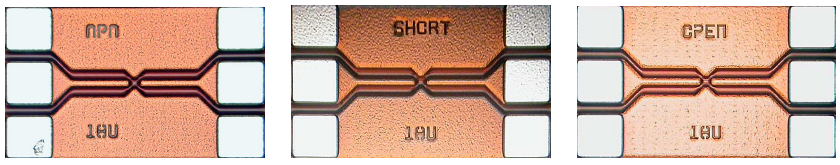
IBM's fourth generation BiCMOS technology, BiCMOS8HP. was fully commercialized in 2005, and is to date IBM's most advanced BiCMOS offering [100]. The process features 0.12μm HBTs with a low-resistance BL sub-collector, a FSA raised extrinsic base, and a B doped SiGe base with carbon doping [95]. Unique features of this process include an in situ doped emitter with phosphorus impurities[4] [102]. Although the Ge concentration is reported to be 25% in [103], elsewhere it is reported that IBM uses a trapezoidal profile [104, 105]. Therefore, it is assumed that the Ge profile is trapezoidal with an average Ge content of 25%. Finally, the transistors have a room temperature f_t/f_{max} of 200/280 GHz and BV_{CE0}/BV_{CB0} of 1.7/5.7 V [95].

Access to the IBM process was available through the Trusted Foundry program and five reticles were taped-out between 2006 and 2008. In addition to integrated circuits, discrete devices were put on each of these reticles along with the de-embedding structures required to extract device models. A die photo of an HBT test structure along with the open/short de-embedding structures appears in Fig. 4.2.

4.1.2 ST Microelectronics

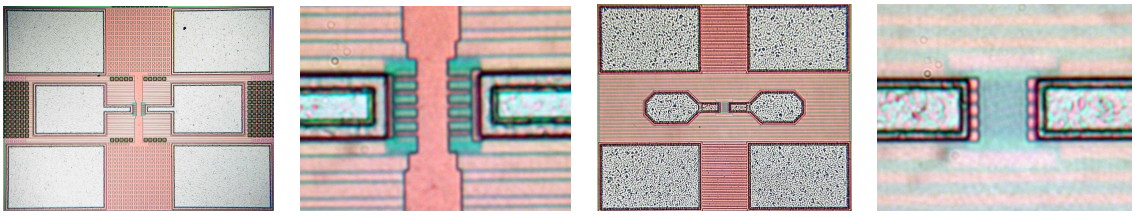
In this work, ST Microelectronics devices from four different process lines were characterized. All of the devices come from a process featuring BL sub-collectors, FSA raised extrinsic bases[5], B doped SiGe bases with C doping, and in situ doped As mono-emitters[6]. The Ge profile for all devices is known to be trapezoidal with the Ge content varying from 20–30% [60, 87]. Of the four devices, three come from experimental processes, BiPX1, BiPX2, and BiPX3. The fourth device comes from the ST BiCMOS9MW process, which is the most advanced BiCMOS technology platform currently available from ST. The experimental devices were fabricated in the research stages leading to the

[4] Arsenic is typical used as the emitter dopant as it performs better during the required RTA stage [101].
[5] The extrinsic bases are aligned using selective epitaxy as opposed to a sacrificial emitter.
[6] The intrinsic emitters are mono-crystalline whereas the extrinsic emitters are polycrystalline



(a) (b) (c)

Figure 4.2: IBM BiCMOS8HP device photographs: (a) transistor, (b) short-circuit, and (c) open-circuit test structures



(a) (b) (c) (d)

Figure 4.3: ST device photos. (a) BiPX1 test structure, (b) close-up of BiPX1 device, (c) BiCMOS9MW test structure, (d) and close-up of BiCMOS9MW device

development of the BiCMOS9MW process.

The BiPX1 and BiPX3 processes are nearly identical and feature devices with 0.17μm emitter widths, room temperature f_t/f_{max} values of 263/290 GHz, and room temperature BV_{CEO}/BV_{CBO} values of 1.44/5.44V. Finally, the effective area of the devices measured from the BiPX1 and BiPX3 processes are $0.17 \times 2 \times 14.79\mu\text{m}^2$.

The device from the BiPX2 process line features a 0.15μm emitter width, room temperature f_t/f_{max} values of 233/249 GHz, and room temperature BV_{CEO}/BV_{CBO} values of 1.50/5.45V. The effective area of the measured BiPX2 device is $0.15 \times 2 \times 14.77\mu\text{m}^2$.

Finally, the device from the BiCMOS9MW process line features a 0.13μm emitter width, room temperature f_t/f_{max} values of 230/280 GHz, and room temperature BV_{CEO}/BV_{CBO} values of 1.6/5.8V. The effective area of the measured ST BiCMOS9MW device is $0.13 \times 9.86\mu\text{m}^2$.

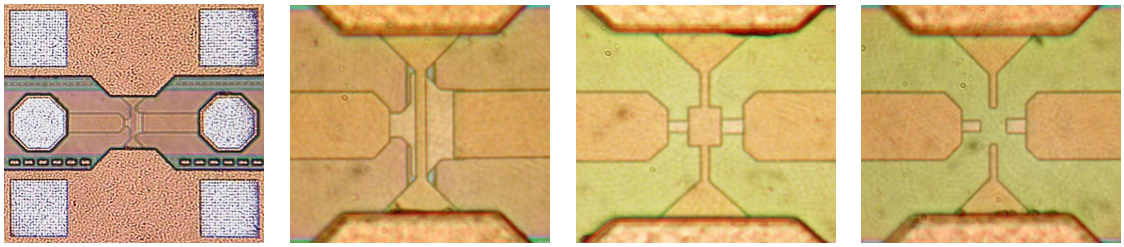
Alongside each of the four devices is an open-circuit structure that can be used for de-embedding purposes. Only the devices from the BiPX2 and BiCMOS9MW processes have associated short-circuit structures[7]. However, due to the close similarity in the layout of the BiPX1 and BiPX3 devices with the BiPX2 device, the short circuit for the BiPX2 was used in de-embedding the test structure parasitics for all three of the experimental transistors.

4.1.3 IHP SG13

For evaluation purposes, sample IHP devices from the SG13 process were obtained. In the SG13 process, 0.12μm HBTs are available with a FSA raised extrinsic base[8], a B doped intrinsic base with C doping, and an in situ As doped polyemitter [97]. What separates this process from the other processes is that the sub-collector is implanted rather than buried under an epitaxial layer, and that

[7]The BiCMOS9MW devices actually have several additional de-embedding structures as detailed in [106]. In this work however, only the open- and short-circuit structures were used.

[8]The extrinsic bases are aligned using selective epitaxy as opposed to a sacrificial emitter.



(a) (b) (c) (d)

Figure 4.4: IHP SG13 HBT test structures. (a) $0.12 \times 16 \times 2.04 \mu m^2$ HBT test structure, (b) enlarged view of the device, (c) short-circuit de-embedding structure, and (d) open-circuit de-embedding structure

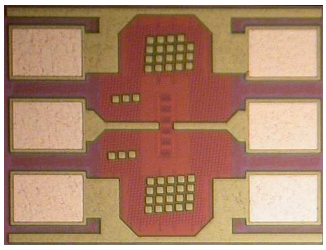


Figure 4.5: Photograph of JAZZ SBC18 test transistor

deep trench isolation is unavailable [107, 108]. The implanted sub-collector will result in an increase in collector resistance and the lack of DTI will result in an increase in C_{CS} as compared to a device with a buried layer and DTI. Reported values of room temperature f_t/f_{max} and BV_{CE0}/BV_{CB0} for the IHP SG13 HBTs are 255/315GHz and 1.8/5.2V, respectively.

The IHP devices evaluated in this investigation were samples provided by the foundry with an effective area[9] of 0.17x16x2.09μm^2 (this is a 16 emitter device). In addition to the device structures, open/short de-embedding structures were also provided. Die photos of the tested devices and de-embedding structures appear in Fig. 4.4.

4.1.4 JAZZ SBC18

The devices evaluated in this study that are manufactured by JAZZ semiconductor are from the 0.18μm SBC18 process line. SBC18 HBTs feature a low-resistance BL sub-collector, a FSA extrinsic-base structure, deep-trench and shallow-trench isolation, and an in situ As doped poly-emitter [109]. Unique features of JAZZ SBC18 devices include an implanted extrinsic base and a triangular Ge profile with a peak Ge content of 30% [62]. Since the JAZZ devices have a triangular

[9]While the drawn emitter width is 0.12μm, the effective emitter width is 50nm larger. Therefore 50nm must be added to all drawn dimensions in determining the effective device size.

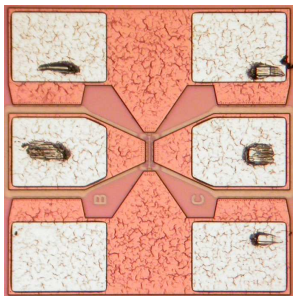


Figure 4.6: Photograph of NXP QUBIC4Xi test transistor

Ge distribution, it is expected that their β_{DC} should have a weaker dependence on temperature than is typical for devices with either trapezoidal or box Ge profiles. Finally, the room temperature peak f_t/f_{max} is 150/190GHz and the BV_{CEO}/BV_{CB0} is 2.2/7.0V.

In order to evaluate the JAZZ devices at cryogenic temperatures, space was allocated on a SBC18 wafer in early 2007. Therefore, custom layouts were made for several HBTs along with the associated de-embedding structures. The tested device area is $0.18 \times 6 \times 10$ μm^2. A Photograph of the test transistor appears in Fig. 4.5

4.1.5 NXP QUBIC4Xi

The final devices evaluated were from NXP's 0.25 μm QUBIC4Xi[10] process, which features SiGe HBTs with BL sub-collectors, in situ As doped poly-emitters, and B doped SiGe bases with carbon doping. Unique features in the QUBIC4Xi process include a QSA raised extrinsic base, and a Ge profile that is stepped rather than grated. Finally, the room temperature peak f_t/f_{max} is 216/177 GHz and the BV_{CEO}/BV_{CBO} is 1.44/5.20V [99, 110, 111].

The QUBIC4Xi devices measured in this work are samples provided by the foundry and are fairly small devices with an effective area of just $0.3 \times 1\mu$m^2. Furthermore, no de-embedding structures were provided. Thus, these devices were only evaluated at dc. A photograph of the test transistor appears in Fig. 4.6.

4.2 Cryogenic Measurement Setup

A closed-cycle cryogenic wafer-probe station was used to characterize the devices, both at dc and RF. A block diagram of the test setup appears in Fig. 4.7 and photos of the cryogenic wafer-probe

[10]It should be noted that packaged NXP BFU725F transistors have also been tested. However, as these devices are not suitable for wafer probing, their performance was not included in the study reported here. For more information regarding the cryogenic performance of these devices, see [43].

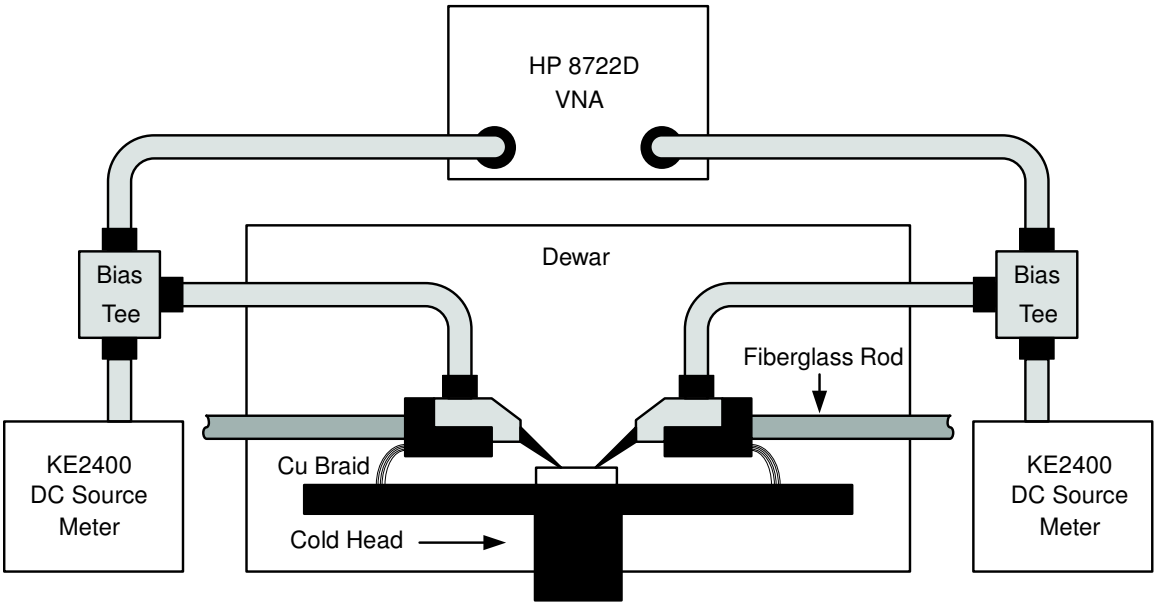
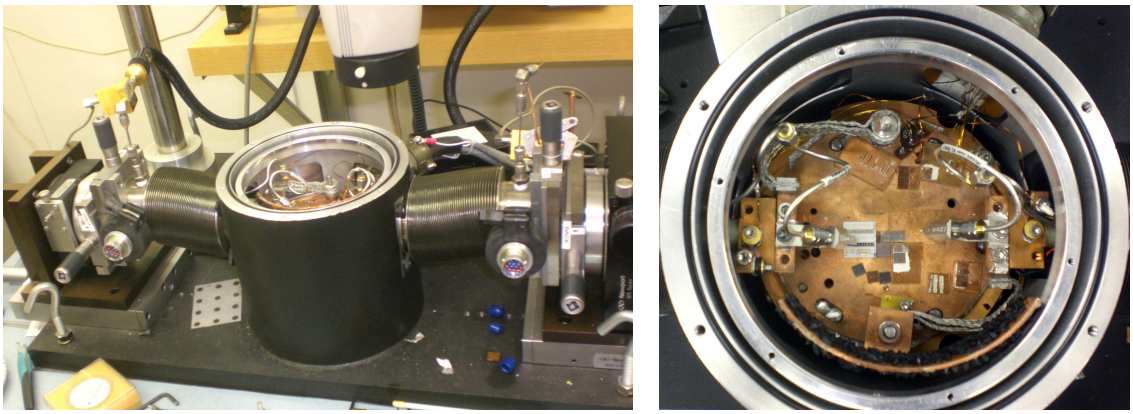


Figure 4.7: Block diagram of cryogenic test setup



(a) (b)

Figure 4.8: (a) side and (b) top view of the cryogenic wafer-probe station located at the Jet Propulsion Laboratory (JPL)

station appear in Fig. 4.8. The probe station can reach temperatures as low as 14 K and an external temperature controller is in place that allows the temperature to be servoed to any temperature from 14 to over 400 K. In order to prevent ice buildup, a strong vacuum is pulled on the chamber using an external vacuum pump, which is in operation when the chamber is at temperatures above 100 K. As the chamber is cooled further, a valve is sealed so that gases cannot enter or escape the chamber and carbon absorber inside of the dewar provides cryo-pumping action to improve the quality of the vacuum. In order to prevent the coaxial cables from presenting an overwhelming

heat load to the refrigerator, stainless-steel 2.4 mm coaxial cables are used to bring in the RF signals and the probe mounts are heatsunk with a copper braid to the copper chuck on which the dies lie. The probes are thermally connected to the probe mounts using a thin layer of indium foil. The temperature is monitored with precision temperature sensors thermally connected via indium foil to the copper chuck as well as each of the probe mounts. Special cryogenic probes rated from 10–300 K and manufactured by SUSS Microtech were used in order to measure the devices as standard probes tend to have contact issues when measuring the same die at multiple temperatures due to a non-uniform expansion of the coplanar probe tips [112, 113]. The dies are either epoxied onto small pieces of copper, which are then thermally connected to the copper chuck using vacuum grease, or placed directly on the copper chuck (with vacuum grease).

DC measurements were made with the VNA source power turned off and all measurements were automated using Matlab scripts. The short-circuit de-embedding structures are used to determine the series resistances as well as the ground return resistances to allow for the effects of these parasitics to be removed from the dc data. For RF measurements, an 8722D VNA was used. A SUSS CS-8 calibration substrate with SOLT standards is used in order to calibrate to the probe-tips and on-wafer open/short structures are used to move the reference plane to the surface of the transistors. More details are given as to the RF measurement scheme in Chapter 6. For all device measurements reported in this work, the emitter and substrate terminals are tied to ground and, unless stated otherwise, V_{CB} is fixed at 0 V. DC measurements were made with the VNA source power turned off and all measurements were automated using Matlab scripts. The short-circuit de-embedding structures are used to determine the series resistances as well as the ground return resistances to allow for the effects of these parasitics to be removed from the dc data. For RF measurements, an 8722D VNA was used. A SUSS CS-8 calibration substrate is used in order to calibrate to the probe-tips and on-wafer open/short structures are used to move the reference plane to the surface of the transistors. More details are given as to the RF measurement scheme in Chapter 6. For all device measurements reported in this work, the emitter and substrate terminals are tied to ground and, unless stated otherwise, V_{CB} is fixed at 0 V.

Chapter 5

DC Modeling

In this chapter, the dc operating characteristics of a variety of state-of-the-art HBTs are discussed and experimental results are compared with the theory introduced in Chapters 2 and 3. The chapter begins with a study of the temperature dependence of the theoretical low-frequency[1] noise performance of state-of-the-art SiGe HBTs, as predicted using the noise theory presented in Chapter 2 in conjunction with measured dc data. This is followed by an investigation in which the temperature dependencies of the terminal currents are studied and compared with theory. Throughout the chapter, similarities and differences between the devices are studied and related to physical properties associated with the various technology platforms. For all measurements reported in this chapter, V_{CB}=0 V and the emitter and substrate terminals are tied to ground.

5.1 Experimental Modeling of Noise in SiGe HBTs, $f \ll f_t$

In this section, the theoretical low-frequency noise parameters of several state-of-the-art HBTs are computed using dc measurements in conjunction with the HBT noise-theory presented in Section 2.2 and the following questions are investigated: 1) What trends can be observed as the devices are cooled? 2) How does the low-frequency cryogenic noise performance of SiGe HBTs vary among foundries? 3) If at all, how is the low-frequency noise-performance at cryogenic temperatures related to the room temperature device parameters?

As discussed in Section 2.2, to first order[2], the low-frequency noise parameters of a SiGe HBT are completely determined by its extrinsic transconductance G_m, extrinsic collector current ideality factor, n_{cx}, and dc current gain, β_{DC}. Explicitly, the low-frequency noise parameters can be

[1] In this work, 'low-frequency' refers to frequencies at which the parasitic capacitances are negligible.
[2] A derivation is given in Appendix E

approximated as[3]

$$T_{min,LF} \approx T_a \frac{n_{cx}}{\sqrt{\beta_{DC}}} = \kappa T_a, \tag{5.1}$$

$$R_{OPT,LF} \approx \frac{\sqrt{\beta_{DC}}}{G_m}, \tag{5.2}$$

and

$$N_{LF} \approx \frac{T_{min,LF}}{2T_0}. \tag{5.3}$$

Referring to equations (5.1)–(5.3), several conclusions can be drawn as to the relationship between the noise parameters of a device and its dc characteristics:

1) $T_{min,LF}$ is reduced by increasing β_{DC} and reducing n_{cx}. It is also linearly dependent upon the ambient temperature. However, the parameter $\kappa = n_{cx}/\sqrt{\beta_{DC}}$ is not explicitly dependent on temperature. Thus, κ is a useful parameter as it can be used to gauge how the dc properties of the device are related to the changes in the low-frequency noise performance as the device is cooled. For instance, if κ is constant with temperature, the implication is that the minimum achievable noise temperature has improved by the same fractional amount as the ambient temperature has dropped. However, if κ decreases as the device is cooled, this means that the noise performance of the device has actually improved by a factor greater than the fractional change in ambient temperature and vice-versa.

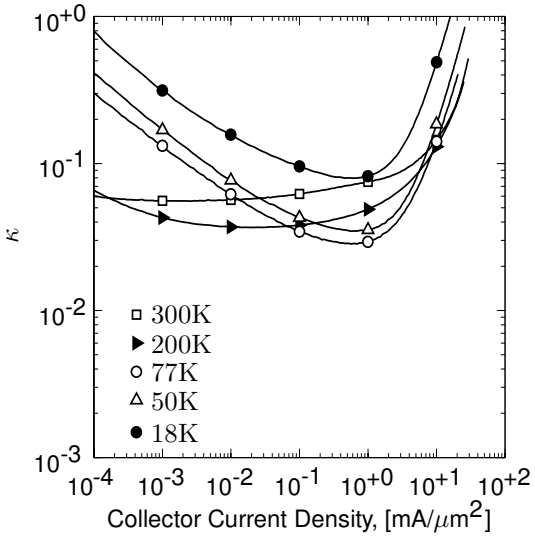
2) R_{OPT} can be increased with either a reduction in G_m or an increase in β_{DC}. Furthermore, $R_{OPT,LF}$ scales inversely with the total device area and therefore can be reduced with an increase in device size. Thus, it is alway possible to transistor with a desired value of $R_{OPT,LF}$ by scaling the total emitter area of the device.

3) N_{LF} is proportional to $T_{min,LF}$ and, to first order, does not need to be determined independently. This is important as it indicates that the low-frequency noise performance of an HBT is completely specified by two noise parameters.

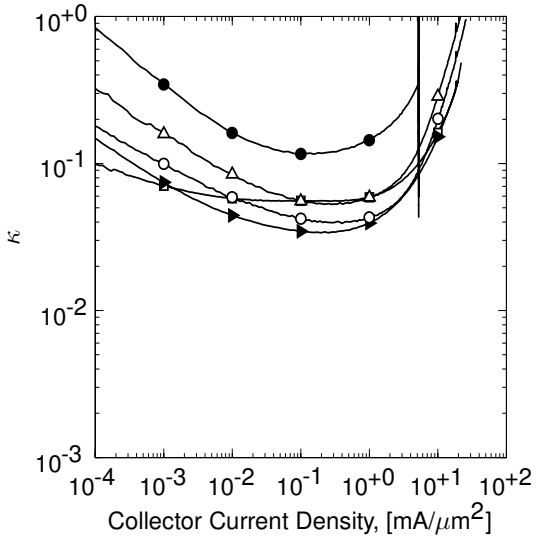
5.1.1 κ and $T_{min,LF}$ Versus Temperature

When optimizing the low-frequency noise parameters, it is of the highest practical importance to minimize $T_{min,LF}$. The reason for this is that both T_{min} and N_{LF} are independent of device size whereas a given value of $R_{OPT,LF}$ can always be obtained by proper device scaling. Furthermore, N_{LF} and $T_{min,LF}$ are coupled and can be minimized simultaneously. Thus, the behavior of κ as a

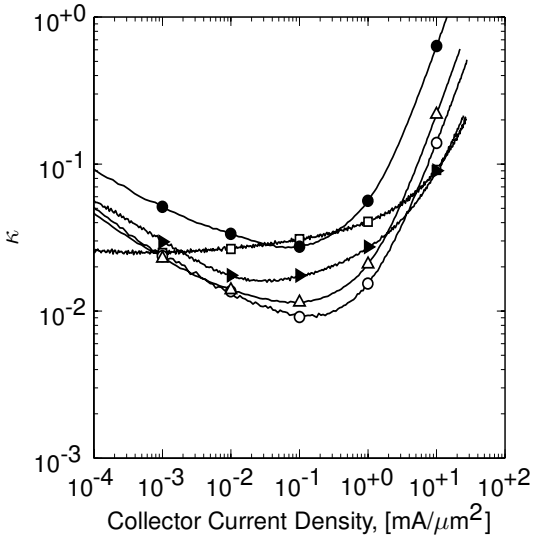
[3] In the following, the noise contributions of r_b and r_e are ignored. This is especially valid at cryogenic temperatures.



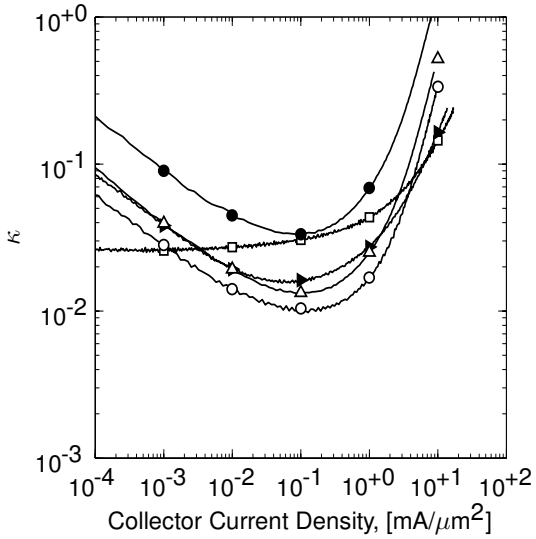
(a) IBM-G4



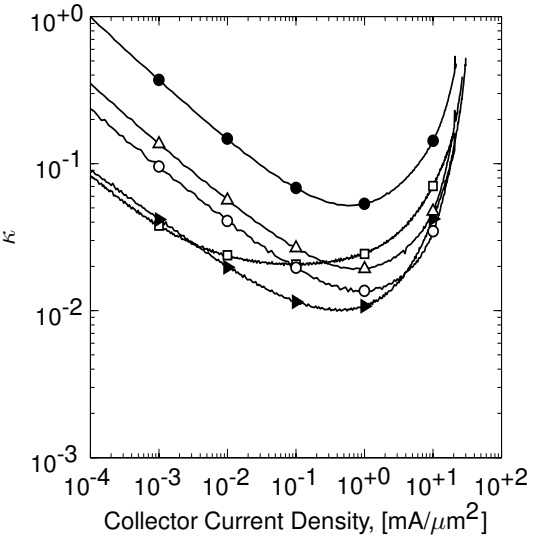
(b) IHP-G4



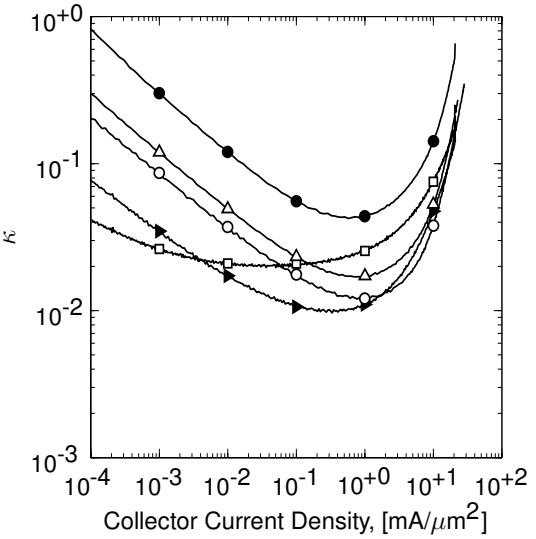
(c) ST-G4



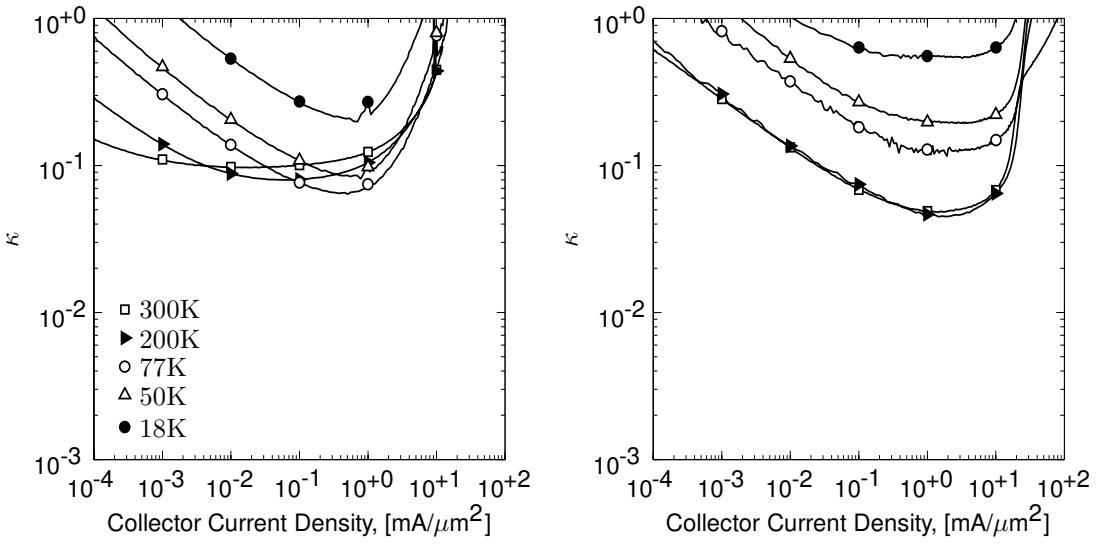
(d) ST-X2



(e) ST-X1



(f) ST-X3



(g) JAZZ-G3 (h) NXP-G3

Figure 5.1: κ versus J_C and T_a for the eight sample transistors

function of temperature is of the utmost importance and serves as a logical starting point for the evaluation of the temperature dependence of the low-frequency noise in SiGe HBTs.

The dc characteristics of the eight different devices were measured at 300, 200, 77, 50, and 18 K and J_C, β_{DC}, G_m, and n_{cx} were extracted as a function of the applied base voltage. These temperatures were chosen for this study as each of these temperatures can be reached with one of four technologies: 200 K can be reached with thermoelectric cooling[4]; 77 K and 50 K can be reached with inexpensive and long-MTF brush-less stirling coolers [115]; finally 18 K can be reached with a standard CTI refrigerator. For these measurements, the base and collector voltages were swept such that the base–collector voltage was fixed at 0 V. From these data, κ was then extracted and the results appear in Fig. 5.1.

Examination of these curves reveals several key features common to all of the devices. To begin with, at room temperature, the minima of κ as a function of J_C is relatively broad, indicating that the room-temperature low-frequency noise performance is not terribly sensitive to bias point. However, as the devices are cooled, it is apparent that the low-frequency noise performance becomes much more sensitive to the bias point as the curves become much more concave[5]. As the curves are directly proportional to $T_{min,LF}$, this indicates that the range of collector current densities over which near optimum noise performance can be achieved grows smaller as the temperature is reduced.

[4]Standard coolers can reach ∼160 K and can be mounted in a vacuum sealed package [114].

[5]The exception is the NXP-G3 device, for which the general shape of the κ versus J_C curve does not change with temperature.

In addition to the general shapes of the curves, there is information in the temperature dependence of the minima of κ with respect to J_C. If $\partial\kappa_{min}/\partial T_a$ is negative, the implication is that the optimum low-frequency noise temperature of the device is dropping more rapidly than the ambient temperature and vice-versa. Referring again to Fig. 5.1, for all of devices, κ_{min} drops as the device is cooled from 300 K to 200 K; in this temperature range, the optimum noise temperature is dropping more rapidly than the ambient temperature for all of the devices. Thus, cooling a SiGe HBT from 300 to 200 K results in a significant enhancement in the low-frequency noise of all eight devices.

As the temperature is lowered from 200 to 77 K, a difference can be identified between the devices as κ_{min} continues to drop for the IBM-G4, ST-G4, ST-X2, and JAZZ-G3 devices but begins to rise for the other four devices. Thus, in cooling from 200 to 77 K, a fractional reduction in optimum noise temperature larger than the fractional reduction in ambient temperature is only observed in half of the devices. Finally, as the devices are cooled below 77K, an increase in the value of κ_{min} is seen in observed for all of the devices, indicating that the fractional reduction in ambient temperature is now greater than the fractional reduction in optimum noise temperature at low-frequencies. In summary, for each of the devices, there is some temperature, $T_{\kappa,min}$, such that $\kappa_{min}\left(T_{\kappa,min}\right) = \min\left\{\kappa_{min}\left(T_a\right)\right\}$. At ambient temperatures above $T_{\kappa,min}$ a fractional change in T_a will produce a larger fractional change in $T_{min,LF}$ and vice-versa.

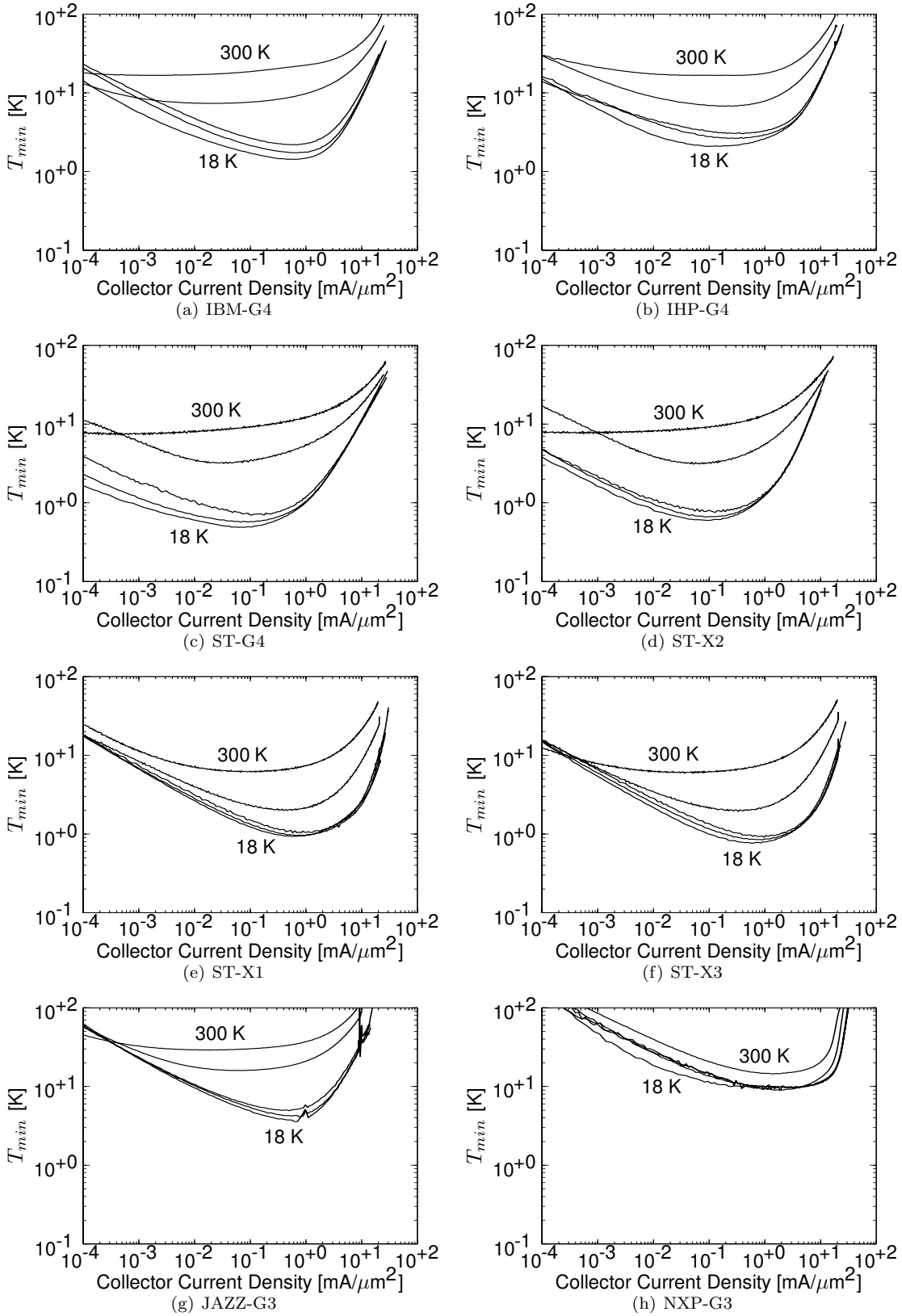
In order to convert the curves in Fig. 5.1 to noise temperature curves, the values of κ simply need to be multiplied by T_a. This exercise was done and the results appear in Fig. 5.2. In addition,

$$T_{min,LF,J_C} \equiv \min_{J_C}\left\{T_{min,LF}\left(J_C\right)\right\} \tag{5.4}$$

was computed for each device and the resulting values along with the associated collector current densities appear in Table 5.1. The spread in noise temperatures among the devices, both in terms of what can be achieved as well as how things change as the devices are cooled is quite remarkable. For instance, at 300K, T_{min,LF,J_C} varies among the devices from 5.9 to 29 K; a factor of nearly five. Furthermore, at cryogenic temperatures, the spread is even wider with the best device being able to achieve a low-frequency input-referred noise temperature of 0.5 K and the worst device only able to reach 9.6 K.

5.1.2 The Temperature Dependence of n_{cx} and β_{DC}

It is desirable to determine the reason that the low-frequency noise performance of one device is better than that of another and the information afforded by the figure-of-merit κ provides insufficient



(a) IBM-G4 (b) IHP-G4

(c) ST-G4 (d) ST-X2

(e) ST-X1 (f) ST-X3

(g) JAZZ-G3 (h) NXP-G3

Figure 5.2: Low-frequency limit of T_{min} for each of the eight devices at 300, 200, 77, 50, and 18 K. The curves are calculated from the extracted values of β_{DC} and n_{cx}.

Table 5.1: Global minima for $T_{min,LF}$ as a function of J_C and the associated J_C for state-of-the-art devices. $T_{min,LF}$ values are computed from measured data using equation (5.1).

	T_{min,LF,J_C}, K					$J_{C,T_{min,LF}}$, mA/μm^2				
T_a, K	300	200	77	50	18	300	200	77	50	18
IBM-G4	16.7	7.3	2.2	1.7	1.4	2.4×10^{-3}	1.4×10^{-2}	5.7×10^{-1}	5.5×10^{-1}	4.6×10^{-1}
IHP-G4	16.6	6.8	3.0	2.6	2.1	2.8×10^{-1}	2.5×10^{-1}	4.0×10^{-1}	2.8×10^{-1}	1.6×10^{-1}
ST-G4	7.3	3.1	0.7	0.6	0.5	3.8×10^{-4}	3.4×10^{-2}	1.1×10^{-1}	6.8×10^{-2}	6.2×10^{-2}
ST-X1	6.9	2.0	1.0	1.0	0.9	6.7×10^{-2}	4.0×10^{-1}	6.9×10^{-1}	6.8×10^{-1}	5.6×10^{-1}
ST-X2	7.5	3.1	0.8	0.7	0.6	5.6×10^{-4}	5.4×10^{-2}	1.3×10^{-1}	8.8×10^{-2}	9.2×10^{-2}
ST-X3	5.9	1.9	0.9	0.8	0.8	5.0×10^{-2}	3.3×10^{-1}	9.7×10^{-1}	8.9×10^{-1}	5.7×10^{-1}
JAZZ-G3	29.0	16.0	4.9	4.1	3.6	2.5×10^{-2}	6.7×10^{-2}	5.1×10^{-1}	7.8×10^{-1}	7.1×10^{-1}
NXP-G3	14.4	9.0	8.9	9.7	9.6	$1.5 \times 10^{+0}$	$1.5 \times 10^{+0}$	$1.9 \times 10^{+0}$	$3.5 \times 10^{+0}$	5.7×10^{-1}

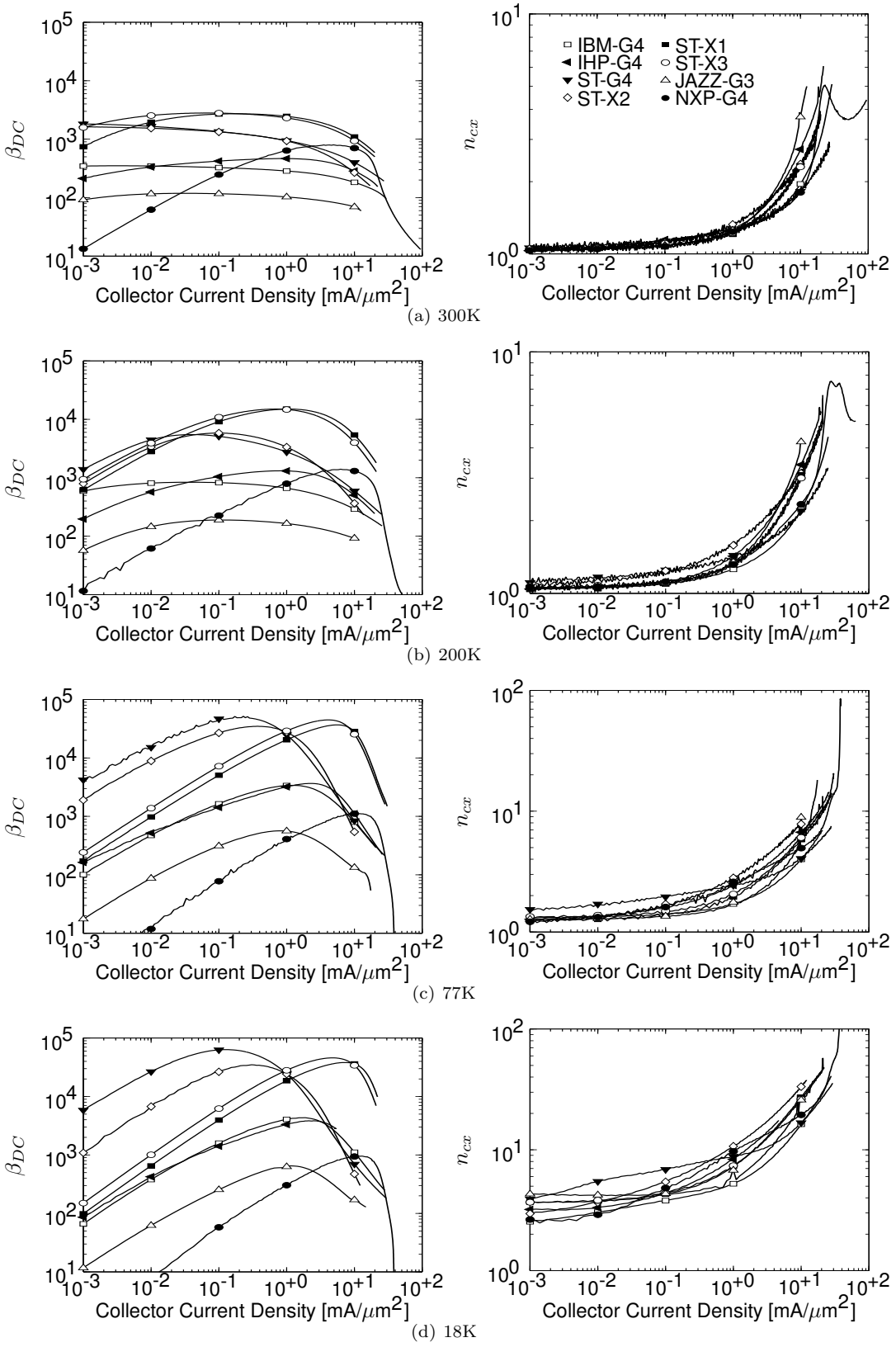
Table 5.2: Peak β_{DC} and the associated value of collector-current density

	$\beta_{DC,pk}$					$J_{C,\beta_{pk}}$, mA/μm^2				
T_a, K	300	200	77	50	18	300	200	77	50	18
IBM-G4	351	839	3400	4040	4340	3.1×10^{-3}	3.5×10^{-2}	$1.2 \times 10^{+0}$	$1.7 \times 10^{+0}$	$1.9 \times 10^{+0}$
IHP-G4	466	1320	3660	3890	3850	8.9×10^{-1}	8.4×10^{-1}	$2.3 \times 10^{+0}$	$2.3 \times 10^{+0}$	$2.5 \times 10^{+0}$
ST-G4	1840	5490	51600	56600	63600	5.0×10^{-4}	4.2×10^{-2}	2.3×10^{-1}	1.5×10^{-1}	1.3×10^{-1}
ST-X1	2720	15000	36800	37800	38200	1.8×10^{-1}	$1.1 \times 10^{+0}$	$5.7 \times 10^{+0}$	$6.3 \times 10^{+0}$	$7.2 \times 10^{+0}$
ST-X2	1630	5830	34800	36000	35000	8.9×10^{-4}	9.9×10^{-2}	3.7×10^{-1}	3.4×10^{-1}	3.0×10^{-1}
ST-X3	2810	15000	44800	45600	45900	6.6×10^{-2}	7.8×10^{-1}	$4.1 \times 10^{+0}$	$4.6 \times 10^{+0}$	$4.7 \times 10^{+0}$
JAZZ-G3	118	189	566	623	651	3.5×10^{-2}	1.1×10^{-1}	7.8×10^{-1}	$1.0 \times 10^{+0}$	$1.2 \times 10^{+0}$
NXP-G3	801	1400	1150	1110	958	$4.9 \times 10^{+0}$	$6.3 \times 10^{+0}$	$10.7 \times 10^{+0}$	$1.1 \times 10^{+1}$	$1.2 \times 10^{+1}$

means to do so, as it is a composite of two different parameters. Thus, in order to understand the behavior of κ with temperature, it is necessary to investigate what happens to the values of β_{DC} and n_{cx} as the devices are cooled. This task is easily accomplished as these parameters are readily available. The dc current gain and extrinsic collector current ideality factor are plotted as a function of collector current density at 300, 200, 77, and 18 K in Fig. 5.3. Referring to the plots of the dc current gain, several important features and trends can be noted:

1) β_{DC} is greatly enhanced by cooling to 18K for all devices except the NXP-G3 transistor. Thus, if the extrinsic collector-current ideality factor were equal to unity (i.e., the collector-current were ideal), then the low-frequency noise performance improvement of each device occurring due to cooling should be determined entirely by the magnitude of the increase in β_{DC}. In this case, significant improvements should be seen in all of the devices except for the sample from the NXP-G3 process.

2) The peak value of β_{DC} varies significantly across processes.



(a) 300K

(b) 200K

(c) 77K

(d) 18K

Figure 5.3: β_{DC} and n_{cx} as a function of J_C for each of the eight devices

3) The collector current density at which peak β_{DC} occurs varies significantly across both temperature and manufacturer.

4) As the devices are cooled, the temperature dependence of β_{DC} becomes stronger.

Similarly, referring to the plots of the extrinsic collector current ideality factors, the following observations can be made:

1) At a fixed ambient temperature, n_{cx} increases monotonically with increasing collector current.

2) At a fixed collector current density, n_{cx} increases monotonically with decreasing temperature.

3) For collector current densities below $\sim$1mA/μm^2, n_{cx} is only weakly dependent upon J_C.

4) At temperatures of $\sim$50 K and lower, the low-current limit of n_{cx} is greater than unity. This implies that the collector-current transport is occurring under non-equilibrium conditions.

5) For collector current densities above $\sim$1mA/μm^2, n_{cx} is highly dependent on J_C. This sharp increase can be attributed to high-injection effects. The onset and magnitude of this effect appears to get worse at cryogenic temperatures. This may be in part due to freeze-out effects if any of the intrinsic semiconductor regions are doped below the Mott-transition.

6) The uniformity of the n_{cx} vs J_C curves for the different devices gets slightly worse as the ambient temperature is lowered. Nonetheless, the overall trend is consistent across the different processes.

At room temperature, since n_{cx} is very close to unity, the minimum achievable low-frequency noise temperature of an HBT is generally determined by where the peak value of β_{DC} occurs.

However, taking into account both the observations about β_{DC} and those made with respect to n_{cx}, it becomes clear that as the device is cooled, both quantities become important as n_{cx} begins to depend upon J_C. Therefore, there will be some tradeoff between minimizing n_{cx} and maximizing β_{DC}. In other words, due to the sharp rise in n_{cx} at modest current levels in HBTs operating at cryogenic temperatures, the optimum bias point may be below the collector current density which produces the peak value of β_{DC}. Thus, simply maximizing the value of β_{DC} without considering the collector-current density at which the peak occurs is a risky way in which to optimize the low-frequency noise performance of the device, as the improvement gained from an increase in β_{DC} can easily be cancelled out due to an increase in n_{cx}.

5.1.3 The Figure-of-Merit, $\Delta T_{min,LF,J_C}(T_a)$

Thus far, the low-frequency noise performance of several state-of-the-art SiGe HBTs has been presented along with a general discussion of what happens to some of the key parameters as the device is cooled. In addition, it has been implied that T_{min,LF,J_C} is not necessarily a monotonically decreasing function with respect to $\beta_{DC,pk}$. However, this study has not yet resulted in a solid understanding of what causes the noise performance of one device to improve more than that of another as the ambient temperature is decreased. In order to clarify this issue, a useful figure-of-merit to investigate is,

$$\Delta T_{min,LF,J_C}(T_a) \equiv \frac{T_{min,LF,J_C}@300\text{ K}}{T_{min,LF,J_C}@T_a} = \frac{300\text{ K}}{T_a}\frac{\kappa_{min}(300\text{ K})}{\kappa_{min}(T_a)}. \tag{5.5}$$

$\Delta T_{min,LF,J_C}$ is physically important as it is a measure of the improvement in the achievable low frequency noise performance that is gained with cooling. In Fig. 5.4(a), $\Delta T_{min,LF,J_C}$ is plotted as a function of T_a for each of the eight devices. It is remarkable how much variation is observed in the values of $\Delta T_{min,LF,J_C}(T_a)$ among the different processes.

In order to verify that the fractional increase in $\beta_{DC,pk}$ does not tell the whole story, the value of $\Delta T_{min}(18\text{ K})$ was plotted for each device as a function of $\sqrt{\beta_{DC,pk}(18\text{ K})}/\sqrt{\beta_{DC,pk}(300\text{ K})}$, and a linear regression of the form $y = \hat{\alpha}_0 + \hat{\alpha}_1\sqrt{x}$ was applied. The coefficients obtained in this fitting exercise are[6] $\hat{\alpha_0} = 0.2 \pm 5.5$ and $\hat{\alpha}_1 = 2.5 \pm 1.4$ and the raw data along with the trend line and normalized residuals (inlay) appears in Fig. 5.4(b). If n_{cx} plays no role in the determination of $\Delta T_{min,LF,J_C}$, then this fitting procedure should produce a good fit to the data. However, while the linear regression certainly implies that the improvement in noise gets better with larger increases in $\beta_{DC,pk}$, the result also implies that the improvement is not determined entirely by the increase in $\beta_{DC,pk}$, as the slope coefficient cannot even be determined with 95% confidence to within a factor of two. As suspected, the fractional increase in $\beta_{DC,pk}$ is only part of the explanation of the noise. Thus, a more complete explanation is desired.

Referring back to equation (5.1), we see that the only remaining possible explanation is that $\Delta T_{min,LF,J_C}$ is also related to the temperature dependence of n_{cx}. Furthermore, at room temperature, it can be observed that n_{cx} is nearly unity even at biases exceeding $1\text{mA}/\mu\text{m}^2$, implying

[6] The coefficients described in this section are specified to a 95% confidence interval using the t-statistic. It is assumed that the residuals are approximately normally distributed (i.e., any higher order moments beyond the variance are negligible). For information regarding confidence intervals of linear estimators, please see a statistics book such as [116].

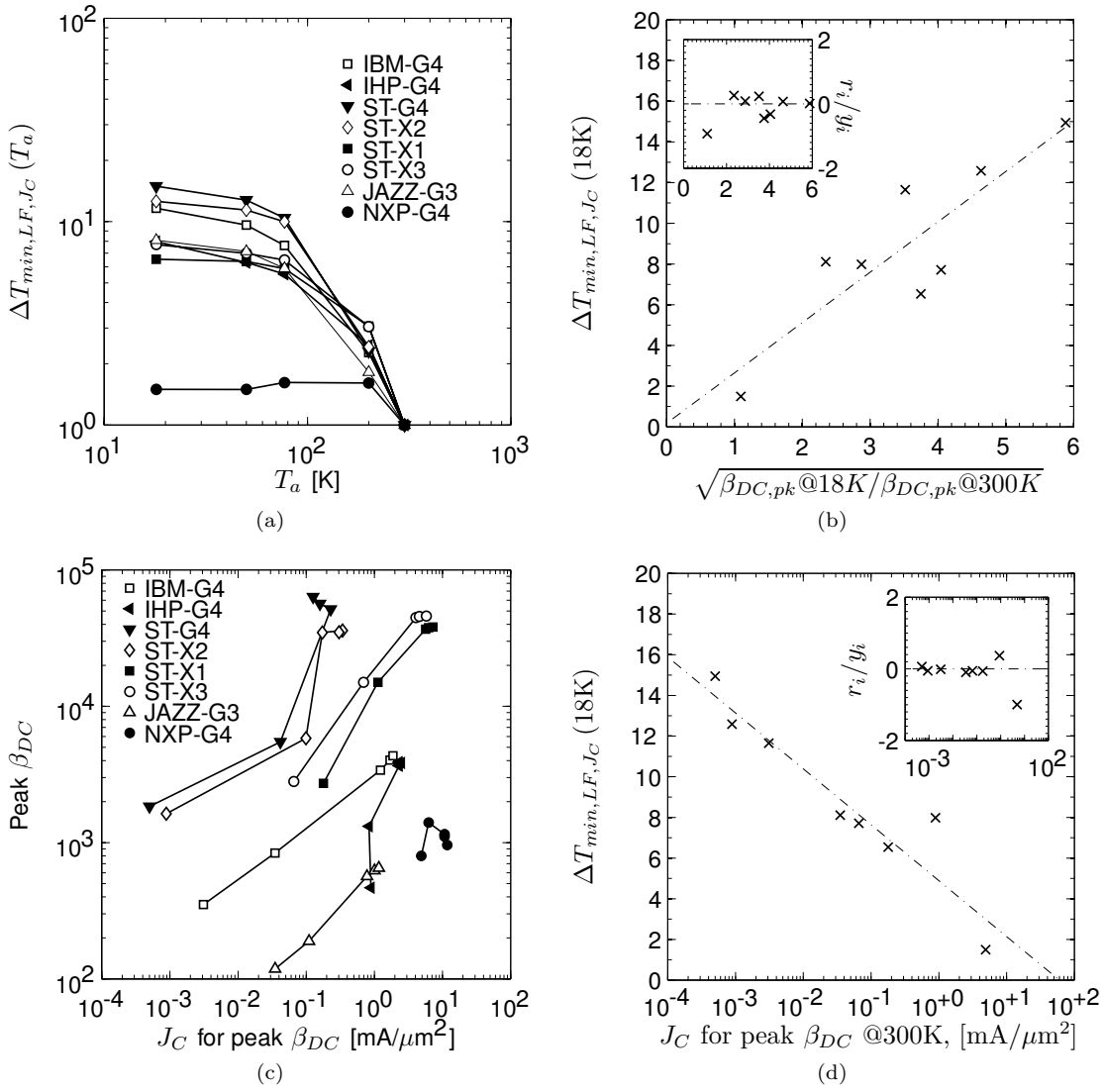


Figure 5.4: (a) $\Delta T_{min,LF,J_C}(T_a)$ as a function of ambient temperature. It should be noted that the IHP-G4 markers are buried beneath the ST-X1 and ST-X3 markers. (b) $\Delta T_{min,LF,J_C}$ (18 K) as a function of the square root of the ratio of β_{DC}@18 K to β_{DC}@300 K for each of the eight devices. The trend line is given as $y = 0.2 + 2.5x$. However, the confidence intervals indicate that the coefficient estimates are quite poor. (c) Peak β_{DC} as a function of the collector current density required to achieve peak β_{DC} for each of the different technologies. Markers are positioned at 300, 200, 77, 50, and 18 K with the leftmost marker being at 300 K on each trace. (d) $\Delta T_{min,LF,J_C}$ (18 K), plotted for each of the eight devices as a function of the collector current density at which the room temperature peak β_{DC} occurs. The trend line is given as $y = 4.89 - 2.75 \log 10\{x\}$.

that $T_{min,J_C}(300\ K)$ is solely[7] determined by $\beta_{DC,pk}(300\ K)$. As mentioned earlier, at cryogenic temperatures n_{cx} is fairly uniform across different technologies and increases rapidly with increasing collector current density. Therefore, it can be assumed that given two devices with comparable

[7]This is a bit of an exaggeration. At room temperature, n_{cx} is marginally degraded above 1 mA/μm^2. However, this change will only impact the NXP-G3 device, for which $\beta_{DC,pk}$@300 K occurs at 4.9 mA/μm^2.

$\beta_{DC,pk}$, the device in which the peak occurs at the lower value of J_C will most likely demonstrate the better low-frequency noise performance of the two devices. To investigate this notion, it is helpful to look at a plot of $\beta_{DC,pk}$ as a function of the associated collector current density. Such a plot appears in Fig. 5.4(c). It is quite apparent that the collector-current density associated with β_{DC} increases significantly as the device is cooled. Thus, it seems possible that there may be a correlation between the collector current density at which the room temperature β_{DC} occurs and $\Delta T_{min,LF,J_C}$.

In order to test this hypothesis, a curve fit was performed to $\Delta T_{min,LF,J_C}\,(18\ K)$ versus the logarithm of the collector current associated with $\beta_{DC,pk}$@300 K. The coefficients obtained are $\hat{\alpha}_0 = 4.9 \pm 1.2$ and $\hat{\alpha}_1 = -2.8 \pm 0.3$, and the results of the fitting process appear in Fig. 5.4(d). It is apparent from the residuals as well as the confidence intervals that there is statistically significant evidence that the improvement in low frequency noise performance associated with cooling the transistors is related to the current density required to achieve $\beta_{DC,pk}$@300 K, with higher associated values of J_C coinciding with smaller improvements in the achievable noise and vice-versa. This is very important because it implies that candidate devices for low-noise cryogenic applications can be screened at room temperature by measuring the collector current density at which peak β_{DC} occurs.

With this analysis complete, the explanation seems fairly trivial; the current density for peak β_{DC} increases with decreasing temperature. Thus, in order to avoid taking a hit on the low-frequency noise due to a degradation in the extrinsic collector-current ideality factor, it is necessary that the room temperature peak value of β_{DC} occur at a reasonably low bias current. For devices which have fairly high bias currents associated with $\beta_{DC,pk}$ at room temperature, the optimum bias point at 18 K will lie below the collector current density required for $\beta_{DC,pk}$ due to a limiting mechanism associated with n_{cx}, which increases rapidly at bias currents above $1\text{mA}/\mu\text{m}^2$. Thus, the actual β_{DC} which results in minimum noise will be significantly lower than peak of $\beta_{DC,pk}$ and the noise performance will not benefit from the full increase in peak dc current gain.

5.2 Gummel Curves and Modeling of Terminal Currents

In this section, we investigate the behavior of the terminal currents of SiGe HBTs operating at low-injection as the devices are cooled. The method used for this study is as follows:

1) Measure the dc Gummel curves for each of the eight devices at 300, 200, 77, 50, and 18 K.

2) Extract physics based large-signal equivalent circuit models that are valid under low-injection.

3) Identify temperature dependent trends in the coefficients well as in the overall characteristics.

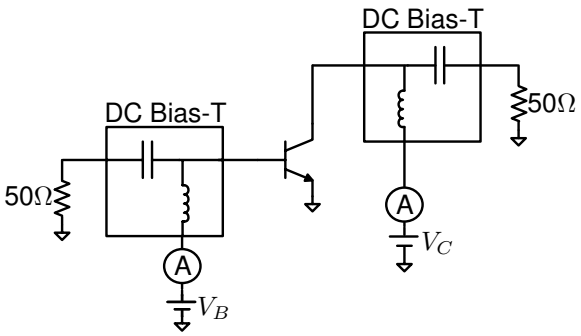


Figure 5.5: Typical setup used to measure Gummel curves

The procedure described above relies upon the analysis of dc Gummel curves. To measure such a curve, both the base and collector are driven with voltage sources and the dc currents are measured as a function of base voltage under the constraint that V_{CB} is fixed[8]. As with all other measurements described in this work, both the emitter and substrate are tied to ground.

A typical test setup for measuring the dc Gummel curves appears in Fig. 5.5. When making dc measurements, it is very important that the device sees a high quality RF termination as oscillations can easily occur at high bias levels if the RF termination impedance is not well defined. Thus, the dc voltages are fed through dc bias-tees with the RF ports terminated with broadband 50 Ω loads. Generally, the voltages are swept in a linear fashion, producing a quasi-logarithmic sampling of currents. Once the data are acquired, they are plotted against V_b on a semi-logarithmic scale. In Fig. 5.6, an example of a typical Gummel curve is shown and four key regions are identified:

1) The collector current is mainly due to the ideal transport current and the base current is mainly due to recombination in the base. The dominant components of the base current density in this region can be described using a diode model: $J_{B,RC} = J_{B0,RC} \exp\{V_{BE}/n_{b,rc}V_T\}$, where J_{BR0} is the base recombination saturation current density and $n_{b,rc}$ is the low-injection base recombination current ideality factor[9]. Similarly, the collector current can be written in terms of a transfer current, which mathematically takes the form of a diode: $J_C = J_{C0} \exp\{V_{BE}/n_{c0}V_T\}$, where J_{C0} is the saturation current density corresponding to the collector current and n_{c0} is the low-injection collector current ideality factor[10]. In all cases, $n_{c0} \geq 1$, with the lower limit being the ideal case.

[8] i.e., the collector voltage and base voltage are swept together such that V_C-V_B is kept constant.

[9] The term "ideality factor" refers to a forward biased diode modeled using Boltzmann statistics in which the current is the result of a diffusion process, and does not have physical meaning with respect to the how ideal the recombination current is.

[10] n_{c0} is equal to n_c in the low-current limit. However, in general, $n_c = I_C/g_m V_T$ has bias dependence whereas n_{c0} is a bias independent parameter.

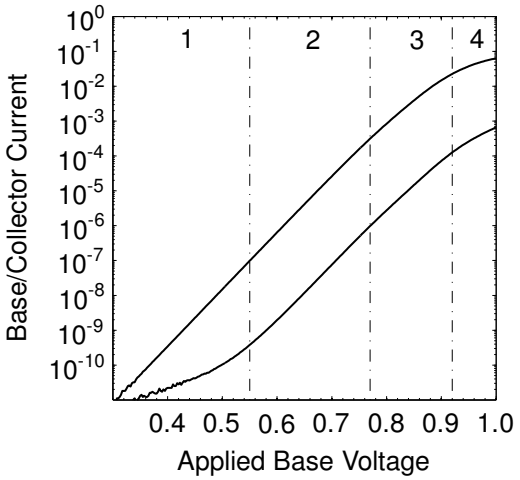


Figure 5.6: Typical Gummel curve. Four regions of interest are roughly specified on the plot. (1) Base current dominated by recombination currents. Collector current ideal. (2) Ideal base and collector currents. (3) Reduction of slope due to resistances. (4) Onset of high-injection effects

.

2) The collector current is still ideal; however, now the ideal component of the base current has begun to dominate. In this region, the base current density can be described by a two diode model, with one diode included to account for the recombination current and the other to describe the ideal current component: $J_B = J_{BI} + J_{B0,RC} = J_{B0} \exp\{V_{BE}/n_b V_T\} + J_{B0,RC} \exp\{V_{BE}/n_{br} V_T\}$. The mathematical description of the collector current is unchanged.

3) In this region, the logarithms of both the collector and base current densities are no longer linear with respect to V_B due to parasitic emitter and base resistances, which result in a reduction of V_{BE} for a given value V_B. If the values of the access resistances are known, then it is possible to plot the currents as a function of the intrinsic base–emitter voltage, V_{BE}, in which case the responses in this region will be linear.

4) In this region, high-injection effects come into play causing the currents to deviate further from their ideal values. Thus they would not appear linear even if plotted against the intrinsic base–emitter voltage.

Upon studying the structure of the Gummel curves, it is possible to construct the equivalent circuit model shown in Fig. 5.7, which captures all of the effects described above. It should be noted that, albeit with a penalty in performance, a HBT can be operated in the inverse mode[11] [117]. Therefore, to fully describe the dc terminal currents, additional diodes and transport current sources are needed to account for the reverse currents. Although these components are negligible under

[11] i.e. with the collector and emitter terminals swapped

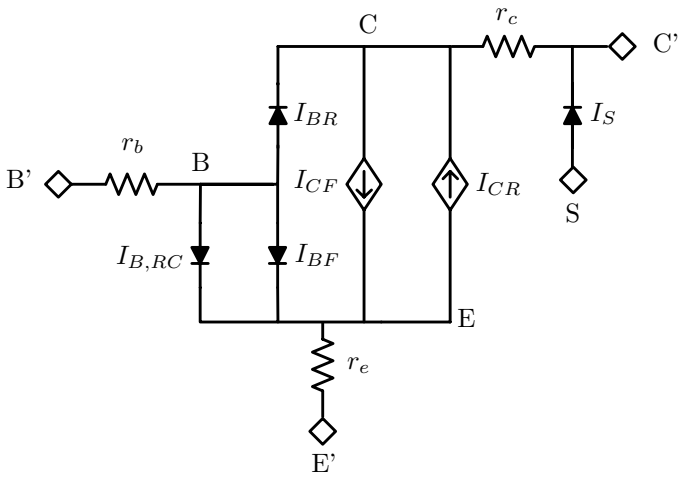


Figure 5.7: The dc equivalent circuit model for a HBT

forward-active operation, and are therefore not extracted here, they are included for completeness as they will become useful later on in Chapter 6.

In order to account for the high-injection effects, the expressions for the ideal diode and transport currents are typically written in a slightly different form as compared to those given above as [60]

$$J_{BI} = \frac{q}{G_E} e^{V_{BE}/n_b V_T} \qquad [\mathrm{A}/\mu\mathrm{m}^2] \tag{5.6}$$

and

$$J_C = \frac{q}{G_B} e^{V_{BE}/n_{c0} V_T} \qquad [\mathrm{A}/\mu\mathrm{m}^2], \tag{5.7}$$

where G_B and G_E are base and emitter *Gummel numbers*, and represent the total mobile charge in these regions [55]. As the Gummel numbers are a function of charge density, they increase under high-injection condition resulting in the degradation of collector and base saturation current densities. In order to maintain consistency with the definitions above, the bias dependencies of the Gummel numbers will be defined as perturbations to their low-injection values:

$$G_B = q \frac{1 + \Phi_C}{J_{C0}} \tag{5.8}$$

and

$$G_E = q \frac{1 + \Phi_E}{J_{B0}}, \tag{5.9}$$

where Φ_C and Φ_E are bias dependent parameters describing the degradation of the currents under high-injection. Furthermore, Φ_C and Φ_E are monotonically increasing parameters that are infinitesimally small in the low-injection regime. It should be noted that a similar approach is used in compact modeling [67]. Thus, the final models of the currents including high injection effects are:

$$J_B = \frac{J_{B0}}{1+\Phi_E} e^{V_{BE}/n_b V_T} + J_{B0,RC} e^{V_{BE}/n_{b,rc} V_T} \tag{5.10}$$

and

$$J_C = \frac{J_{C0}}{1+\Phi_B} e^{V_{BE}/n_{c0} V_T}. \tag{5.11}$$

The extraction and interpretation of the low-injection coefficients is the focus of much of the remainder of Section 5.2. Before continuing on with the extraction of the coefficients that describe the base and collector currents in the low-injection regime, we will begin by examining the general characteristics of the Gummel curves and how they change with temperature.

5.2.1 Gummel Plots: General Observations

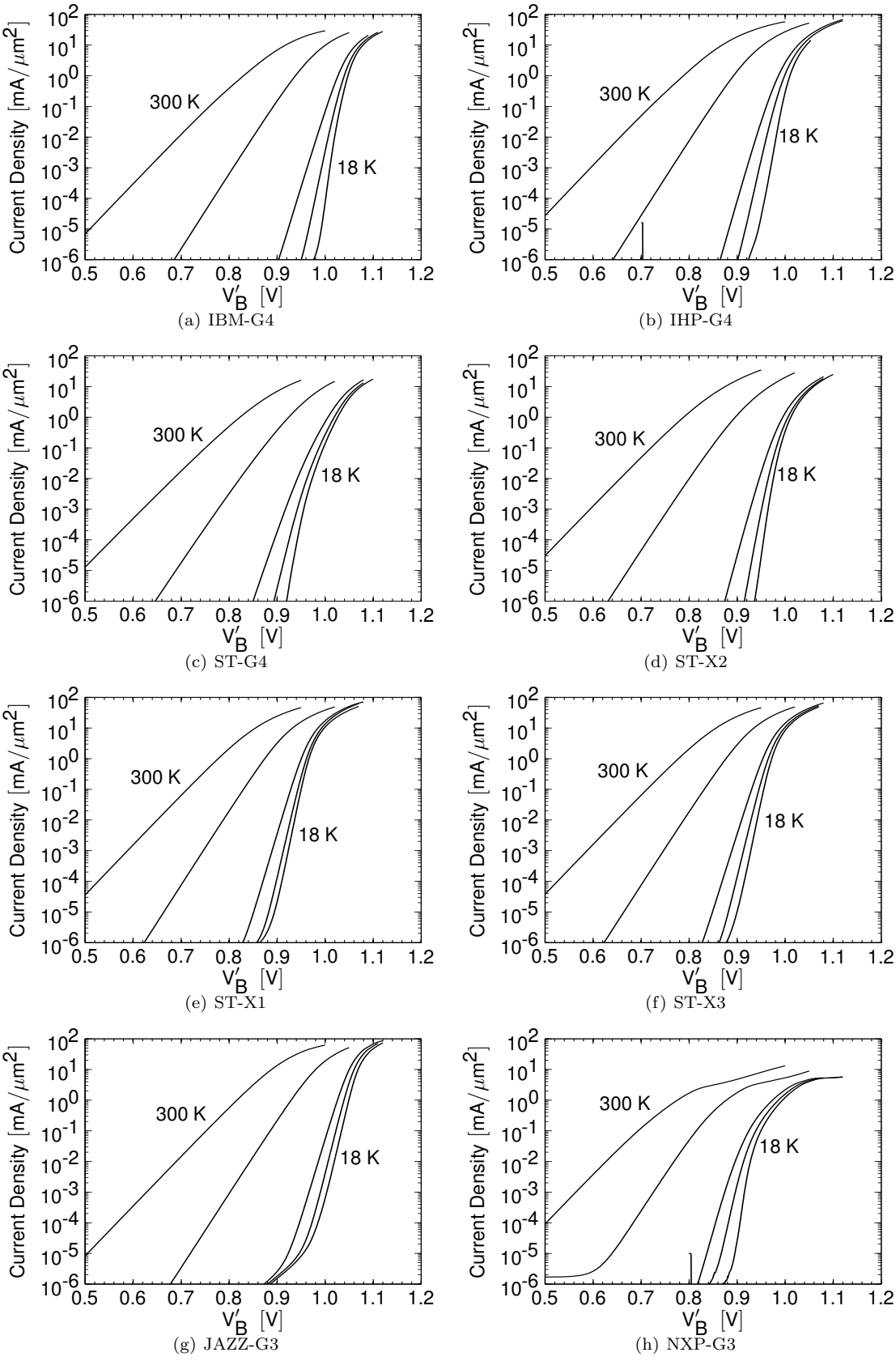
The Gummel curves were measured for each of the eight devices at 300, 200, 77, 50, and 18 K. In Fig. 5.8, the Gummel curve associated with the collector currents of each of the eight devices is shown; inspection of these plots reveals several interesting features:

1) The base–emitter voltage for a given collector current density increases monotonically with decreasing temperature, with the greatest fractional increase occurring during cooling from 300 to 77 K. In Chapter 3, we saw that J_C is proportional to n_{io}^2, which theoretically should drop by approximately 600 orders of magnitude when going from 300 to 18 K. While doping and Ge-induced bandgap narrowing effects should make this effect considerably smaller, they do not explain the fact that the change is getting weaker at lower temperatures. Thus, the fact that V_{BE} increases less as the temperature drops below 77 K is indicative of the presence of non-equilibrium transport phenomena. This topic will be revisited later in the chapter.

2) The slope of the collector current density as a function of base voltage is getting steeper as the device is cooled. This is expected as the low-injection transconductance of the devices increases with cooling[12].

3) While the majority of the devices have a rather smooth roll-off at high injection, the NXP-G3

[12] The increase was seen indirectly through the observed changes in n_{cx}.



(a) IBM-G4 (b) IHP-G4

(c) ST-G4 (d) ST-X2

(e) ST-X1 (f) ST-X3

(g) JAZZ-G3 (h) NXP-G3

Figure 5.8: Gummel plots for collector currents at 300 K, 200 K, 77 K, 50 K, and 18 K

device exhibits abnormal behavior and undergoes quasi-saturation at cryogenic temperatures.

4) The collector current density at which resistive and high-injection effects begin to dominate appears to drop a bit as the temperature is lowered.

5) In some of the low temperature curves, such as those corresponding to the JAZZ device, it appears that there is an additional component to the collector current at low bias voltages. This current is assumed to be a leakage current and will require an extra transport current in the model.

The Gummel curves were also measured for the base currents and the resulting plots are shown in Fig. 5.9. Inspection of these curves leads to several notable observations:

1) For most cases, the base–emitter voltage for a given base current density increases monotonically with decreasing temperature. However, there are some instances at low-injection where this is not the case.

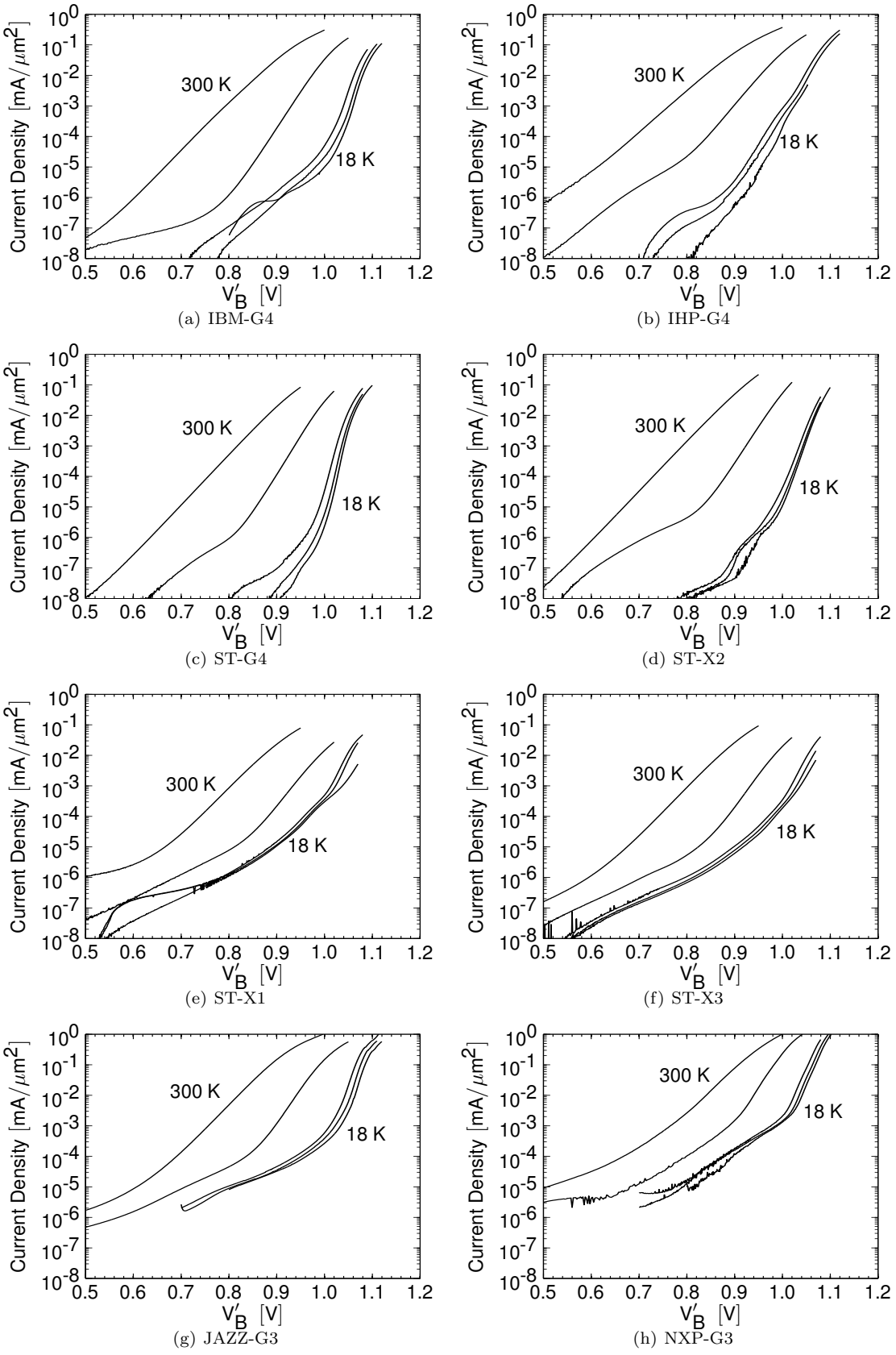
2) The neutral base recombination current appears to have a bigger influence at lower temperatures. This is not surprising as the magnitudes of the ideal base currents drop considerably with cooling.

3) In general, the curves seem to get more structure as the temperature is lowered.

Finally, to allow for identification of differences in the way they change as function of temperature, the base and collector currents of all the devices are plotted side by side at 300, 200, 77, and 18 K in Fig. 5.10. In general, the collector currents appear to be much more well behaved than the base currents. For instance, while the turn on voltage varies among the devices, the general shape of the J_C curves is similar among all devices, even as the temperature is lowered to 18 K. On the other hand, the base current contours vary tremendously among the various devices, and this lack of uniformity grows worse as the temperature is lowered. In order to better understand the behavior of the terminal currents as a function of temperature, a quantitative analysis will be conducted in the following section.

5.2.2 Modeling of Currents at Low-Injection

As described above, at low-injection, the base current is modeled as two diodes, with the first diode representing the ideal current component and the second representing a recombination current. Thus, four coefficients are required in order to fully describe the base current under low-injection



(a) IBM-G4 (b) IHP-G4

(c) ST-G4 (d) ST-X2

(e) ST-X1 (f) ST-X3

(g) JAZZ-G3 (h) NXP-G3

Figure 5.9: Gummel plots for base currents at 300 K, 200 K, 77 K, 50 K, and 18 K

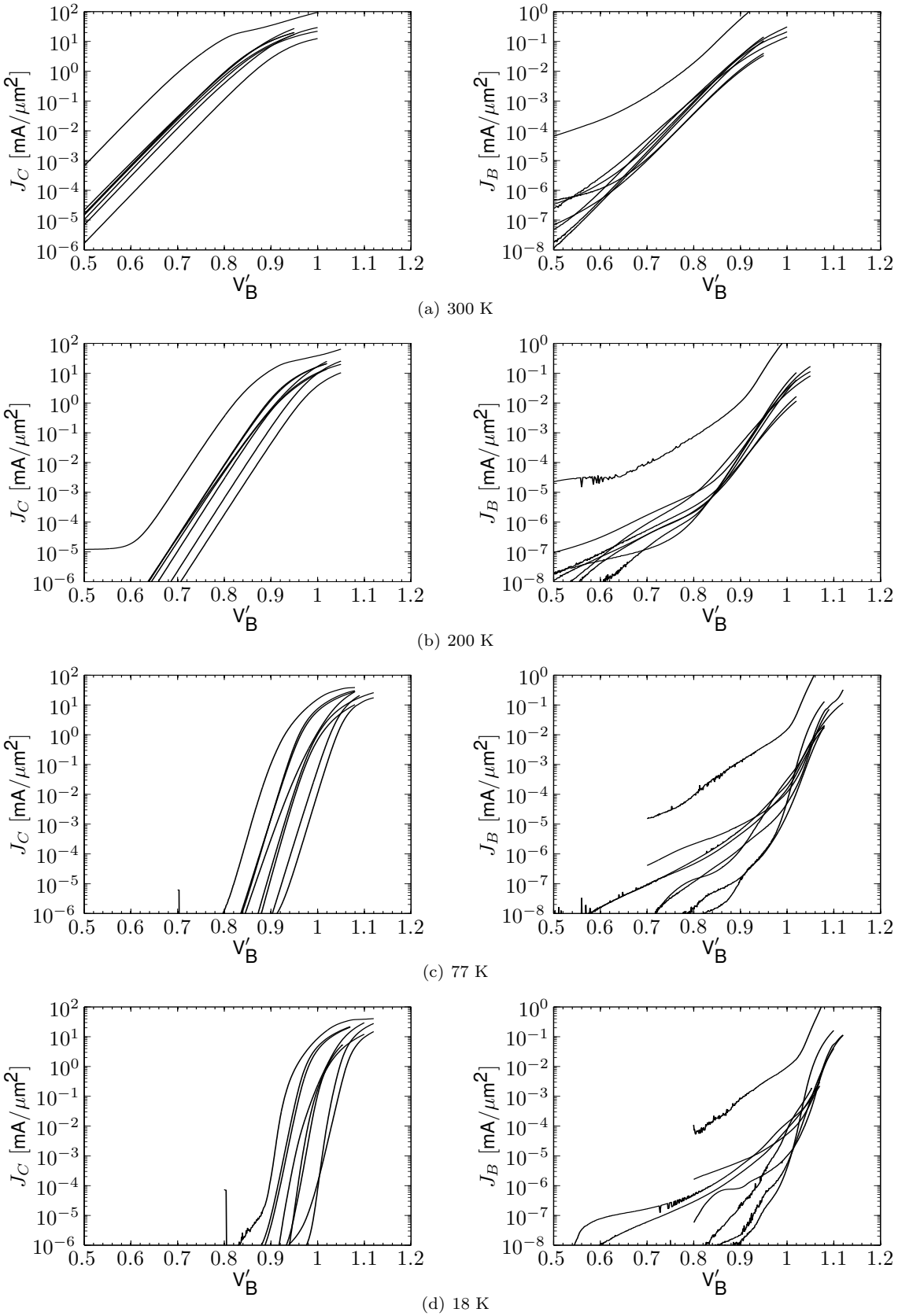


Figure 5.10: Collector (left) and base Gummel plots with data for all devices at (a) 300, (b) 200, (c) 77, and (d) 18 K

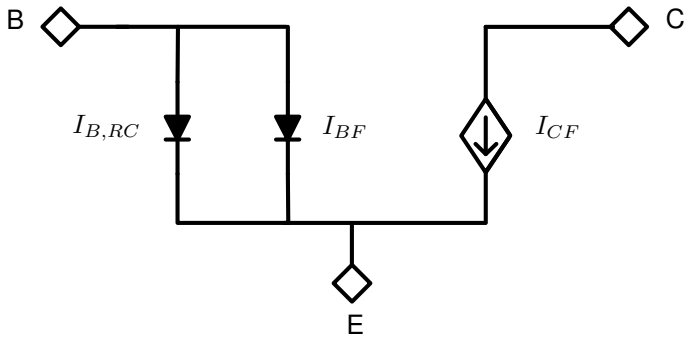


Figure 5.11: Equivalent large-signal circuit model in the low-injection regime

conditions. Similarly, the collector current under low-injection conditions can be modeled as a transport current source which has the mathematical form of a diode. Furthermore, in the case of a collector leakage current, an additional transport current source can be used in order to capture this effect. Thus, in general, four coefficients need to be determined in order to mathematically describe the collector current at low-injection. Following this analysis, it can be deduced that eight coefficients[13] are required to fully specify the terminal currents under low-injection: 1) J_{B0}, 2) n_b, 3) $J_{B0,RC}$, 4) $n_{b,rc}$, 5) J_{C0}, 6) $J_{C0,P}$, 7) n_{c0}, and 8) $n_{c,p}$.

Now, under low-injection conditions, the dc voltage drops across the access resistances are negligible. Thus, the equivalent large signal model in the absence of collector leakage currents simplifies to that shown in Fig. 5.11. Thus, as V_{BE} is unaffected by parasitic resistances in this regime, it is possible to determine the unknown coefficients by performing an optimization routine on the raw Gummel-curve data. This optimization was performed for each of the devices at 300, 200, 77, 50, and 18 K using a Matlab minimization routine. The full set of coefficients appears in Appendix C.

An example of typical results of the extraction procedure appears in Fig. 5.12. Clearly, the proposed model works well in capturing the ideal current components in the low-injection region. In the following sections, the change in the ideal coefficients, J_{B0}, J_{C0}, n_b, and n_{c0}, will be studied and compared with theory. To begin with, we will investigate the ideality factors, n_b and n_{c0}.

5.2.3 Low Injection Ideality Factors and Effective Temperatures

The base and collector current ideality factors are important figures of merit for study at cryogenic temperatures because they give a measure of the realized device performance in comparison

[13]In most cases, six coefficients are sufficient since the effect of the collector leakage current is usually negligible rendering the determination of $J_{C0,P}$ and $n_{c,p}$ unnecessary.

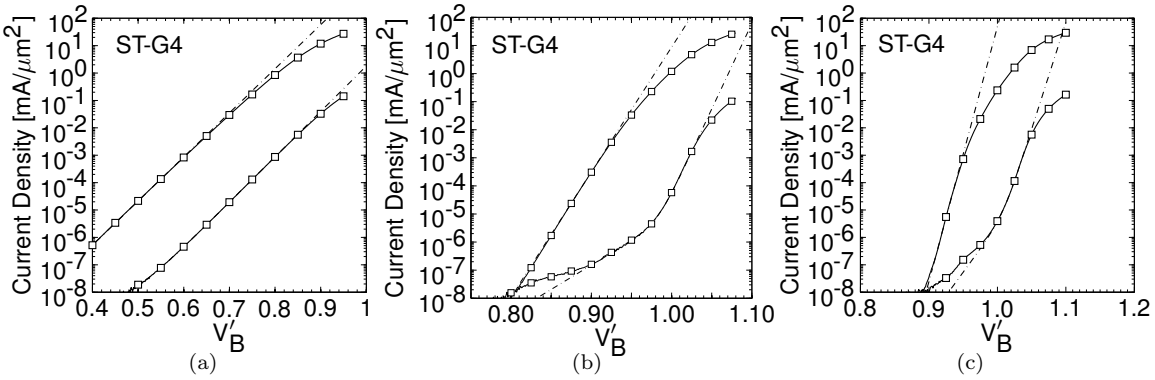


Figure 5.12: Examples of dc fit to low-injection Gummel data at (a) 300, (b) 77, and (c) 18 K

to the theoretical performance, as predicted by classical drift–diffusion theory. In particular, the ideality factors at low-injection are especially revealing because they are constant with respect to bias and thus cannot be of thermal origin. Using these coefficients, it is possible to define effective electron and hole temperatures, $T_{eff,n} = n_{c0} \times T_a$ and $T_{eff,p} = n_b \times T_a$, which are the equivalent temperatures which would result in the same low injection current slopes under the standard drift-diffusion formulation[14]. Thus, the temperature dependence of $T_{eff,n}$ and $T_{eff,p}$ provides physical information regarding the transport mechanisms and deserves serious investigation. As noted in Chapter 3, several researchers have reported effective electron temperatures that are higher than T_a for SiGe devices operating at cryogenic temperatures, which they have attributed to ballistic transport through the thin base [29, 30]. However, there is little information in the literature regarding the behavior of the effective hole temperature in SiGe devices operating at cryogenic temperatures.

The extracted low-injection ideality factors, n_b and n_{c0}, were converted to effective temperatures and are plotted against the ambient temperature in Fig. 5.13. When comparing these figures side by side, it is evident that the behavior of the collector current is more ideal than is that of the base current. In order to gain further insight into how $T_{eff,n}$ and $T_{eff,p}$ compare, all measured pairs are plotted against one another in Fig. 5.14. It is clear that the carriers of the base and electron currents act as if they are at different temperatures and that of the two currents, the collector current behaves as if it is at a lower temperature. This is puzzling as the non-equilibrium transport phenomena with which $T_{eff,n}$ is associated is attributed to ballistic transport through the thin base at low-temperatures, which can be modeled as a rise in the effective temperature of the electrons [29].

[14] i.e. if the device were behaving perfectly at T_{eff}, then the slope of the current would be what was measured at T_a. Thus, the electrons (holes) behave as if they are at $T_{eff,n(p)}$, even though the lattice temperature is T_a

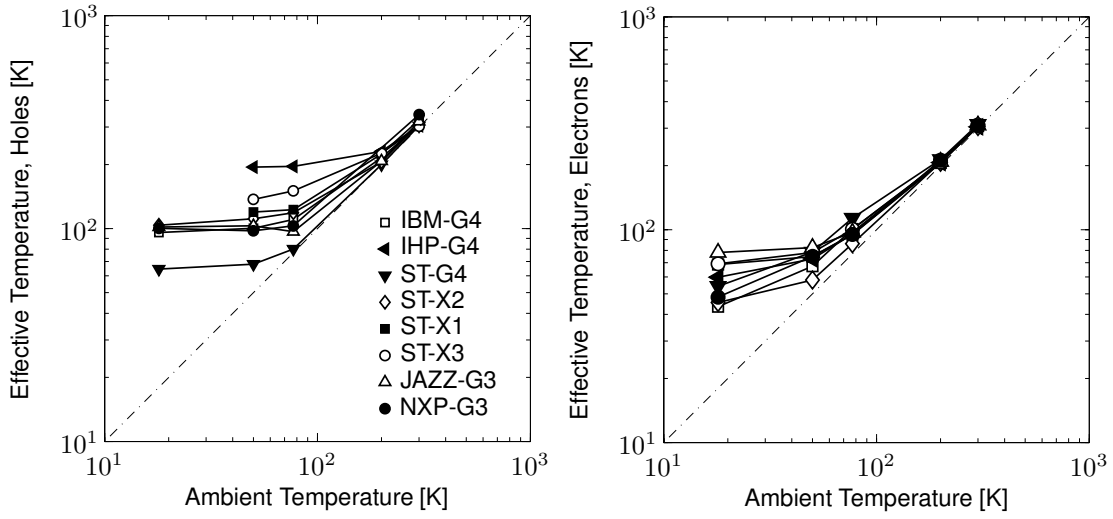


Figure 5.13: Effective temperature for (a) hole and (b) electron currents at low-injection. The hole current corresponds to the base current whereas the electron current corresponds to the collector current.

Thus, the fact that the $T_{eff,p}$ is consistently higher than $T_{eff,p}$ seems to contradict the notion that the rise in $T_{eff,n}$ is due to ballistic transport as the hole current cannot be ballistic. Nonetheless, it is possible that the physics behind $T_{eff,p}$ are consistent the current theory of non-equilibrium transport. This is an open issue and certainly deserves further study.

5.2.4 Saturation Currents

In the previous section, we saw that both the collector-current and base-current ideality factors change significantly at cryogenic temperatures. In this section, we will look at how the saturation current coefficients J_{C0} and J_{B0} change with cooling. To begin, we will examine J_{B0}, which, as we saw in Chapter 3, theoretically should change as

$$\frac{J_{B0}\left(T_a\right)}{J_{B0}\left(300\right)} \approx 7.83 \times 10^{18} \left(\frac{T_a}{300}\right)^{7/2} e^{-E_g(T_a)/kT_a}. \tag{5.12}$$

Now, as it has been determined that the holes act as if they are hotter than the ambient temperature, our expectation is that the fractional change in saturation current will be dictated by equation (5.12), with the variable T_a replaced by the effective hole temperature, $T_{eff,p}$. The ratio $J_{B0}\left(T_a\right)/J_{B0}\left(300\right)$

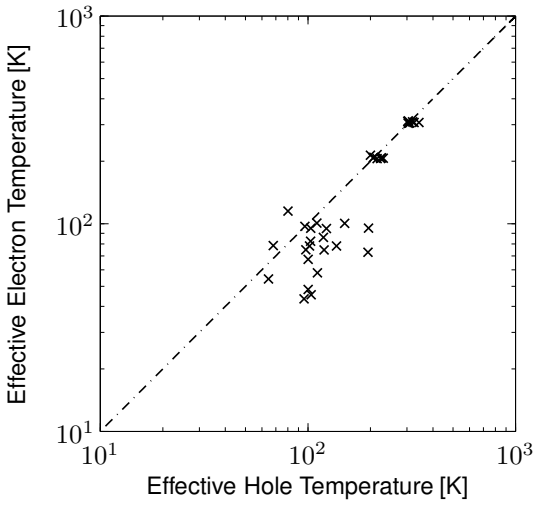


Figure 5.14: Scatter plot of $T_{eff,p}$ vs $T_{eff,n}$. For ambient temperatures of 77 K and below, $T_{eff,n}$ is consistently lower than $T_{eff,p}$ with the exception of two points.

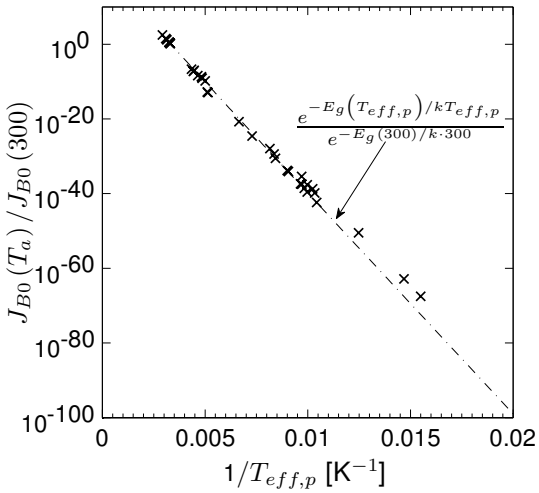


Figure 5.15: Normalized J_{B0} plotted against $1/T_{eff,p}$. The data from all eight of the devices is included in the plot.

is plotted in Fig. 5.15. Inspection of the data reveals that a better fit is

$$\frac{J_{B0}\left(T_a\right)}{J_{B0}\left(300\right)} \approx 7.83 \times 10^{18} e^{-E_g(T_{eff,p})/kT_{eff,p}}. \tag{5.13}$$

This trend line also appears in 5.15 and is clearly a very good fit to the data. Thus, there is a disagreement between the theory and measurement. While equations (5.12) and (5.13) differ by a factor of $(T_a/300)^{3/2}$, with the sheer scale of the problem a much more important issue is why the base current acts as if it is at $T_{eff,p}$ as opposed to T_a. Quasi-ballistic transport is generally used

to explain why the low-injection collector current ideality factor is significantly greater than unity at cryogenic temperatures, with the justification being that the base layer is extremely thin [29]. However, this rationale does not apply to holes being back injected into the emitter. Thus, the explanation as to why the holes behave as if they are at an elevated temperature deserves some study using a device simulator.

The collector current saturation currents were also determined as a function of temperature. However, as discussed in Chapter 3, there are multiple factors affecting the collector saturation current, which makes its temperature dependence difficult to tie to physical characteristics of the device without detailed information regarding the device structure. Thus, the temperature dependence of the collector saturation current will not be discussed here. However, for reference, the extracted values appear in Appendix C.

5.3 Summary

In this chapter, the dc operating characteristics of a variety of state-of-the-art SiGe HBTs have been studied systematically at temperatures ranging from 300 to 18 K. In the first half of the chapter, the theoretical low-frequency noise performance of the devices was investigated. As a result of the study, it was found that the low-frequency noise performance of SiGe HBTs is enhanced significantly with cooling to 18 K. Furthermore, it was revealed that the improvement in low-frequency noise performance at 18 K can be predicted using room temperature Gummel curves. In the second half of the chapter, the Gummel characteristics of the devices were studied. In this investigation, among other things, it was discovered that, even at low-injection, both the electrons and holes behave as if they are at elevated temperatures when the devices are cooled to cryogenic temperatures. In the next Chapter, the RF characteristics of the devices are investigated.

Chapter 6

Small-Signal Characterization and Modeling

An understanding of the RF small-signal performance of a device is a critical step towards the evaluation of a microwave transistor. While SiGe HBTs have been modeled extensively at 300 K, little research has been conducted on how their small-signal model parameters vary as the devices are cooled to cryogenic temperatures. Instead, RF studies have generally been focused on the high frequency figures of merit f_t and f_{max} [31, 32, 33, 34, 35]. Thus, a study of the small-signal model at cryogenic temperatures is warranted. In this chapter, the small-signal modeling of the devices described in Chapter 4 is presented. Of particular interest in this investigation is the relative change that occurs in the device parameters as the ambient temperature is swept from room temperature to deep cryogenic temperatures.

The chapter begins with a look into how f_t and f_{max} depend upon temperature. Next, a small-signal model extraction procedure is first presented and then applied to the modeling of the seven[1] devices described in Chapter 4. During the discussion of the extracted parameters, expectations of the changes anticipated with temperature are discussed and compared with the measured results. In addition, where possible, the differences between the various devices are linked to physical differences in the processes. All measurements in this section were made using the cryogenic probe station setup described in Chapter 4.

6.1 The Temperature Dependence of f_t and f_{max}

As a springboard to discussing the RF performance of SiGe HBTs at cryogenic temperatures, it is intuitive to look at what happens to the f_t and f_{max} of a SiGe HBT as the device is cooled.

[1]The NXP-G3 device was excluded from this study as RF de-embedding structures were not available.

Referring back to the expression for f_t, we have:

$$f_t = \frac{1}{2\pi}\left[\tau_e + \tau_b + \tau_{cbd} + \frac{kT_a}{qI_c}\left(C_{jbe} + C_{jcb}\right) + r_c C_{jcb}\right]^{-1}. \tag{6.1}$$

Therefore, one would expect an increase in f_t due to:

1) A drop in τ_e due to an increase in β_{DC},

2) A drop in kT_a/qI_c,

3) A drop in τ_{cbd} due to an increase in v_{sat}.

Furthermore, as f_{max} is proportional to $\sqrt{f_t}$, one would expect that an increase in f_{max} will occur with cooling as well.

The unity-current-gain cutoff-frequency, f_t can be determined by extrapolating the ac current gain,

$$h_{21} = \frac{Y_{21}}{Y_{11}}, \tag{6.2}$$

which is assumed to roll-off at 20 dB/decade when plotted versus f. Similarly, the maximum frequency of oscillation, f_{max} can be found by extrapolating the unilateral gain, defined as [118]

$$U = \frac{\left|Y_{21} - Y_{12}\right|^2}{4\left(\Re\left\{Y_{11}\right\}\Re\left\{Y_{22}\right\} - \Re\left\{Y_{12}\right\}\Re\left\{Y_{21}\right\}\right)}, \tag{6.3}$$

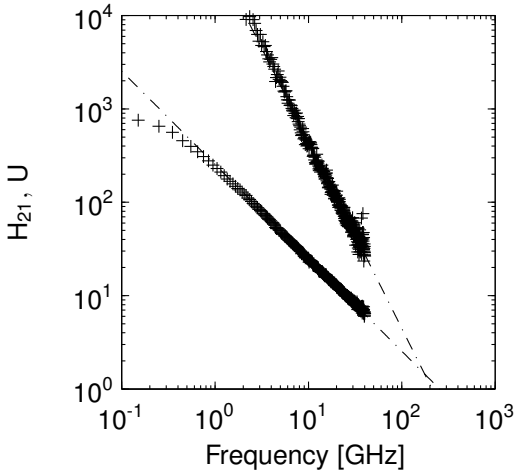


Figure 6.1: Sample f_t and f_{max} extraction data. The data which demonstrates sharper roll-off is the unilateral gain.

at 20 dB/decade when plotted against $\sqrt{f}$. In general, the measurement of f_t is more reliable than that of f_{max}. The reason for this is that the measurement of f_{max} is dependent upon the difference between two potentially small numbers, $\Re\{Y_{11}\}\Re\{Y_{22}\}$ and $\Re\{Y_{12}\}\Re\{Y_{21}\}$.. Nonetheless, if care is taken in choosing an appropriate frequency range over which to extrapolate the unilateral gain[2], reliable estimates of f_{max} can be made.

In order to estimate f_t and f_{max} the following procedure is used:

1) The open- and short-circuit de-embedding structures are measured.

2) S-Parameter measurements are made over a wide range of bias points from low currents to over 20 mA/μm^2.

3) The effects of the feed structure and pads are removed from the active bias measurements using the procedure summarized in Appendix B.

Typical extraction data appears in Fig. 6.1. It is apparent from the curves that the measured data follows the expected slopes. Using this procedure, the values of f_t and f_{max} were extracted for each of the seven devices at 18, 50, 77, 200, and 300 K and, at each temperature the peak value was located. The values of $f_{t,pk}$ and $f_{max,pk}$ as well as the associated collector current densities are given numerically in Appendix C and the ratio of the 18 K values to room temperature values is plotted for each device in Fig. 6.2. What is particularly interesting about these data is that, while there is significant variation in $f_{max,pk}$, the ratio of the cryogenic value of $f_{t,pk}$ to room temperature $f_{t,pk}$ is

[2]At higher frequencies, the measurement is more reliable.

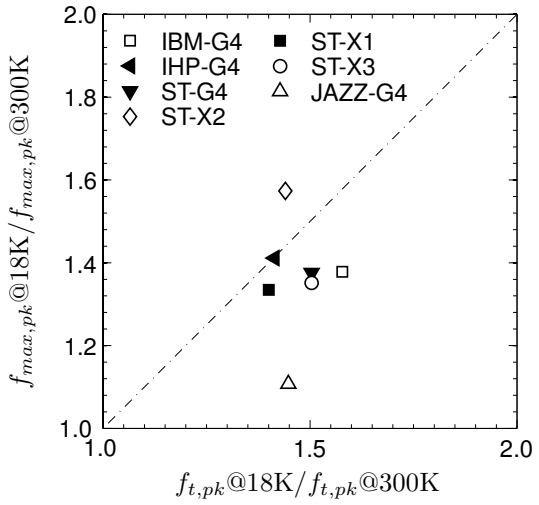
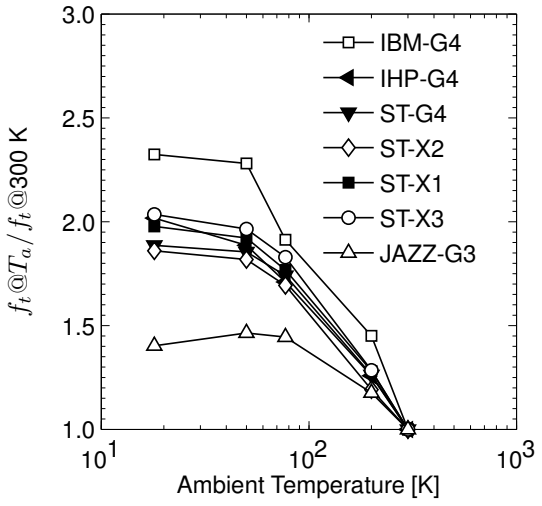


Figure 6.2: Ratio increase in $f_{t,pk}$ and $f_{max,kp}$ with cooling



(a)

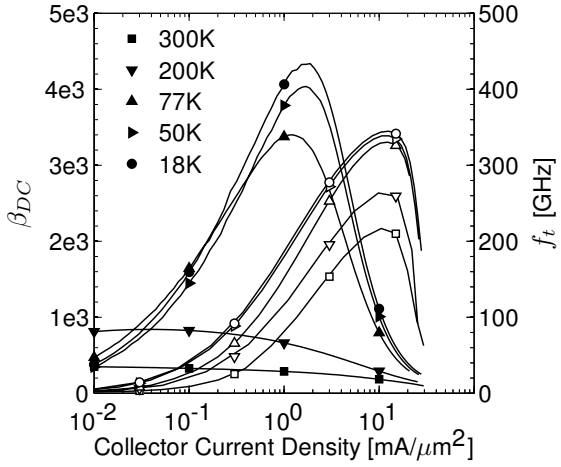
Figure 6.3: Ratio of increase in f_t vs temperature for $J_C = 1\text{mA}/\mu\text{m}^2$

approximately 1.5±0.1. Furthermore, this trend is in good agreement with recently reported values given in [31, 32, 33, 34, 35, 37]. Thus, it seems that cooling improves the value of peak f_t by about 50%, regardless of the device structure.

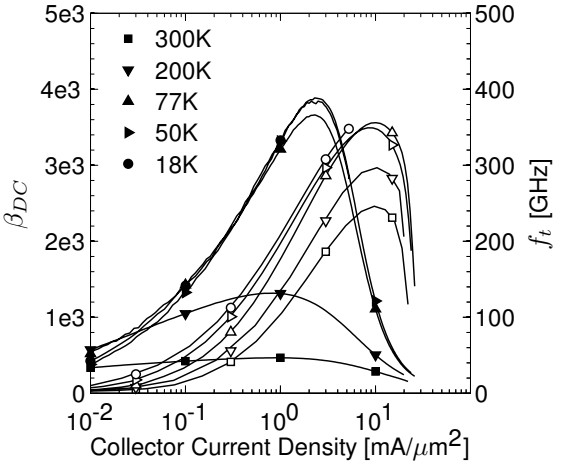
In addition to studying the increase in $f_{t,pk}$ and $f_{max,pk}$, it is also interesting to study the increase in f_t at a fixed collector current density. In Fig. 6.3, the ratio of $f_t\,(T_a)$ to $f_t\,(300)$ is shown as a function of temperature for a fixed collector current density of 1 $\text{mA}/\mu\text{m}^2$ for each of the seven devices. Inspection of these curves reveals two interesting features: 1) f_t increases monotonically for all of the devices with the exception of the JAZZ-G3 device, and 2) the increase in f_t@1 $\text{mA}/\mu\text{m}^2$ is substantially greater than the increase in $f_{t,pk}$, with the exception of the JAZZ-G3 device in which the increase is the same in both cases. It is believed that the reason for a stronger increase at lower biases is that the device is at a lower effective temperature, and hence the effect of cooling on the terminal characteristics is more pronounced.

In addition to the increase in f_t and f_{max} with cooling, it is important to examine how the collector current required for peak f_t and f_{max} changes with cooling. Referring to Appendix C, it can be seen that the locations of peak f_t and peak f_{max} are more or less concurrent. Therefore, it is reasonable to consider just the location of peak f_t and apply the conclusions to f_{max}. In Fig. 6.4, f_t and β_{DC} are plotted together in order to gauge if it is possible to achieve both high dc current gain and high f_t simultaneously. Inspection reveals several interesting features:

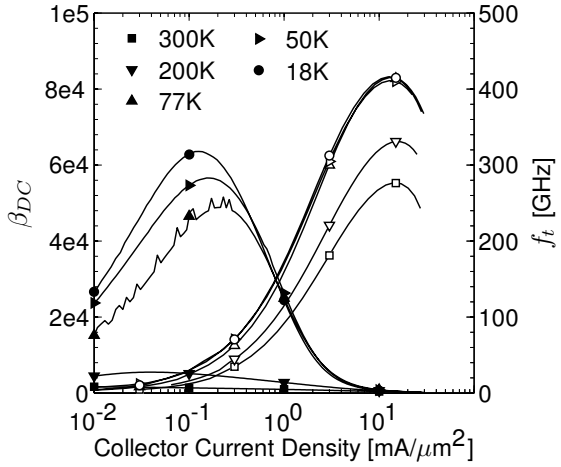
1) For each of the seven devices, and at all temperatures, $\beta_{DC,pk}$ occurs at a considerably lower



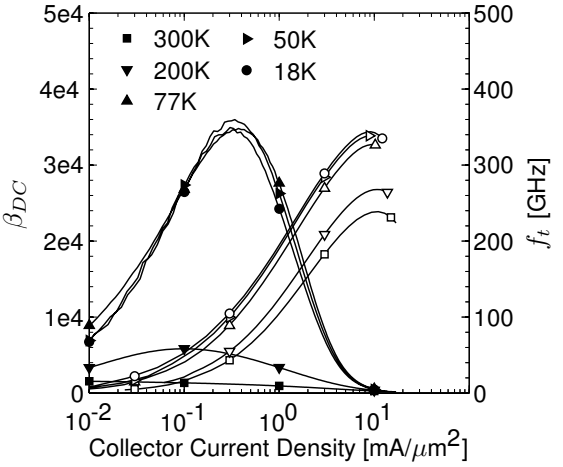
(a) IBM-G4



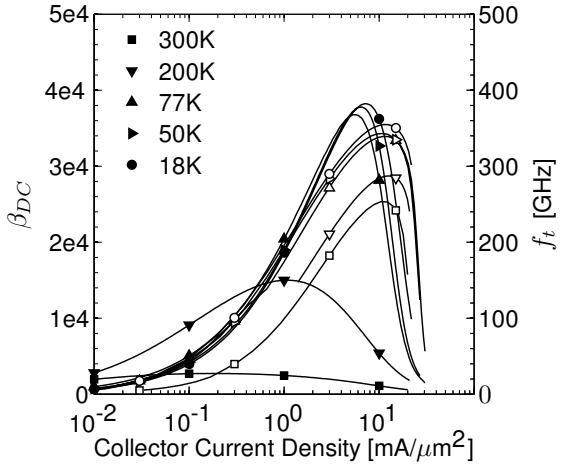
(b) IHP-G4



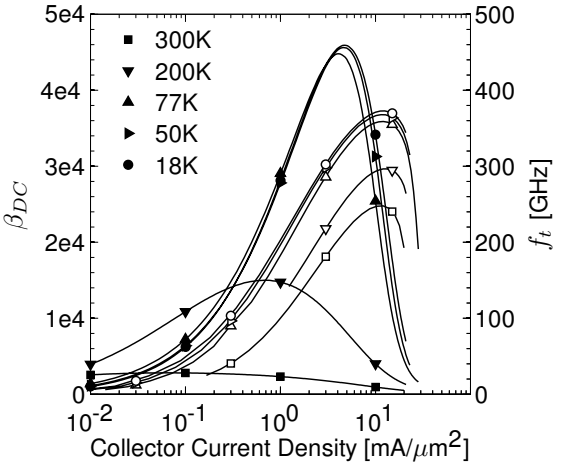
(c) ST-G4



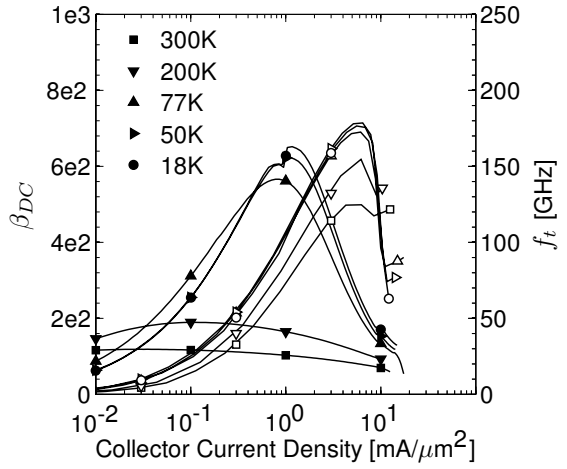
(d) ST-X2



(e) ST-X1



(f) ST-X3



(g) JAZZ-G3

Figure 6.4: f_t and β_{DC} versus J_C for each of the seven devices under study. The f_t curves are denoted by white markers whereas the β_{DC} curves are denoted by solid markers.

current density than $f_{t,pk}$.

2) The general shape of the f_t curves is independent of temperature whereas the shape of the β_{DC} curves depends strongly on temperature.

3) The collector current density required to achieve $f_{t,pk}$ does not change significantly as the devices are cooled.

4) The collector current density required to achieve $\beta_{DC,pk}$ increases as the devices are cooled.

As the low-frequency noise performance of a bipolar device depends primarily on the dc current gain and collector current ideality factor, and the high-frequency noise performance depends upon the value of f_t, the observations made above imply that there will be a tradeoff in the bias point based upon the desired frequency of operation; that is, as the frequency is increased, the bias point which optimizes the noise will also increase. This topic has profound implications in terms of the broadband noise performance of SiGe HBTs and will be revisited in Chapter 7.

6.2 Small-Signal Model Parameter Extraction Techniques

Design and simulation of microwave amplifiers requires accurate device models. Furthermore, prediction of the noise performance of HBTs is not possible without models that relate physically to the device structure (i.e., with accurate representations of physical resistances). While foundries typically supply these models for temperatures in the −55–125 C range [119], models are not available

at lower temperatures. In this section, techniques to extract a small-signal equivalent circuit model which accurately captures the physics of the device operation will be reviewed. In section 6.3.2, these small-signal modeling procedures are applied to several state-of-the-art devices, and a comparison of the performance of the modeled devices is presented.

The schematic diagram of the small-signal equivalent circuit of a SiGe HBT is shown in Fig. 6.5 along with the identification of sub-circuits[3] Y_I, Y_{II}, Z_{III}, and Y_{IV}. Several authors have proposed models for SiGe HBTs that are considerably more complex than the model shown here. For instance, it is common for r_b and C_{CB} to be broken up into intrinsic and extrinsic components [120, 121, 122]. However, for the large devices that are required for microwave LNAs that are optimized for 50 Ω generator impedance, the resistances in the circuit are small, which makes the extraction of base and collector capacitance splitting unreliable, leading to extraction parameters that are not unique and thus cease to reveal physical information regarding the device structure. Therefore, the simple model shown in Fig. 6.5 was chosen for this work as it is believed that unique component values can be determined.

The extraction of bipolar junction transistor small-signal model parameters is a subject that has received a great deal of attention in the literature over the years. Several authors have reported methods of determining the entire equivalent circuit model using RF measurements [120, 121, 122, 123, 124]. While determining all of the parameters using RF network measurements permits one to generate models which accurately capture the RF performance, it is diffucult to extract resistances which are on the order of an ohm in this manner. For measuring small resistances, it is believed that dc measurements provide much greater accuracy than high-frequency measurements, and methods to measure each of the resistances have been presented in the literature [67, 125, 126, 127, 128, 129, 130, 131].

In this work, a combination of dc and RF measurements are used to determine the component values. Inspection of the small model reveals that there are nine parameters to be determined: 1) r_b, 2) g_{be}, 3) C_{BE}, 4) C_{CB}, 5) r_c, 6) r_e, 7) g_m, 8) τ_d, and 9) C_{CS}. An overview of the extraction procedure is as follows:

1) Determine C_{CS} and C_{CB} using off-bias measurements, where $V_{BE} = 0$ V.

2) Determine r_e and r_c using dc measurements. If desired, these resistances can also be measured using RF techniques to ensure consistency. As r_e degenerates the transistor, it is particularly important that its value be accurately determined.

[3]The network parameters for the important sub-circuits is given in Appendix D.

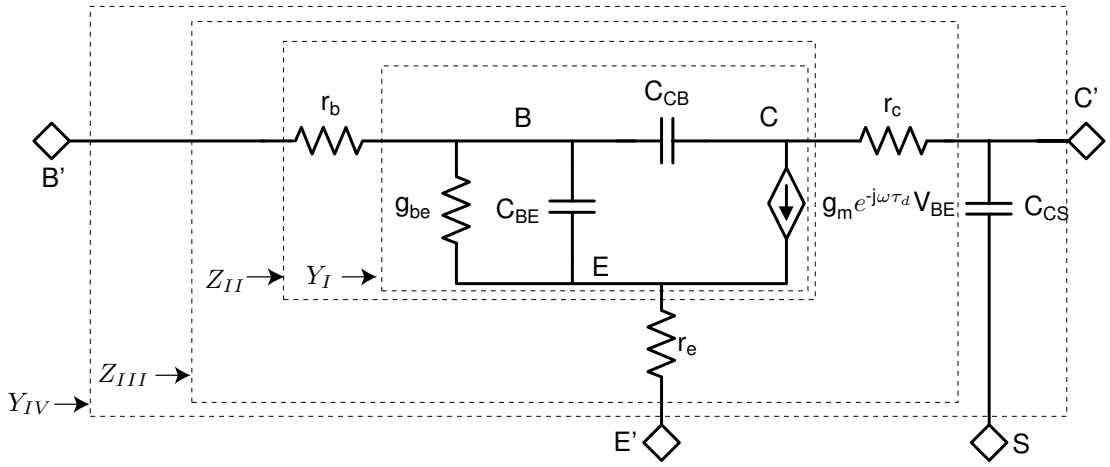


Figure 6.5: SiGe HBT small-signal model with sub-network blocks identified

3) De-embed C_{CS}, r_c, and r_e from the active-bias measurements to end up with the network labeled Z_{II} in Fig. 6.5.

4) Extract and de-embed r_b to end up with the intrinsic network (Y_I in Fig. 6.5).

5) Determine the remaining parameters, g_m, τ_d, g_{be}, and C_{BE}.

For all measurements, it is assumed that both the substrate and emitter terminals are tied to ground. Furthermore, for all active-bias measurements, V_{CB} =0 V.

Prior to processing S-parameter data, the effects of the bondpads and feedlines are removed using the procedure described in Appendix de-embedding. Once these effects have been removed, the reference plane is at the contacts to the device. This is important as it will allow the extracted models to be scaled and used with arbitrary wiring configurations. It is also important that any systematic resistances be removed from the dc data prior to processing. Thus, the series and return resistances are determined using the short circuit test structures.

6.2.1 Determination of C_{CS} and C_{CB}

The collector–base and collector–substrate capacitances are junction capacitances that can be determined using off-bias measurements, in which the collector voltage is swept while the base is held at 0V. Under this excitation, both g_{be} and g_m become infinitesimally small and can be neglected in the analysis. The resulting circuit appears in Fig. 6.6. The Y-parameters for the transistor in the

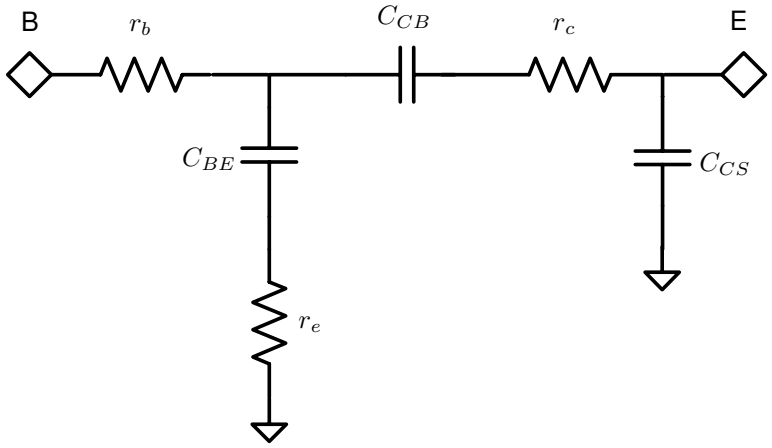


Figure 6.6: Equivalent circuit for SiGe HBT in the off-bias state

off-bias state can thus be written as:

$$Y_{off} = \begin{bmatrix} \frac{j\omega C_{CB}(1+j\omega\tau_{BE})+j\omega C_{BE}(1+j\omega\tau_{CB})}{\Delta} & \frac{-j\omega C_{CB}(1+j\omega\tau_{BE})}{\Delta} \\ \frac{-j\omega C_{CB}(1+j\omega\tau_{BE})}{\Delta} & \frac{j\omega C_{CB}(1+j\omega(r_b+r_e)C_{BE})}{\Delta} + j\omega C_{CS} \end{bmatrix}, \tag{6.4}$$

where

$$\tau_{BE} = r_e C_{BE}, \tag{6.5}$$

$$\tau_{CB} = r_c C_{CB}, \tag{6.6}$$

and

$$\Delta = 1 + j\omega\left(\tau_{BE}\left(1 + r_b/r_e\right) + \tau_{CB}\left(1 + r_b/r_c\right)\right) - \omega^2\tau_{BE}\tau_{CB}\left(1 + r_b/r_e + r_b/r_c\right). \tag{6.7}$$

Inspection of (6.4) allows the determination of C_{CB} and C_{CS} as:

$$C_{CB} = -\lim_{\omega\to 0}\frac{\Im\left\{Y_{12,off}\right\}}{\omega} \tag{6.8}$$

and

$$C_{CS} = \lim_{\omega\to 0}\frac{\Im\left\{Y_{12,off} + Y_{22,off}\right\}}{\omega}. \tag{6.9}$$

Thus, C_{CB} and C_{CS} can be found for a single collector voltage by determining the y-intercept of equations (6.8) and (6.9) plotted as a function of ω. Example data used to determine the collector-base and collector–substrate capacitances appear in Fig. 6.7 along with the linear fits used to determine the y-intercepts. Once the capacitances have been extracted over a wide range of collector voltages,

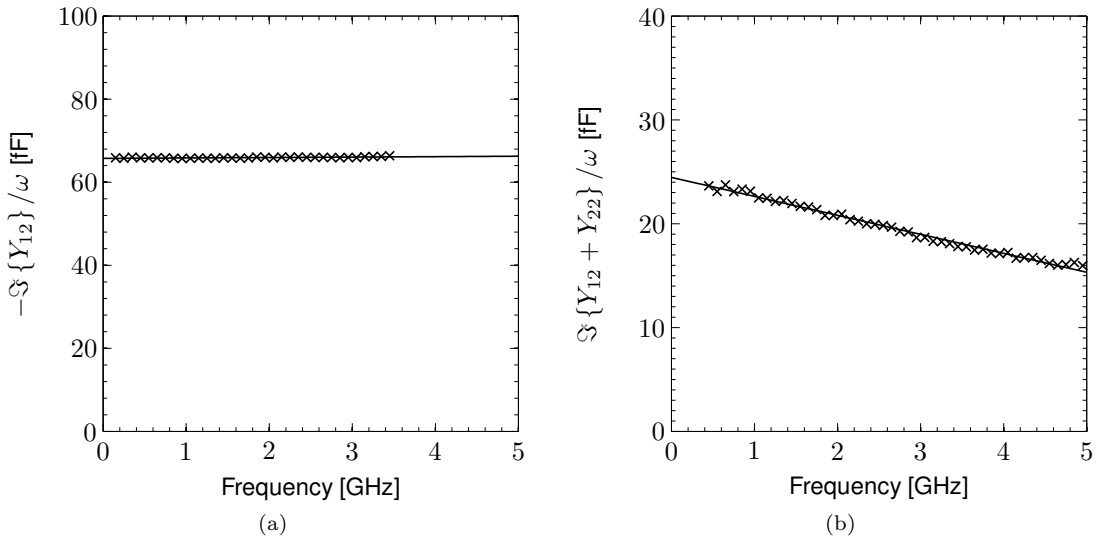


Figure 6.7: Example of data used for extraction of (a) C_{CB} and (b) C_{CS}. This particular data was taken for a ST BiP-X3 device at 300 K physical temperature.

it is possible to fit smooth curves to the data as the bias dependence of depletion capacitances is well known [132]:

$$C_{CB}\left(V_{CB}\right) = \frac{C_{CB0}}{\left(1 + V_{CB}/V_{CB0}\right)^{m_{cb}}} \tag{6.10}$$

and

$$C_{CS}\left(V_{CS}\right) = \frac{C_{CS0}}{\left(1 + V_{CS}/V_{CS0}\right)^{m_{cs}}}. \tag{6.11}$$

In Fig. 6.8, some example curves are shown demonstrating the effectiveness of the equations (6.10) and (6.11) at representing the voltage dependence of the capacitances. Fitting the data to this form is useful in that it allows the capacitances to be represented in a very compact form that can easily be ported to a simulator.

Once the collector–substrate capacitance is known, it can easily be de-embedded[4], resulting in the equivalent circuit marked Z_{III} in Fig. 6.5. The next step is to determine the emitter and collector resistances.

6.2.2 Determination of r_e

As the emitter resistance degenerates the transistor, its accurate determination is essential in determining an HBT model that has physical ties to the device structure. Therefore, a variety of different techniques are used to determine r_e, and the results are compared in order to check for

[4]See Appendix B for information on de-embedding procedures.

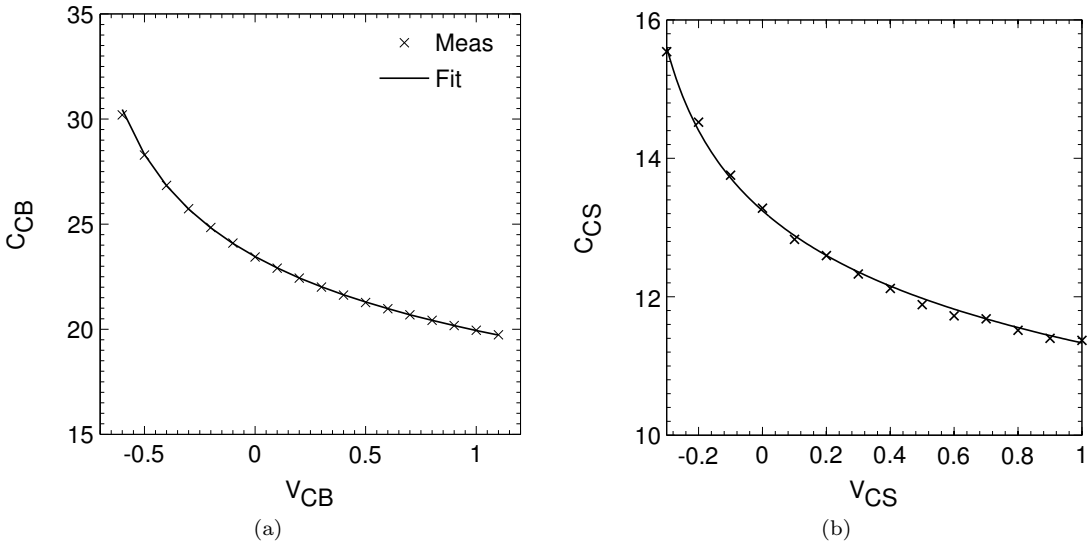


Figure 6.8: Example fits to (a) C_{CB} and (b) C_{CS} data

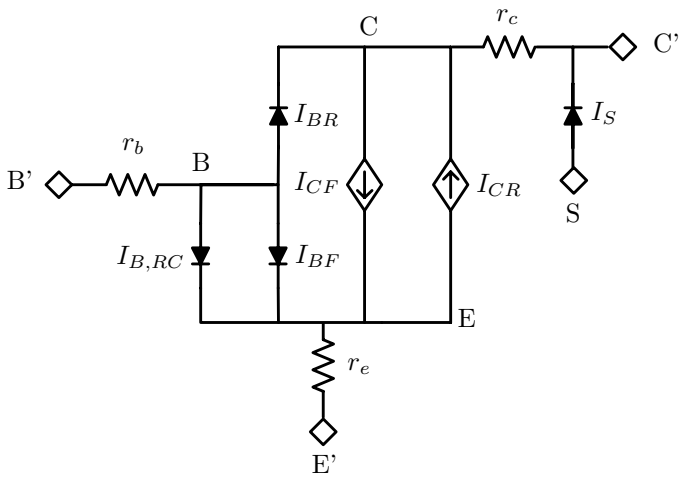


Figure 6.9: Equivalent circuit diagrams used for open-collector and gummel parameter extraction methods

consistency. Namely, the emitter resistance is extracted using dc forward-gummel measurements, dc open-collector measurements, and active-bias RF S-parameter measurements. Each measurement method is discussed below.

6.2.2.1 r_e Extraction: Forward-Gummel Method

The dc emitter current of a SiGe HBT operating in the forward-active regime is given as

$$I_E \approx I_{B0}\left(e^{V_{BE}/n_bV_T}-1\right)+I_{C0}\left(e^{V_{BE}/n_cV_T}\right)\approx I_{C0}\left(e^{V_{BE}/n_cV_T}\right), \tag{6.12}$$

where $V_{BE} \approx V_B - I_E r_e$. Thus, an expression for the applied base voltage can be written as

$$V_B \approx I_E r_e + n_c V_T \ln\{I_E/I_{C0}\}. \tag{6.13}$$

Differentiating equation (6.13) with respect to the emitter current, we arrive at the following expression:

$$\frac{\partial V_B}{\partial I_E} \approx r_e + \frac{n_c V_T}{I_E}. \tag{6.14}$$

6.2.2.2 r_e Extraction: Open-Collector Method

An alternative way in which the emitter resistance can be extracted is to use the open-collector method as suggested by Rudolph in [67]. A block diagram of the test setup used for this measurement appears in Fig. 6.11. The base terminal is driven with a swept current source and the collector voltage is measured while the collector current is forced to be zero using a second current source. Under these conditions, the device is in saturation as both pn junctions are forward biased. Thus, the collector and base currents can be approximated as [67] Therefore, the emitter resistance can be found by plotting $\partial V_B/\partial I_E$ as a function of $1/I_E$ and extrapolating to determine the value of the y-intercept [133]. An example plot demonstrating this technique appears in Fig. 6.10(a). It should be noted that in the derivation it has been assumed that β_{DC} is large, and that both I_{C0} and n_c are independent of bias. These assumptions lose validity under high-injection conditions (>10mA/μm^2 in modern SiGe HBTs), so care must be taken to exclude data taken under operation in the high-injection regime.

$$I_C \approx I_{C0}e^{V_{BE}/n_cV_T} - I_{BR0}e^{V_{BC}/n_{br}V_T} \tag{6.15}$$

and

$$I_B \approx I_{BR0}e^{V_{CB}/n_{br}V_T}, \tag{6.16}$$

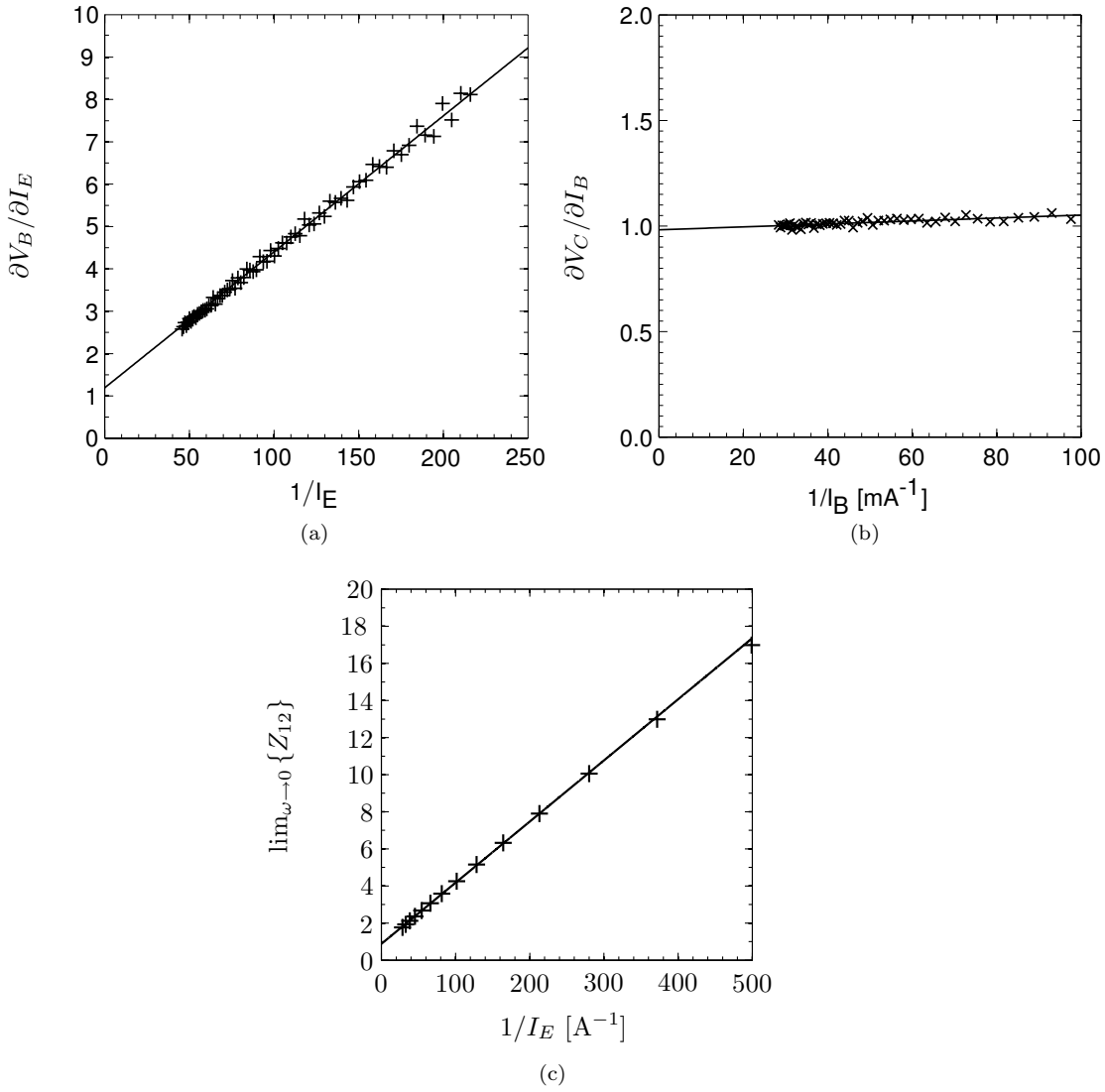


Figure 6.10: Sample extraction data using (a) the gummel method, (b) the open collector method, and (c) the RF active bias method. These data are for a 0.13x9.86μm^2 ST9MW device at 300K and the extracted values of r_e are 1.054 Ω for the gummel measurement, 0.982 Ω for the open-collector measurement and 0.887 Ω for the AC s-parameter measurement.

where I_{BR0} and n_{br} are the saturation coefficient and ideality factor associated with current flow through the base–collector junction. Thus, the collector voltage can be approximated as

$$V_C \approx I_B r_e + V_T \ln\left\{\left(\frac{I_B}{I_{C0}}\right)^{n_c}\left(\frac{I_{BR0}}{I_B}\right)^{n_{br}}\right\}. \tag{6.17}$$

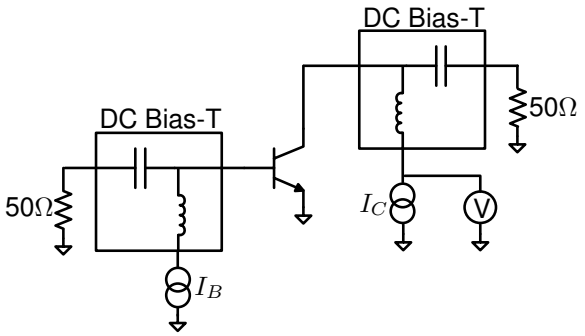


Figure 6.11: Test setup used for open-collector measurements

Finally taking a derivate of V_C with respect to I_B, we arrive at a linear polynomial function of $1/I_B$:

$$\left.\frac{\partial V_C}{\partial I_B}\right|_{I_C=0} \approx r_e + \frac{V_T}{I_B}\left(n_c - n_{br}\right). \tag{6.18}$$

Thus, the emitter resistance can be determined by fitting a line to $\partial V_C/\partial I_B$ as a function of $1/I_B$. An example plot demonstrating the effectiveness of this method appears in Fig. 6.10(b).

6.2.2.3 r_e Extraction: Active-Bias RF S-Parameter Method

After removal of the collector–substrate capacitance, the reverse transimpedance of an HBT under active bias is given as

$$Z_{12} = r_e + \frac{1}{\widetilde{g_m} + g_{be} + j\omega C_{BE}}, \tag{6.19}$$

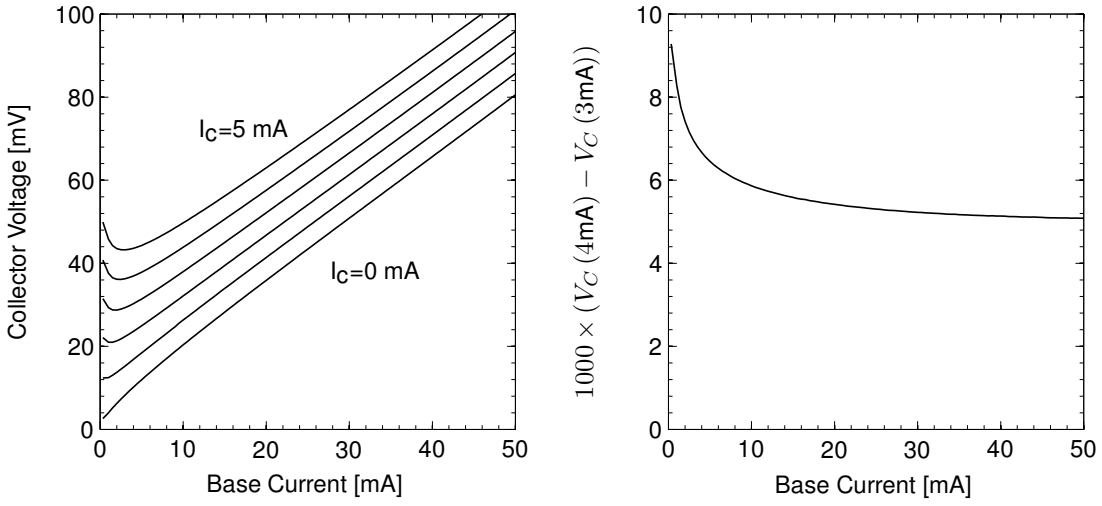
where $\widetilde{g_m} = g_m e^{-j\omega\tau_d}$. In the limit as $\omega \to 0$, the real part of equation (6.19) simplifies to

$$\lim_{\omega\to 0} \Re\{Z_{12}\} = \frac{1 + r_e\left(g_m + g_{be}\right)}{g_m + g_{be}} = \frac{1}{g_m + g_{be}} + r_e. \tag{6.20}$$

Finally, since both g_m and g_{be} increase monotonically with emitter current,

$$r_e = \lim_{I_E\to\infty}\left(\lim_{\omega\to 0} \Re\{Z_{12}\}\right). \tag{6.21}$$

Therefore, r_e can be found by first extrapolating the y-intercept of $Re\{Z_{12}\}$ versus frequency and then fitting a first order polynomial to the extrapolated intercepts as a function of $1/I_E$ [134]. An example plot showing the fitting process for emitter resistance determination using the AC Z-parameter method appears in Fig. 6.10(c).



(a) (b)

Figure 6.12: Example of r_c extraction data. (a) The collector voltage plotted as a function of base current for forced collector currents of 0–5 mA. The collector current step size is 1 mA. (b) The increase in collector voltage plotted as a function of base current. The curve represents the sum of r_c and r_e. Nearly identical curves are obtained with other values of collector current.

6.2.3 Determination of r_c

The collector resistance can be extracted using the dc open-collector method described by Rudolph in [67]. In this method, both the collector and base terminals are driven with a current source. If the device is in saturation, then the collector voltage can be written as

$$V_C \approx I_C\,(r_c + r_e) + I_B\,(r_e) + V_T \ln\left\{\left(\frac{I_C + I_B}{I_{C0}}\right)^{n_c}\left(\frac{I_{BR0}}{I_B}\right)^{n_{br}}\right\} \tag{6.22}$$

Now, taking the derivative of equation (6.22) with respect to I_C, we arrive at the following equation:

$$\frac{\partial V_C}{\partial I_C} \approx r_b + r_e + \frac{n_c V_T}{I_B + I_C}. \tag{6.23}$$

Thus the collector resistance can be determined as

$$r_c \approx \frac{\Delta V_C}{\Delta I_C} - r_e. \tag{6.24}$$

Example data obtained using this procedure appears in Fig. 6.12.

6.2.4 Determination of r_b

Extraction of base resistance is a complicated subject and has received considerable attention in the literature [121, 123, 127, 135]. Recently, Raya et al. have proposed the fabrication of special test structures to allow for direct measurement of the base resistance [127]. However, in this work, the goal is to evaluate commercial devices and special test structures are not available. Therefore, in situ extraction techniques are required.

Unlike the emitter and collector resistances, it is unreasonable to assume that r_b is independent of bias[5] due to its distributed nature across the junction [136]. Therefore, extraction of the base resistance should be carried out at each active bias point. To determine the base resistance at a given bias, we begin by de-embedding the collector to substrate capacitance as well as the emitter and collector resistances from the measured active-bias data to end up with the network parameters of a circuit that can be represented by the block labeled $Z_{\rm II}$ in Fig. 6.5. At this point, the base resistance can be computed from the Z-parameters as

$$r_b = \Re\left\{Z_{11}^{\rm II} - Z_{12}^{\rm II}\right\}. \tag{6.25}$$

If there are errors of ϵ_{re} and ϵ_{Ccs} associated with the extraction of the emitter resistance and collector–substrate capacitance, then it can be shown that the resulting error in the estimate of the base resistance due to the use of equation (6.25) is given as

$$\epsilon_{rb} \approx \frac{\epsilon_{Ccs}}{C_{CB}}\left(\frac{1}{g_m} + \epsilon_{re}\right) \approx \frac{\epsilon_{Ccs}}{g_m C_{CB}}. \tag{6.26}$$

As there will certainly be small errors in C_{CS}, there will always be a finite error proportional to ϵ_{Ccs}/C_{CB}. For the fairly large devices being investigated in this study, the base resistance and emitter resistances are expected to be on the order of 1 Ω. If we take the numerical example of $\epsilon_{Ccs}/C_{CB} = 0.01$, then the error in extracting a base resistance of 1 Ω is less that 10% only when the transconductance is greater than 100 mS[6]. Thus, the computation of the base-resistance using the Z-parameter method is particularly error-prone at low biases, where $1/g_m$ is large. However, at high-biases this method should offer good accuracy.

An alternative base-resistance extraction method is to make the computation using $H_{11}^{\rm II} = 1/Y_{11}^{\rm II}$.

[5] For completeness, it should be noted that the intrinsic collector resistance is actually bias dependent. However, in this work, it is assumed that the collector resistance is dominated by the resistances of the buried sub-collector and collector reach-through implants. Thus the bias dependence of r_c is ignored.

[6] To jump ahead, the values of C_{CB} are on the order of five times larger than those of C_{CS}. Thus, in this numerical example, we have assumed that there is a 5% error in the extraction of C_{CS}.

Because Y_{11}^{II} is determined with a short circuit on the collector terminal, this method is insensitive to errors in the extraction of C_{CS}. Explicitly, H_{11}^{II} can be written as

$$H_{11}^{\mathrm{II}} = \frac{1}{Y_{11}^{\mathrm{II}}} = R_b + \frac{g_{be} - j\omega\left(C_{CB} + C_{BE}\right)}{g_{be}^2 + \omega^2\left(C_{CB} + C_{BE}\right)^2}. \tag{6.27}$$

Thus the base resistance can be determined as

$$r_b = \lim_{\omega \to \infty} \Re\left\{\frac{1}{Y_{11}}\right\}. \tag{6.28}$$

When equation (6.28) is used to extract r_b, an error term arises due to any error in the extraction and de-embedding of r_c and r_e. In the limit as the frequency becomes infinite this error reduces to the parallel combination of the error in extraction of these two resistances.

6.2.5 Determination of g_m, τ_d, and g_{be}

After removing the parasitic resistances along with the collector–substrate capacitance, the transconductance and its associated delay time can easily be determined from the Y-parameters of the intrinsic circuit using the method provided in [123]:

$$g_m e^{-j\omega\tau_d} = Y_{21}^{\mathrm{I}} - Y_{12}^{\mathrm{I}}, \tag{6.29}$$

so

$$g_m = \left|Y_{21}^{\mathrm{I}} - Y_{12}^{\mathrm{I}}\right| \tag{6.30}$$

and

$$\tau_d = -\frac{\angle\left(Y_{21}^{\mathrm{I}} - Y_{12}^{\mathrm{I}}\right)}{\omega}. \tag{6.31}$$

Alternatively, g_m can be determined using the dc Gummel data as

$$g_m = \frac{\partial I_C}{\partial\left(V_B - r_e I_E\right)}. \tag{6.32}$$

It should be noted that determination of g_m through dc measurements is error-prone at high biases due to parasitic ground return resistances. Therefore, it is recommended that g_m be determined using RF measurements as this is believed to be more reliable. On the other hand, it is believed

that g_{be} is best computed using dc measurements as

$$g_{be} = \frac{g_m}{\beta_{AC}} = \frac{\partial I_B}{\partial V_{BE}}. \tag{6.33}$$

At high biases, the method presented in [8] can also be used to determine g_{be}.

6.2.6 Determination of C_{BE}

The final parameter which must be extracted is the base–emitter capacitance, which can be extracted from the intrinsic network as

$$C_{BE} \approx -\frac{1}{\omega \Im\left\{1/Y_{11}^{\mathrm{I}}\right\}} - C_{CB} \tag{6.34}$$

or

$$C_{BE} \approx \Im\left\{Y_{11}^{\mathrm{I}} + Y_{12}^{\mathrm{I}}\right\}. \tag{6.35}$$

Once the base–emitter capacitance is known as a function of bias, it is possible to fit a smooth curve to the extracted values. As it turns out, the base–emitter capacitance can be modeled mathematically as a depletion capacitor in parallel with a fixed capacitor [134]:

$$C_{BE} = C_{BE0} + \frac{C_{BE1}}{\left(1 - V_{BE}/V_{BE0}\right)^{m_{be}}}. \tag{6.36}$$

With a method for the determination of C_{BE} in place, a procedure for the extraction of each of the component values is known. In the next section, this procedure will be applied to several state-of-the-art HBTs.

6.3 Small-Signal Modeling Results

In the previous section, the techniques used to extract a complete small-signal model were presented. These techniques have been applied to each of the seven devices and small-signal models were extracted at 300, 200, 77, 50, and 18 K. In order to verify that the extraction procedure produces sufficiently accurate models, a comparison of typical extraction results with measured S-parameters is given at the beginning of the section. The small-signal models are then analyzed in terms of their temperature dependence and uniformity across the different technologies. In order to make a comparison as to how the different technologies change with temperature, the parameters

which depend upon bias will be compared at both a fixed current density, which is interesting from the device physics point of view, and a fixed transconductance, which is interesting from the circuit design point of view.

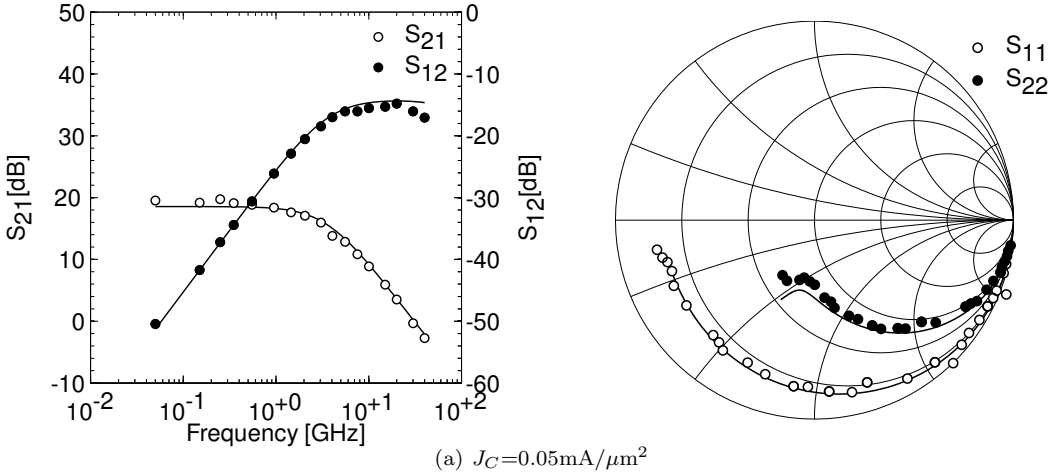
An example of some of the typical results that can be achieved when using the modeling procedure presented in this section is shown in Fig. 6.13. This measurement was taken at 18 K and the bias point in the presented data varies by a factor greater than 100:1. Clearly the agreement between measurement and modeling is excellent. Thus, we will proceed to discuss the behavior of each of the small-signal component values as a function of the ambient temperature.

6.3.1 Capacitances

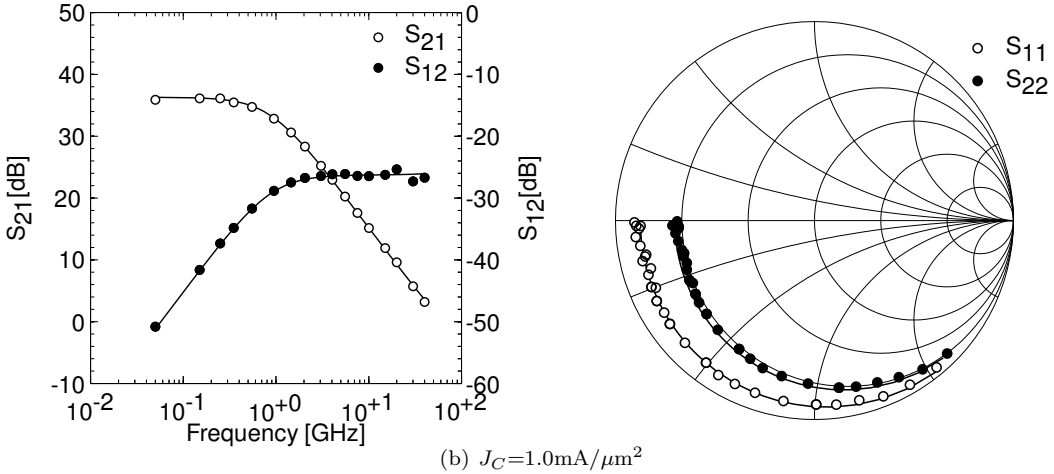
The capacitances were extracted as a function of bias, and smooth curves were fit to the data. The base–collector and collector–substrate capacitors were modeled as junction capacitances with unknown parameters C_{j0}, V_{j0}, and m_j, whereas the base–emitter capacitance was modeled as the parallel combination of a junction capacitance and a bias independent capacitor. Fits were performed at each temperature and all of the coefficients can be found in Appendix C.3.1. As the capacitances play a key role in the high-frequency performance of SiGe HBTs, their temperature dependencies are of great interest to this study. In the following, each of the three capacitances will be investigated as a function of temperature.

6.3.1.1 Collector-Base Capacitance

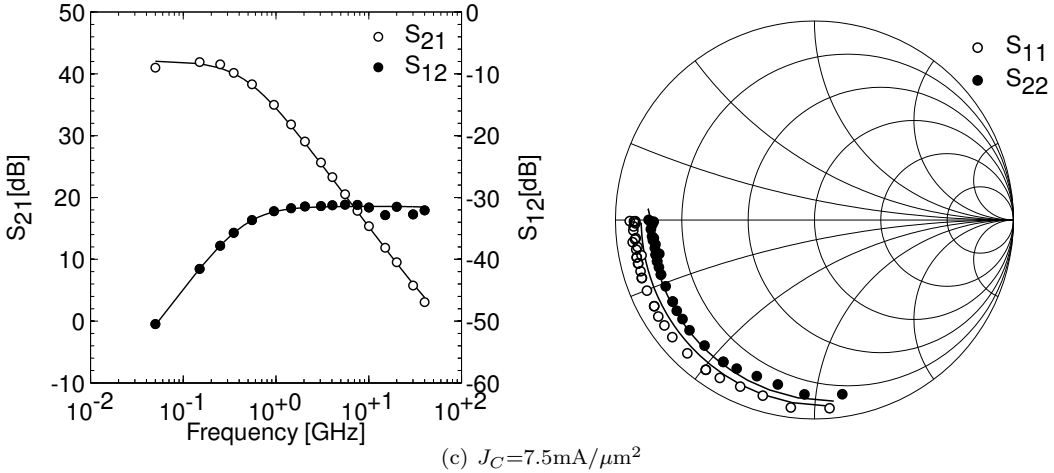
Physically, C_{CB} is a junction capacitance across the collector-base space-charge region (SCR). When the device is cooled, we expect this region to grow a little wider due to a decrease in the concentration of ionized dopants in the collector region. Thus, the value of C_{CB} should drop proportionally to the change in the collector-base SCR thickness [37]. The extracted capacitances are shown in Fig. 6.14(a) for a fixed base–collector voltage of 0 V. It is interesting to note that the device with the lowest base–collector capacitance per unit area (JAZZ-G3) is also the device which demonstrates the poorest performance in terms of f_t and f_{max}, whereas the device with the highest base–collector capacitance per unit area (ST-G4) is also the best device in terms of high frequency performance. This is an indication that the collector of the ST-G4 device is doped much more heavily than the JAZZ device; presumably this was done in order to delay the onset of the Kirk effect. This interpretation is supported by the fact that the collector current density required to achieve peak f_t and f_{max} is consistently higher for the ST-G4 device than any of the other devices whereas,



(a) J_C=0.05mA/μm^2

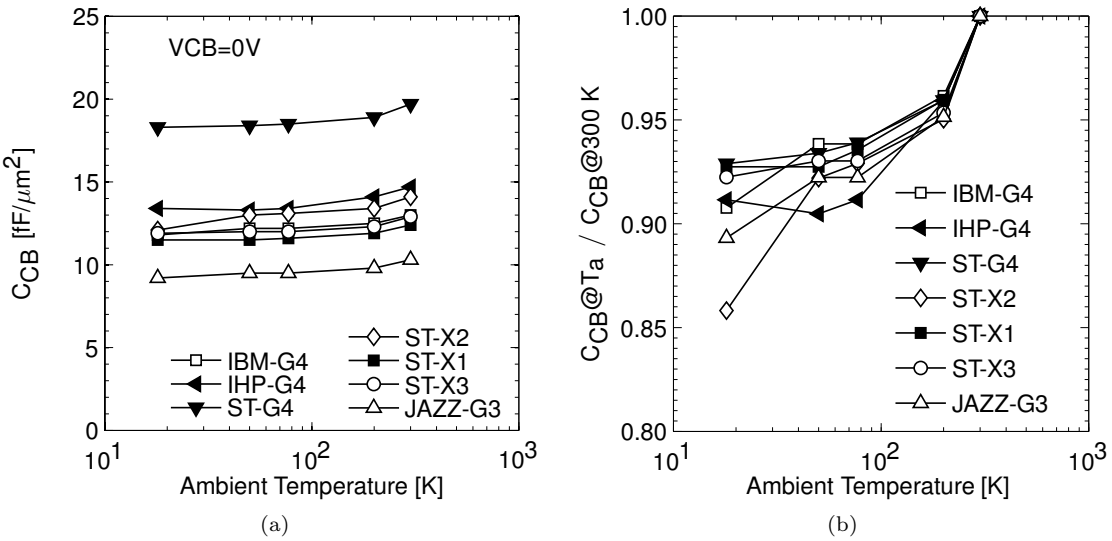


(b) J_C=1.0mA/μm^2



(c) J_C=7.5mA/μm^2

Figure 6.13: Example extraction results. The data is for a 3x0.17x14.79μm^2 ST-X3 device at 18 K. Excellent agreement is seen for all four scattering parameters over this very wide range of bias currents ($> 100:1$).



(a) (b)

Figure 6.14: (a) C_{CB} and (b) normalized C_{CB} as a function of ambient temperature for each of the devices

for the JAZZ device, it is consistently lower than the other devices.

The base–collector capacitances normalized to their room temperature values are shown as a function of ambient temperature in Fig. 6.14(b). As expected, a decrease in C_{CB} was observed for each of the seven devices as the temperature was lowered. Quantitatively speaking, the average device showed about a 10% decrease in capacitance as the ambient temperature was lowered from 300 to 18 K. The standard deviation in the measurements was 2.5%, indicating that, in general, the base–collector capacitance will drop. Furthermore, with cooling to 77 K, an average decrease in base–collector capacitance of 7.5% was observed with a standard deviation of 1%. These numbers are consistent with the changes observed with cooling to 77 K by Pruvost et al. [37].

6.3.1.2 Collector–Substrate Capacitance

As a direct consequence of the fact that the devices lie on a lightly doped substrate, a collector–substrate capacitance exists. As discussed in Section 3.1.5, semiconductor materials which are doped below the Mott-transition will show carrier freeze-out effects at temperatures below 30 K or so. This incomplete ionization will result in a drop in the depletion capacitance at cryogenic temperatures as $C_{CS} \propto \sqrt{N^-_{A,sub}}$. Furthermore, it will result in the conductivity of the substrate dropping by orders of magnitude, rendering the substrate terminal of C_{CS} essentially open-circuited. Thus, it is expected that the measured collector–substrate capacitance should drop enormously at cryogenic temperatures.

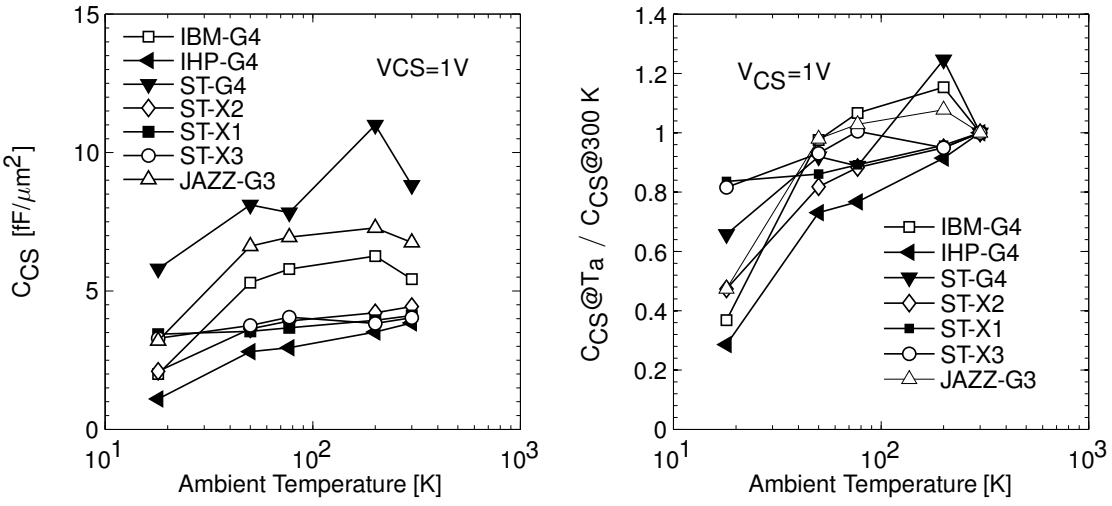


Figure 6.15: (a) C_{CS} and (b) normalized C_{CS} as a function of ambient temperature for each of the devices

The collector–substrate capacitances and the normalized collector–substrate capacitances are plotted against ambient temperature for each of the seven devices in Fig. 6.15. As expected, a sharp decline in the collector–substrate capacitance is observed in the majority of the devices as the temperature is lowered from 50 to 18 K. It is believed that the remaining capacitances measured at 18 K are residual second-order effects, such as fringing capacitance from the collector terminal to the metallic ground plane. Referring to the absolute values of the capacitances[7], it can be seen that they are on the order of 3-8 fF/μm^2, which is quite small.

6.3.1.3 Base–Emitter Capacitance

As discussed in Chapter 2, the base–emitter capacitance is the parallel combination of a depletion capacitance, $C_{BE,depl}$, and a diffusion capacitance, $C_{BE,diff}$. As the depletion capacitance is a junction capacitor, its temperature dependence is determined by the doping levels, which for modern SiGe HBTs are high enough to render $C_{BE,diff}$ nearly completely independent of temperature. On the other hand, the diffusion capacitance does depend on temperature as $C_{BE,diff} = g_m\tau_f =$

[7]It should be noted that a high value of C_{CS} was expected for the IHP device due to the lack of DTI in the SG-13 process. However, referring to Fig. 6.15(a), the measured value of C_{CS} was found to be lower per unit-area than any of the other devices. From the documentation provided with the devices, it looks as if the low-impedance ground plane is not ohmically connected to the substrate. This would then explain the lower capacitance as the capacitance from the ground plane to the substrate would be in series with C_{CS} and is only expected to be on the order of a couple of fF.

$qI_C\tau_f/n_c k\mathrm{T_a}$. For a SiGe HBT, $\tau_f = \tau_b + \tau_e$, where [30]

$$\tau_b \approx \frac{W_B^2}{2\widetilde{\mu}_{nb}\Delta E_{g,Ge}\,(grade)} \tag{6.37}$$

and

$$\tau_e \approx \frac{1}{\beta_{AC}}\left(\frac{W_e}{S_{pe}} + \frac{W_e^2}{2D_{pe}}\right), \tag{6.38}$$

where $\widetilde{\mu}_{nb}$ is the position averaged minority carrier mobility in the base, $\Delta E_{g,Ge}$ is the Ge grating in V (as opposed to eV; q has been factored out), S_{pe} is the hole surface recombination velocity at the emitter contact and W_e is the width of the neutral emitter. In devices with high β_{AC}, the forward transit time is limited by τ_b, and $C_{BE,diff}$ can be written as

$$C_{BE,diff} \approx g_m \frac{W_B^2}{2\widetilde{\mu}_{nb}\Delta E_{g,Ge}\,(grade)}. \tag{6.39}$$

Referring back to Fig. 3.4(b), it can be seen that there is little change in the mobility as a function of temperature for the impurity concentrations present in the base of modern SiGe HBTs. Thus, for a fixed collector-current density, we would expect $C_{BE,diff}$ to increase proportionally to the transconductance as the device is cooled, resulting in a maximum fractional increase in C_{BE} of $(300/T_a) \times (n_c(300)/n_c(T_a))$.

The base–emitter capacitance for each of the devices was extracted as a function of temperature and bias current, and smooth curves were fit to the data at each temperature. A complete listing of the coefficients used to model the capacitors as a function of V_{BE} appears in Appendix C.3.1. In order to study the temperature dependence of C_{BE}, the extracted values were plotted as a function of temperature for a fixed collector-current density, and the results are shown in Fig. 6.16(a). As expected, the JAZZ-G3 device has considerably higher capacitance per unit-area than the other devices.

In Fig. 6.16(b), the fractional increase of C_{BE} for a fixed collector-current density of $1\mathrm{mA}/\mu\mathrm{m}^2$ is plotted as a function of temperature. It can be seen that the increase in C_{BE} as the ambient temperature drops from 300 to 18 K varies from approximately 1.4 to 2.5. These values are considerably lower than the observed fractional increase of 2.4 to 4.3 in the transconductance as the temperature was lowered. This indicates that the depletion capacitance must be a significant portion of the overall capacitance. In order to verify the hypothesis that the temperature dependence of the base–emitter capacitance is completely contained in the transconductance term in equation (6.39), the

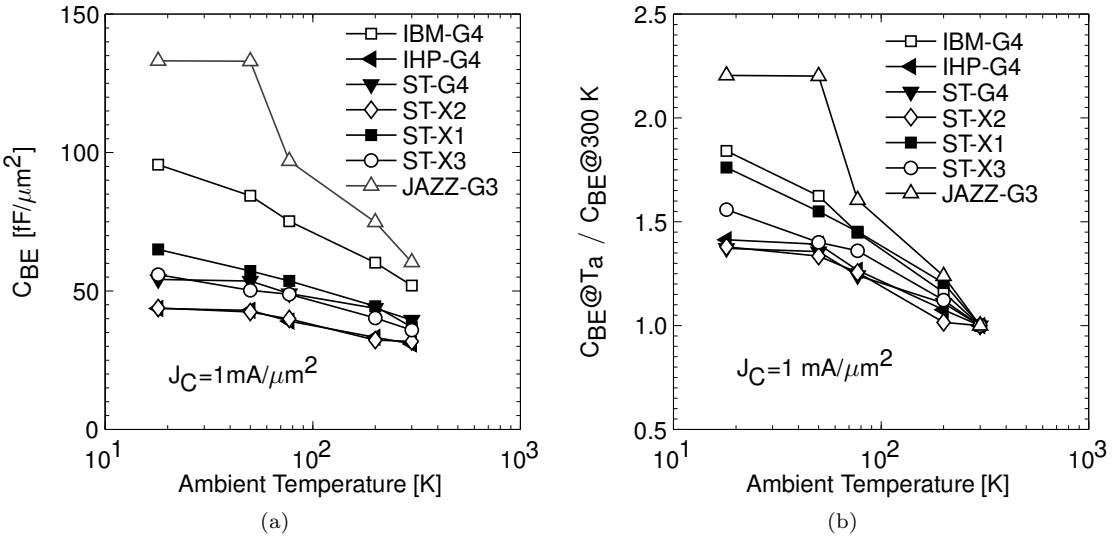


Figure 6.16: (a) C_{BE} and (b) normalized C_{BE} as a function of ambient temperature for each of the devices

low-temperature[8] collector-base capacitance values for each of the devices was extrapolated based upon the values measured at 200 and 300 K in conjunction with the cryogenic transconductance values[9]. The results of this procedure are plotted for several of the devices in Fig. 6.17 and indicate that, to first order, the temperature dependence of C_{BE} is in fact strongly coupled to that of the transconductance. Thus, it is possible to roughly determine the base–emitter capacitance at cryogenic temperatures by knowing only the cryogenic transconductance and the values of g_m and C_{BE} at two higher temperatures.

In addition to studying the temperature dependence of C_{BE} at a fixed collector current density, it is also important to study it for a fixed transconductance. In Fig. 6.18(a), C_{BE} is plotted as a function of temperature for a fixed transconductance of 100 mS/μm^2, and in Fig. 6.18(b) the values of C_{BE}@18 K normalized to their room temperature values are plotted as a function of transconductance. Clearly, the assumption that room temperature collector–base capacitance values can be used at 18 K is invalid. This is very important as it implies that designs cannot be carried out using room temperature models with the intention of scaling the collector current at cryogenic temperatures to achieve the transconductance that was simulated at 300 K. Instead, in order to accurately predict circuit performance, it is necessary that cryogenic circuits be designed using models developed at cryogenic temperatures.

[8] i.e. 77 K and below

[9] The extrapolation used is $\widehat{C}_{BE}(T_a) = m\,(g_m(T_a) - g_m(300\text{ K})) + C_{BE}(300\text{ K})$, where $m = (C_{BE}(300\text{ K}) - C_{BE}(200\text{ K}))\,/\,(g_m(300\text{ K}) - g_m(200\text{ K}))$. Physically, the slope coefficient, m, is related to the change in τ_b with cooling and is dependent upon the base doping.

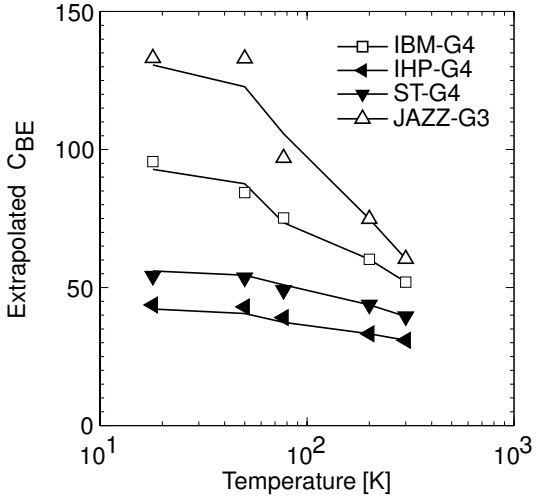


Figure 6.17: Low-temperature C_{BE} as predicted from $g_m(T_a)$ and C_{BE} @ 200 and 300 K. $J_C = 1\text{mA}/\mu\text{m}^2$

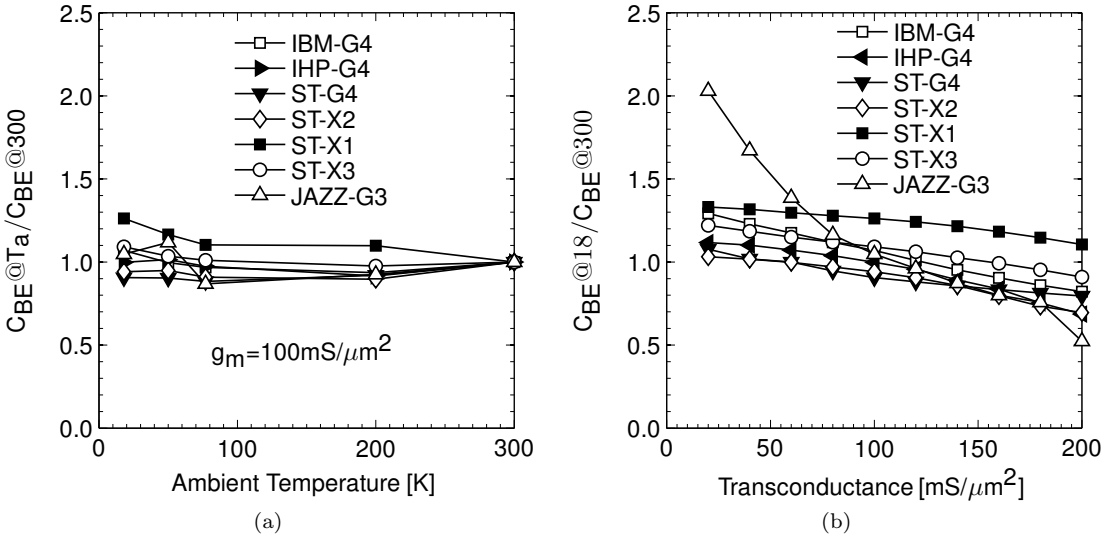


Figure 6.18: The temperature dependence C_{BE} for a fixed value of g_m. (a) The normalized value of C_{BE} as a function of ambient temperature for a fixed transconductance of 100 mS/μm^2. (b) The normalized value of C_{BE}@18 K as a function of transconductance.

6.3.2 Resistances

The resistances in a SiGe HBT arise as a consequence of the finite conductivity of the tungsten contacts, polysilicon layers, and doped silicon layers. A "back of the envelope" calculation[10] as to the magnitude of the resistive component due to the tungsten contacts at room temperature gives a result on the order of 0.4 $\Omega\mu\text{m}^2$ [96]. When cooled below 77 K, the conductivity of tungsten increases by several orders of magnitude [137]. Thus, one can expect that a small component of each of the resistances will go away when the device is cooled due to an increase in the conductivity of the tungsten contacts.

In order to consider what will happen to the resistances associated with the doped silicon and polysilicon regions, we can begin by writing an expression for the sheet resistance of a thin doped region of semiconductor:

$$R_S = \frac{1}{\int_0^W q\, N_{A(D)}^{-(+)}(x)\, \mu_{p(n)}(x) dx}, \ [\Omega/\square] \tag{6.40}$$

where W is the thickness of the region of semiconductor and x is the depth in the material [55].

The evaluation of equation (6.40) is not trivial since the position dependent mobility is a function of both temperature and ionized impurity concentration. However, closed form expressions for all required variables were provided in Chapter 3 (see equations (3.14) and (3.18)). Thus, it will be very useful to investigate how the sheet-resistance of the materials varies for a fixed level of doping.

This calculation was completed for As, P, and B dopants and the resulting curves are shown in Fig. 6.19. For polysilicon layers, whose doping level is on the order of 1×10^{21} cm^{-3}, the calculations show that regardless of which dopant is used, one can expect a decrease in sheet resistance of approximately 25% for cooling from 300 to 77 K and an additional 15% decrease when going all the way down to 18 K. It should also be noted that these calculations are in line with the experimental result of $\sim 20\%$ reported by Clark et al. for cooling to 77 K [138].

While the situation with the polysilicon resistances is fairly straightforward, the temperature dependence of the lightly to moderately doped semiconductor regions is a bit more complicated due to carrier freeze-out effects. For instance, if we take the case of As dopants, we see that for impurity concentrations below about 1×10^{19} cm^{-3}, the sheet resistance ceases to be a monotonically decreasing function with respect to temperature. Instead, there is some temperature between 300 and 18 K at which the sheet resistance is minimum. Furthermore, for As acceptor impurity concentrations

[10] For an IBM BiCMOS8HP device, there are 46 contacts in a $0.12 \times 18\mu$ m^2 region. As each contact is 9 Ω, the resistance is 0.4 $\Omega\mu\text{m}^2$.

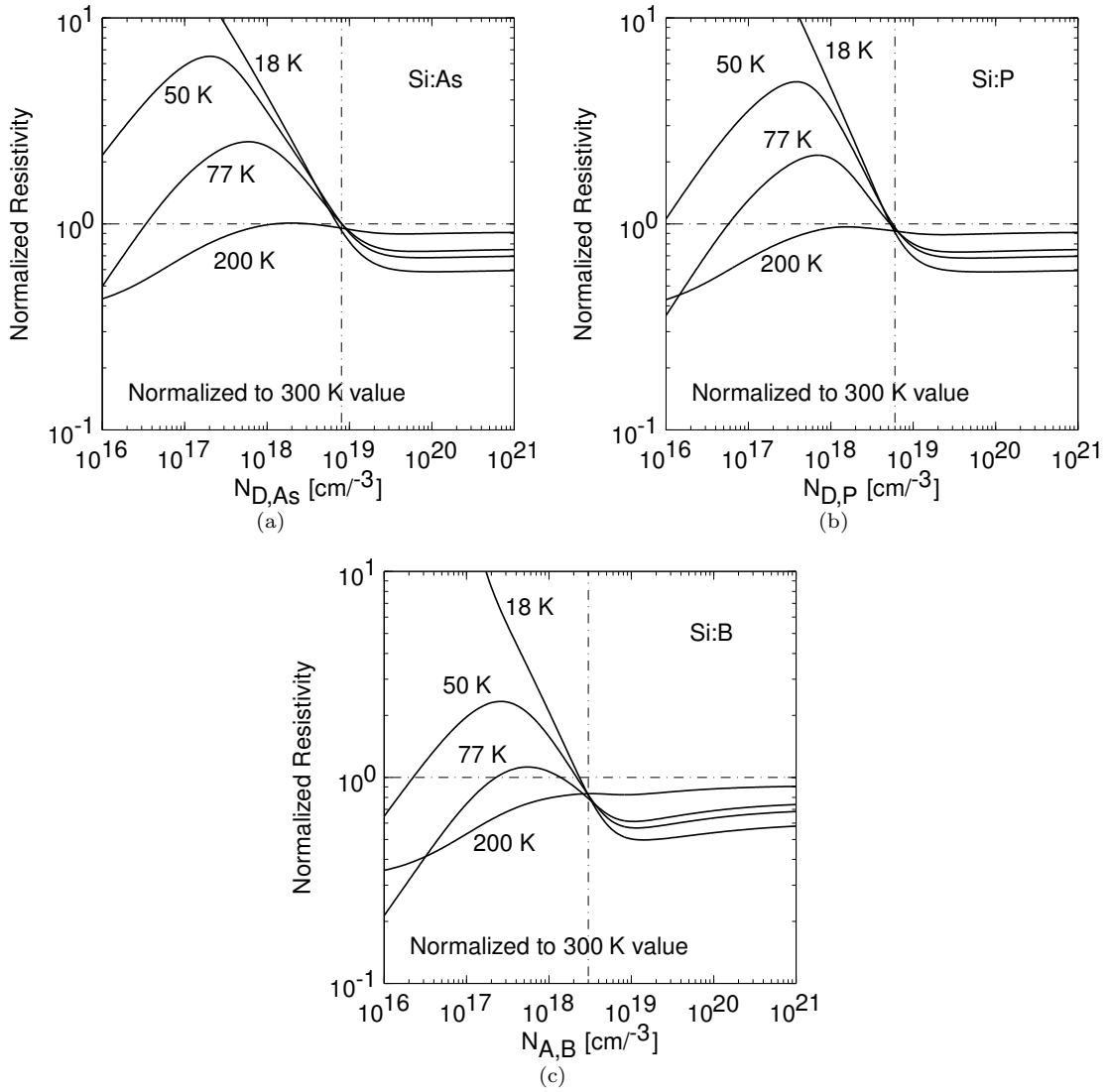


Figure 6.19: Normalized resistivity as a function of doping and temperature for a Si sample doped with (a) As donor impurities, (b) P donor impurities, and (c) B acceptor impurities. The emitter is usually doped with As, whereas the collector is usually doped with P. B is used in the base. The effects of incomplete ionization and temperature were taken in to account when computing mobility.

lower than about 8×10^{19} cm^{-3}, carrier freeze-out effects cause a steep increase in sheet resistance at cryogenic temperatures. Thus, the behavior as a function of temperature of the resistances intrinsic to a SiGe HBT will be highly dependent upon the impurity concentration in each of the regions of the device. As this is only important at cryogenic temperatures, one would not expect the foundries to take carrier freeze-out effects into account when optimizing the doping profiles. Thus, the behavior of the resistances as a function of temperature is inherently difficult to predict. In the next section, we will examine the behavior of the base, emitter, and collector resistances as a function of temperature.

6.3.2.1 Base Resistance

The base resistance is a series combination of extrinsic resistances due to the polysilicon base contact ($r_{bx,p}$) and the tungsten base contact resistance ($r_{bx,t}$) and an intrinsic spreading resistance (r_{bi}). The extrinsic components are assumed to be bias independent whereas the intrinsic part of the base resistance decreases with increasing bias due to current crowding (distributed) effects [54]. As discussed above, a decrease of $\sim 40\%$ is expected in $r_{bx,p}$ with cooling from 300 to 18 K. The effect of $r_{bx,t}$ is assumed to be very small. Thus, the major unknown is the intrinsic base resistance, r_{bi}.

The base resistance was extracted as a function of bias for each of the seven devices and an example result from each of the foundries appears in Fig. 6.20. At low biases, the IBM-G4 and IHP-G4 devices have a minimum base resistance at a temperature in the vicinity of 50 K whereas the minima is closer to 77 K for the ST-X1 and JAZZ-G3 devices. At current densities above 1 mA/μm^2, the improvement in r_b with cooling becomes less significant, and in the case of the IBM-G4 and JAZZ-G3 devices, there is even a penalty in base resistance for cooling all the way to 18 K.

The percentage change in r_b with cooling from to 300 to 18 K is shown for each of the devices at 0.1 and 3.0 mA/μm^2 in Fig. 6.21. At both the low- and high-injection bias points there is considerable variation among the different devices. Recently, a study was published by Garcia et al. in which the base resistance of eight different transistor were reported at 300 and 40 K [33], with an average decrease in base resistance of 22% measured with cooling. The standard deviation of the data is 10%. Unfortunately, the bias point was not stated. This is in agreement with the extracted 50 K low-injection data reported here as the mean decrease in base resistance was found to be 18.7% with a standard deviation of 16%. It should be noted that the standard deviation in the work reported here is expected to be higher than that reported in [33], as the result of a wider variety of devices under investigation.

6.3.2.2 Emitter Resistance

In modern SiGe HBTs, a p-doped silicon capping layer is grown on top of the SiGe base in order to apply strain to the Ge film, thereby increasing the Ge-induced bandgap narrowing. To form the emitter, an extrinsic polysilicon layer is deposited on top of the capping layer and a thermal annealing process is carried out, resulting in the diffusion of a large concentration of As, or in the case of the IBM devices, P, dopants into the capping layer. This layer then serves as the intrinsic n+

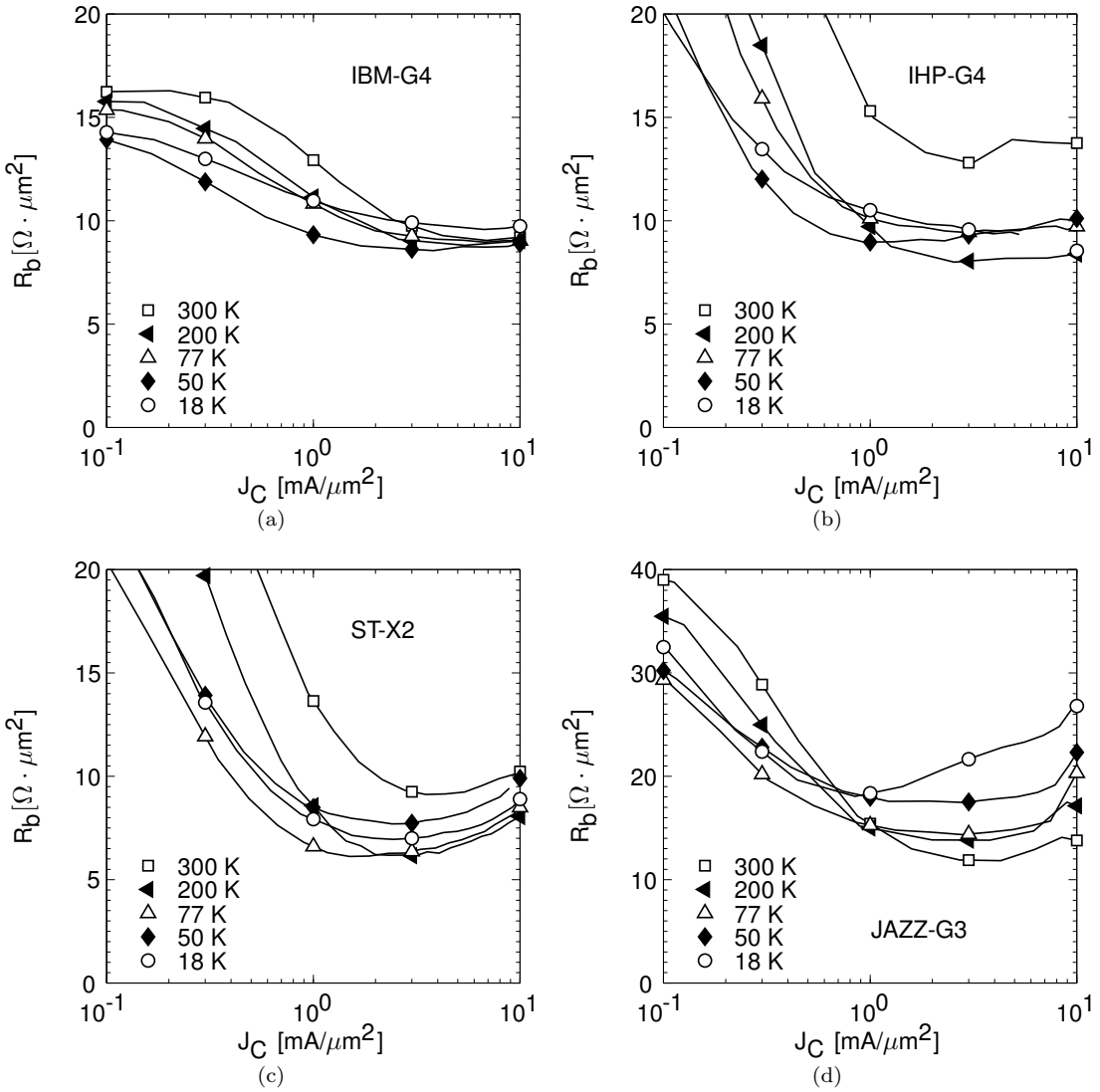


Figure 6.20: Extracted base resistance as a function of J_C

emitter [54]. While the dopant concentration in the extrinsic emitter is quite high, the concentration of impurities in the intrinsic emitter is dependent upon the annealing process and can be considerably lower.

Thus, there are two components to the emitter resistance; the external polyemitter, which we expect to decrease by about 40% with cooling to 18 K, and the intrinsic emitter, which may or may not change significantly depending on the exact concentration of dopants in the region. The emitter resistance has been extracted for each of the seven devices under study and the results appear in Fig. 6.22. Interestingly enough, the majority of the devices have an increase in emitter resistance as they are cooled from 300 K to 18 K. The physical explanation for this must be that the emitter

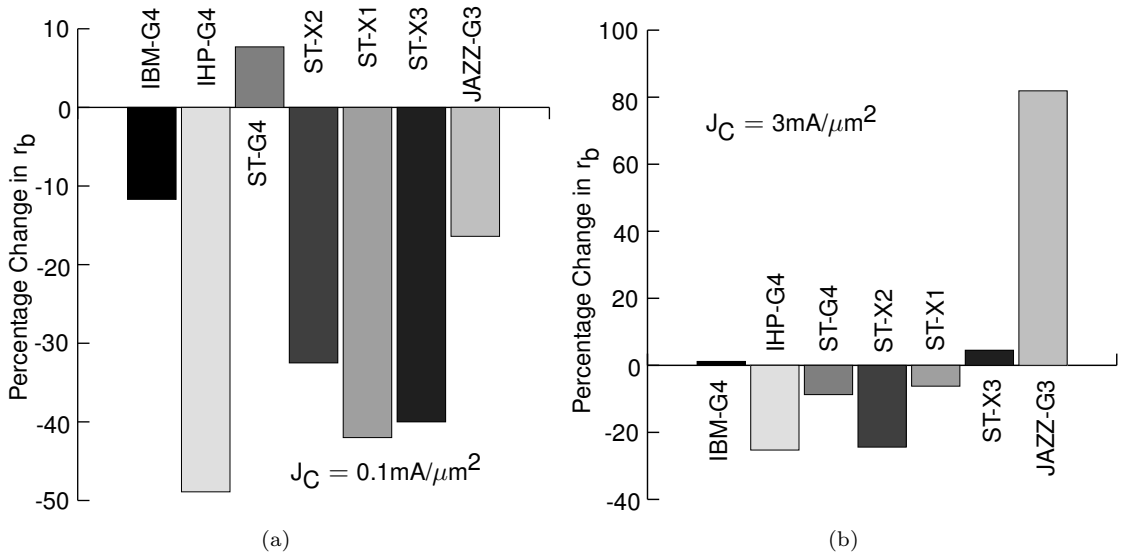


Figure 6.21: Percentage change in r_b with cooling. (a) $J_C = 0.1$ mAμm^2 and (b) $J_C = 3$ mA/μm^2

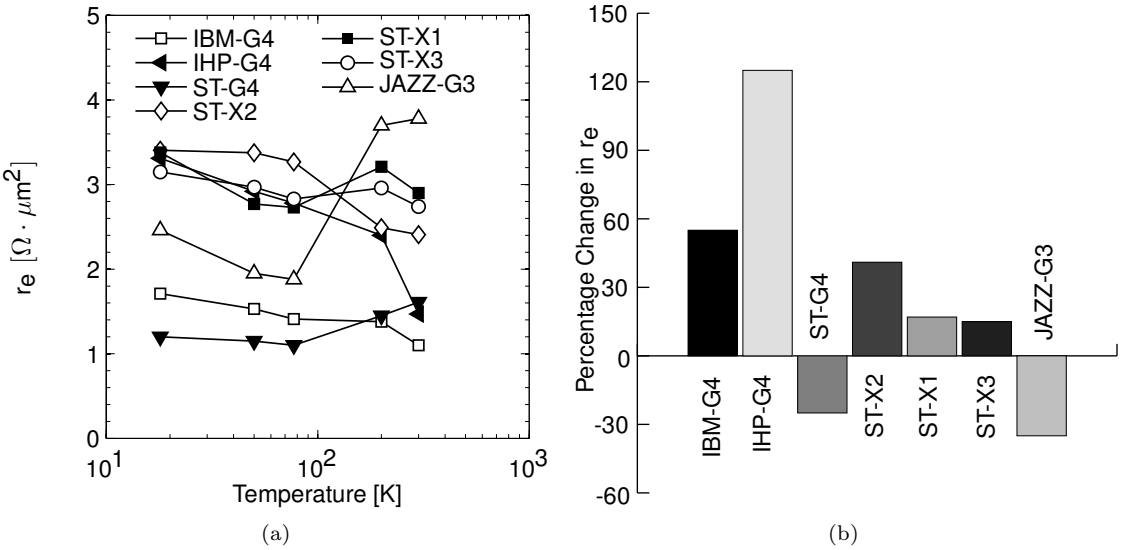


Figure 6.22: (a) Extracted emitter resistance for each of the devices and (b) percentage change with cooling from 300 to 18 K

region closest to the metallurgical base–emitter junction of the devices is doped slightly below the Mott transition and is freezing-out at cryogenic temperatures.

6.3.2.3 Collector Resistance

The collector resistance results from the intrinsic collector, as well as the access resistance between the SIC and the collector terminals. As mentioned in Chapter 4, the IHP devices are

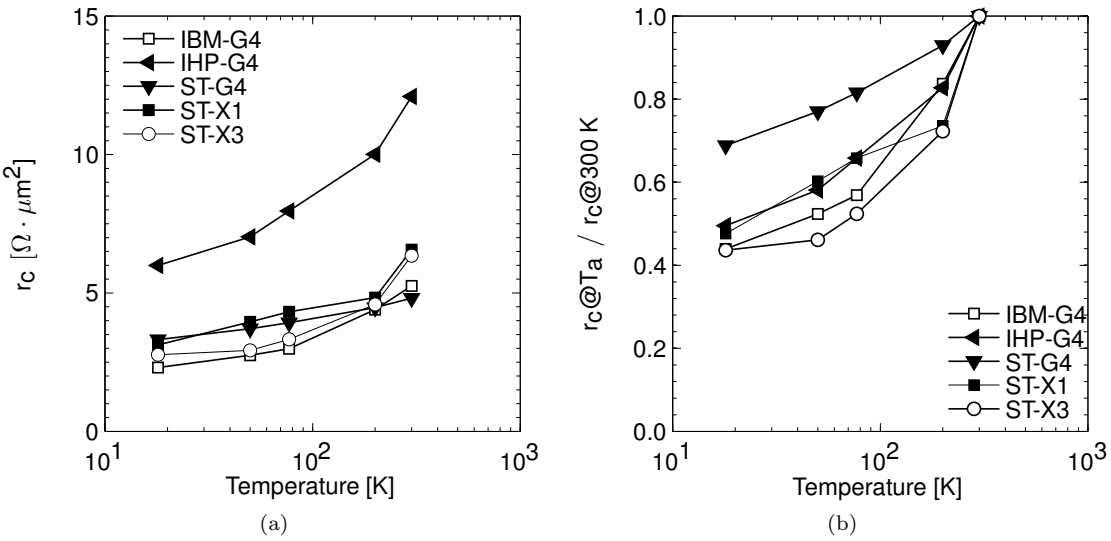


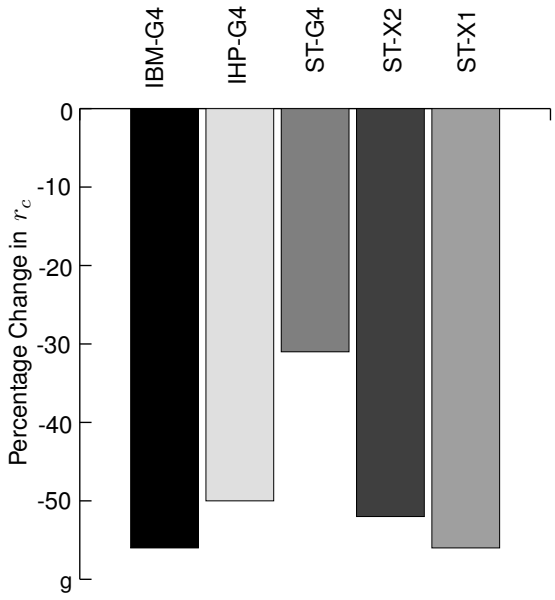
Figure 6.23: (a) Collector resistance and (a) normalized collector resistance as a function of ambient temperature

expected to have a higher extrinsic collector resistance in comparison to the other technologies, due to use of an implanted sub-collector. As seen in Fig. 6.23, the measured resistance confirms this expectation. However, it is interesting to note that the fractional improvement seen in the collector resistance of the IHP device with cooling is similar to that of the devices with high quality buried sub-collectors. Furthermore, referring to Fig. 6.24 the percentage change that was observed with cooling from 300 to 18 K was 50 %, which is indicative that main sources of collector resistance are the heavily doped buried sub-collector and reach-through implants.

6.3.3 Transconductance, g_m, and Base–Emitter Conductance, g_{be}

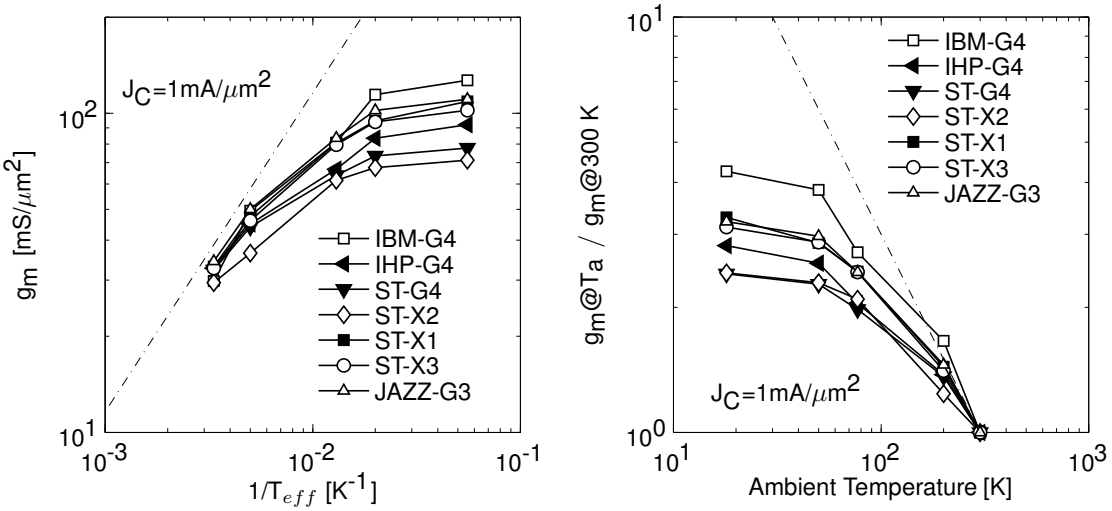
The extracted intrinsic transconductance and input conductance are plotted in figures 6.25 and 6.26 for a collector current density of 1mA/μm^2. The extrinsic transconductance was studied in Chapter 5, and, for low biases, is equal to the intrinsic transconductance. At 1mA/μm^2, an increase in the transconductance ranging from 2.4 to 4.4 was observed. As discussed in Chapter 5, non-equilibrium effects are responsible for limiting the achieved transconductance.

The base–emitter conductance was computed and the results appear in Fig. 6.26. If β_{AC} and g_m rise together with cooling, there should be no change in g_{be} as a function of temperature. However, as β_{AC} rises much more rapidly than g_m, g_{be} drops accordingly. The benefit of this drop is quite profound. As g_{be} is vanishingly small at 18 K, it can be neglected in the small-signal model for all biases except under very high injection. Furthermore, as the early effect is negligible in SiGe



(a)

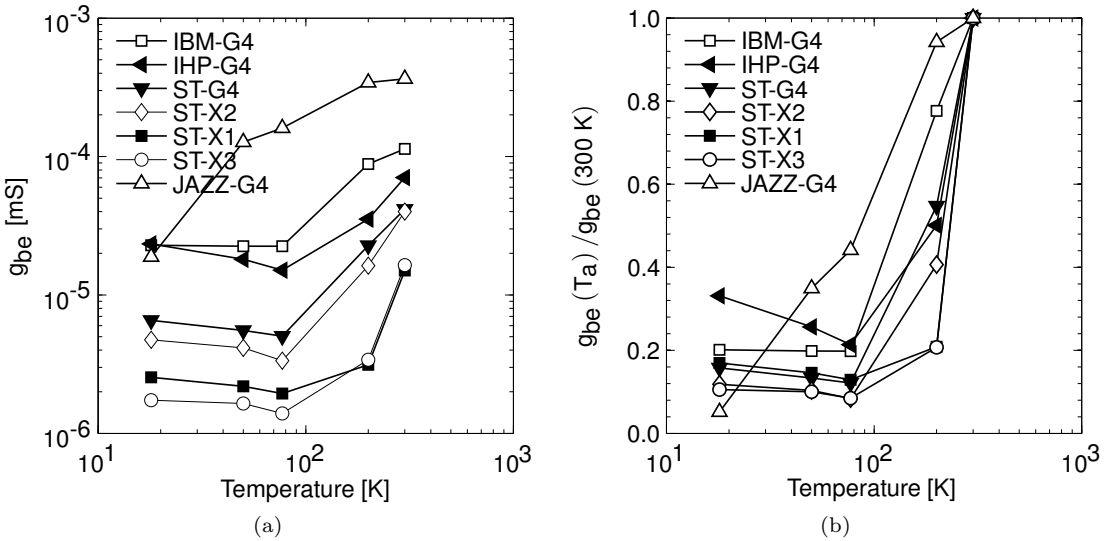
Figure 6.24: Percentage change in r_c



(a) (b)

Figure 6.25: (a) Transconductance and (b) normalized transconductance as a function of ambient temperature

HBTs [30], the output impedance is also very high. In comparison to FETs, which have finite output resistance, HBTs are a much more ideal device at cryogenic temperatures in terms of input and output resistances.



(a) (b)

Figure 6.26: (a) Base–emitter conductance and (b) normalized base–emitter conductance as a function of ambient temperature

6.4 Summary

In this chapter, a detailed investigation of the small-signal performance of SiGe HBTs has been presented. The study began with a brief look into the behavior of f_t and f_{max} with cooling in which it was seen that the peak value of f_t is improved by about 50% with cooling. The behavior of f_{max} with cooling was found to be process dependent. After discussing the high-frequency figures-of-merit, a small-signal extraction procedure was presented and applied to model several state-of-the-art SiGe HBTs at a variety of temperatures. Next, the characteristics of each component of the small-signal model were discussed in terms of temperature dependence both with respect to how the results compare with theoretical expectations and in terms of how the device properties vary among the different processes. In the next chapter, the noise of the devices will be modeled using the small-signal and dc models which have been developed.

Chapter 7

Noise Modeling

In this chapter, the models and theory developed in the preceding sections of the dissertation are applied to determine the ultimate limitations of SiGe HBTs in terms of noise performance. Furthermore, by analyzing the noise performance in terms of a figure of merit in which both noise and gain are taken into account, an accurate limit on achievable system noise performance is determined. The main goals of this chapter are 1) to determine the fundamental noise performance limitations of SiGe HBTs in the ambient temperature range of 0–300 K, 2) to understand how the noise properties of the devices change with temperature, and 3) to evaluate how SiGe devices compare with state-of-the-art InP HEMT devices in terms of cryogenic noise performance.

The chapter begins with a brief review of the general concepts used to describe noise performance as well as a description of the specific representation used to model the noise in a SiGe HBT. Next, a procedure for the systematic determination of the metrics of interest is presented. With the necessary framework in place, the noise modeling procedure is then applied to the devices that were modeled in Chapter 6, and the results are discussed thoroughly. Finally, the chapter is drawn to a close with a comparison of the noise performance of SiGe HBTs with that of state-of-the-art InP HEMTs. Prior to describing the procedure of determining the noise properties, we will begin with a brief review of noise parameters.

7.1 A Brief Review of Noise Parameters and the Concept of Cascaded Noise Temperature

The spot noise-temperature of a linear two-port can be fully described by the formula [68]

$$T_e = T_{min} + T_0 \frac{N\left|Y_S - Y_{OPT}\right|^2}{\Re\left\{Y_{OPT}\right\}\Re\left\{Y_S\right\}}, \tag{7.1}$$

where $N = R_n \Re\{Y_{OPT}\}$. While there are several alternative forms in which the equation (7.1) is commonly cast, from a device modeling point of view this particular representation is ideal due to its fundamental nature [68, 69]. In particular,

1) T_{min} and N are invariant with respect to lossless transformations applied at both the input and output of the two-port. Thus, they not affected by lossless packaging.

2) T_{min} and N are independent of device size[1].

3) Y_{OPT} varies with lossless transformations to the input of the two-port as a typical admittance would. Furthermore, Y_{OPT} scales with size in the same manner as other admittances (i.e., a factor of two increase in device size will result in a factor of two increase in Y_{OPT}.)

Thus, this set of noise parameters permits a fair comparison of devices, as T_{min} and N do not depend upon device size, and Y_{OPT} scales with device size and can be normalized to a device of unit area.

7.1.1 The Concept of Cascaded Noise Temperature

While the noise temperature and noise figure have become the standard figures of merit for low-noise amplifiers and devices, they do not provide the information required to determine system noise performance. The reason for this is straightforward; the noise temperature and noise figure do not take gain into account. Thus, knowledge of only T_e does not give enough information to determine how the amplifier will behave in a system. The importance of gain in terms of system noise temperature can be understood by considering Friss's formula for the cascading of noisy networks [139]:

$$T_{sys} = T_{e1} + \frac{T_{e2}}{G_{a1}} + \frac{T_{e3}}{G_{a1}G_{a2}}..., \tag{7.2}$$

where T_{sys} is the overall input referred noise temperature of the cascade, T_{e1} and G_{a1} are the noise temperature and available gain of the first amplification stage, T_{e2} and G_{a2} are the noise temperature and available gain of the second amplification stage, and so on. Referring to equation 7.2, it is evident that the gain of an LNA has a serious impact on the system noise temperature, even though the noise temperature of the LNA is independent of gain.

A far more fundamental figure of merit is the cascaded noise temperature, which is a variant of

[1]It should be noted that R_n is a much more commonly quoted noise parameter in the literature. However, for the purpose of device characterization, it is a biased figure of merit, as it is not independent of device area. R_n is inversely proportional to the device size whereas Y_{OPT} scales with device size. Therefore, $N = R_n \Re\{Y_{OPT}\}$ is independent of device size.

the noise measure introduced by Haus and Adler [8, 140], and is given as[2]

$$T_{CAS} = T_0 M = \frac{T_e}{1 - 1/G_a}, \tag{7.3}$$

where M is the noise measure and $G_a = |Y_{21}|^2 \Re\{Y_S\} / \left(\Re\{Y_{22}\} |Y_{11} + Y_S|^2 - \Re\{Y_{12} Y_{21} (Y_{11}^* + Y_S^*)\}\right)$ is the available gain of the network, and is a function of source admittance, Y_S [141]. The cascaded noise temperature is a very fundamental property of a network since its minima is invariant to any lossless transformation to the network; that is, it is invariant to lossless input, output, and feedback networks. Furthermore, T_{CAS} can be interpreted physically as the noise temperature of an infinite chain of identical devices. Thus, it sets a limit on the system noise performance that is achievable using a given device. Finally, for high values of G_a, the cascaded noise temperature simplifies to the standard noise temperature. In this work, the cascaded noise temperature will be reported rather than the standard noise temperature.

7.2 Noise Model

As discussed in Section 2.2, the noise parameters and T_{cas} of a SiGe HBT are completely determined by the small-signal model and dc bias currents, so long as the shot-noise sources are uncorrelated. Furthermore, if the shot-noise sources are correlated, then the impact is a reduction in the minimum achievable noise of the device. Therefore, it is possible to estimate the noise performance of a SiGe HBT knowing just the dc currents, which determine the magnitudes of the shot-noise sources, the values of parasitic resistances, which contribute thermal noise, and the small-signal model, which determines the frequency response of the device. Once the noise parameters for a given device are determined, they can then be converted to T_{cas} allowing for the evaluation of the ultimate performance limitations of the HBT.

The full small-signal noise model used in this work is shown in Fig. 7.1. The model has the same lumped components as the model presented in Chapter 6, with additional thermal noise sources for each of the physical resistances and shot noise sources associated with the base and collector currents. It should be noted that the shot-noise sources are considered to be uncorrelated in this work and their magnitudes are tabulated directly from dc measurements. In addition the input conductance does not have thermal noise associated with it as it is not a physical resistance. In the next section, the procedure for determining the noise performance from the small-signal model and

[2]equation (7.3) does not apply when there are negative resistances in the circuit. In this case, it is necessary to replace G_a with G_e, which is the exchangeable gain of the network.

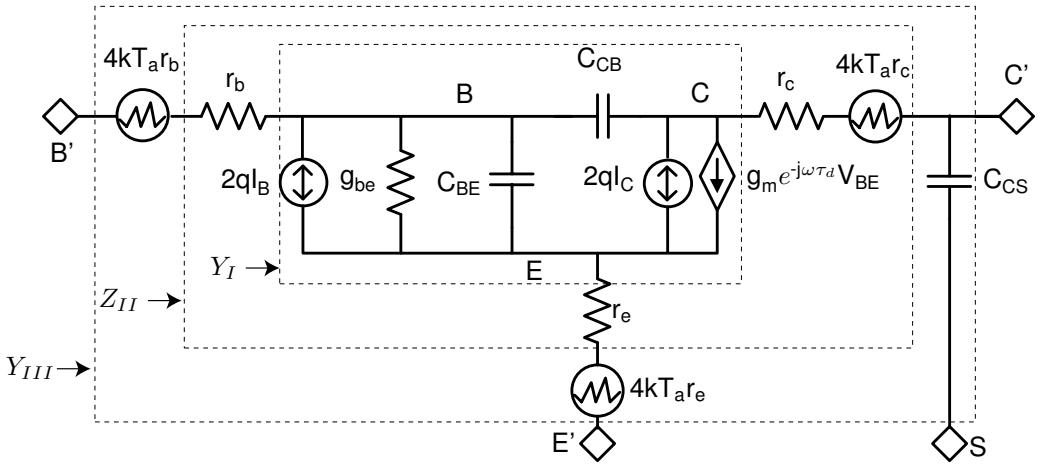


Figure 7.1: Complete small-signal HBT noise model

dc terminal currents is summarized.

7.3 Noise Modeling Procedure

The HBT noise modeling procedure begins with the measurement of the dc terminal characteristics and the extraction of the small-signal model over a wide range of bias points. These topics were discussed at length in Chapters 5 and 6 and will not be revisited here. Assuming the bias and small-signal characteristics of the device are known, the next step is to use a noise imbedding procedure in order to determine the equivalent chain-representation of the noise, as referred to the input terminals of the device. Once these noise spectra are known, determination of the noise parameters is straightforward. Finally with the noise parameters known, the cascaded noise temperature can be computed. In this section, this procedure will be detailed and pertinent equations will be provided. Derivations of the key equations appear in Appendix E.

7.3.1 Computation of Noise Parameters

With the values of the small-signal model and internal noise sources known, determination of the noise parameters is a straightforward exercise. To begin with, the noise parameters of the intrinsic network are written in admittance (short-circuit) representation as

$$\mathbf{N_I^Y} = \begin{bmatrix} 2qI_B & 0 \\ 0 & 2qI_C \end{bmatrix}. \tag{7.4}$$

Similarly, the admittance parameters of the intrinsic network can be written as

$$\mathbf{Y_I} = \begin{bmatrix} g_{be} + j\omega\left(C_{BE} + C_{CB}\right) & -j\omega C_{CB} \\ g_m - j\omega C_{CB} & j\omega C_{CB} \end{bmatrix}. \tag{7.5}$$

Now, with the noise and network parameters of the intrinsic network known, the next step is to convert the matrices to the impedance (open-circuit) representation and add the effects of the series resistances. The conversion of $\mathbf{Y_I}$ to $\mathbf{Z_I}$ is trivial as $\mathbf{Z_I} = \mathbf{Y_I}^{-1}$. The noise parameters can be converted using $\mathbf{N_I^Z} = \mathbf{Z_I N_I^Y N_I}^{\dagger}$, where the † operator refers to the hermitian transpose operation [142].

Now, with the intrinsic network in the impedance representation, the imbedding of the series resistances is straightforward:

$$\mathbf{Z_{II}} = \mathbf{Z_I} + \begin{bmatrix} r_b + r_e & r_e \\ r_e & r_c + r_e \end{bmatrix} \tag{7.6}$$

and

$$\mathbf{N_{II}^Z} = \mathbf{N_I^Z} + 4k\mathrm{T}_a \begin{bmatrix} r_b + r_e & r_e \\ r_e & r_c + r_e \end{bmatrix}. \tag{7.7}$$

At this point, all of the external noise sources are accounted for. Referring back to the equivalent circuit, one might think that it is necessary to take the filtering effect of C_{CS} into account before converting the noise spectral densities to the desired set of noise parameters. However, the way that the noise properties of a network are defined is in terms of available noise power at the output of the network. Thus, as the collector–substrate capacitance can always be tuned out, it has no effect on the noise parameters, and need not be accounted for in their calculation. Nonetheless, it may be desired to calculate the full set of network parameters, so the shunt capacitance at the output will be added for completeness:

$$\mathbf{Y_{III}} = \mathbf{Y_{II}} + \begin{bmatrix} 0 & 0 \\ 0 & j\omega C_{CS} \end{bmatrix} \tag{7.8}$$

and $\mathbf{N_{III}^Y} = \mathbf{N_{II}^Y}$. As the final step, the following equations are used to convert to the desired noise representation [30, 142, 143, 144]:

$$\mathbf{N_{III}^A} = \begin{bmatrix} \overline{|N_{v_n}|^2} & \overline{N_{v_n,i_n^*}} \\ \overline{N_{v_n^*,i_n}} & \overline{|N_{i_n}|^2} \end{bmatrix} = \begin{bmatrix} 0 & -1/Y_{21}^{\mathrm{III}} \\ 1 & -Y_{11}^{\mathrm{III}}/Y_{21}^{\mathrm{III}} \end{bmatrix} \mathbf{N_{III}^Y} \begin{bmatrix} 0 & 1 \\ -1/Y_{21}^{\mathrm{III}*} & -Y_{11}^{\mathrm{III}*}/Y_{21}^{\mathrm{III}*} \end{bmatrix}, \tag{7.9}$$

$$T_{min} = \frac{1}{2\mathrm{k}} \left(\sqrt{\overline{|N_{v_n}|^2}\,\overline{|N_{i_n}|^2} - \Im\left\{\overline{N_{v_n^* i_n}}\right\}^2} + \Re\left\{\overline{N_{v_n^* i_n}}\right\} \right), \tag{7.10}$$

$$N = \frac{1}{4\mathrm{kT}_0} \sqrt{\overline{|N_{v_n}|^2}\,\overline{|N_{i_n}|^2} - \Im\left\{\overline{N_{v_n^* i_n}}\right\}^2} \tag{7.11}$$

$$G_{OPT} = \frac{1}{\overline{|N_{v_n}|^2}} \sqrt{\overline{|N_{v_n}|^2}\,\overline{|N_{i_n}|^2} - \Im\left\{\overline{N_{v_n^* i_n}}\right\}^2}, \tag{7.12}$$

and

$$B_{OPT} = -\frac{\Im\left\{N_{v_n^* i_n}\right\}}{\overline{|N_{v_n}|^2}}. \tag{7.13}$$

Therefore, given a small signal-model equivalent circuit model, the noise parameters can be determined in a systematic fashion that is amenable to automation.

7.3.2 Computation of $T_{cas,min}$

Upon determining the noise and network parameters of the device, it is possible to compute T_{CAS} as a function of source impedance at any particular bias. A particularly insightful formula for the noise measure was provided by Fukui and is converted to T_{CAS} to give [141, 145]:

$$T_{CAS} = \frac{T_{min} + T_0 \frac{N}{\Re\{Y_{OPT}\}\Re\{Y_S\}} \left|Y_S - Y_{OPT}\right|^2}{1 - \frac{1}{G_{a,max}} - \frac{\Re\{Y_{22}\}}{\Re\{Y_S\}|Y_{21}|^2} \left|Y_S - Y_{OPT,G}\right|^2}, \tag{7.14}$$

where $G_{a,max}$ is the maximum available gain of the network and $Y_{OPT,G}$ is the generator impedance required to achieve this gain. Referring to equation (7.14), it is apparent that, for finite $G_{a,max}$, $T_{CAS} > T_{min}$, and its minimization requires orchestrating a tradeoff between gain and noise matching. In [141], Fukui showed the loci of source impedances producing constant noise measure form circles when plotted on the smith chart, and that there is a unique optimum source impedance that minimizes the noise measure[3].

A closed-form expression for the minimum noise measure as a function of source impedance has been derived by Poole and Paul and once converted to T_{CAS} is given as [146, 147]

$$T_{CAS,min} = T_0 \frac{-M_b + \sqrt{M_b^2 - 4M_a M_c}}{2M_a}, \tag{7.15}$$

where M_a, M_b, and M_c are defined in Table 7.1. Furthermore, Poole and Paul also provide the

[3]The existence of the optimum was also shown earlier by Haus and Adler. Furthermore, it was shown that if the network is analyzed using a matrix formulation, then this minima is the smallest positive eigenvalue of the characteristic noise matrix, with optimum source impedance corresponding to the associated eigenvector [140].

Table 7.1: Formulas required to compute M_{min}

Coefficient	Equation or description
M_a	$\left\|1+\Gamma_{on}\right\|^4\left(PQ+\left\|C_1\right\|^2\right)$
M_b	$\left\|1+\Gamma_{on}\right\|^2\left\|S_{21}\right\|^2\left(8r_n\Re\left\{\Gamma_{on}C_1\right\}-\left(4r_n\left\|\Gamma_{on}\right\|^2+W\right)P-\left(W-4r_n\right)Q\right)$
M_c	$\left\|S_{21}\right\|^4 W\left(W-4r_n\left(1-\left\|\Gamma_{on}\right\|^2\right)\right)$
P	$\left\|S_{21}\right\|^2+\left\|S_{11}\right\|^2-\left\|\Delta\right\|^2$
Q	$\left\|S_{21}\right\|^2+\left\|S_{22}\right\|^2-1$
W	$\left\|1+\Gamma_{on}\right\|^2\left(F_{min}-1\right)$
C_1	$S_{11}-S_{22}^*\left(S_{11}S_{22}-S_{12}S_{21}\right)$
Δ	$S_{11}S_{22}-S_{12}S_{21}$
r_n	Normalized noise resistance, R_n/Z_0
Γ_{on}	Optimum source reflection coefficient for noise match.

source reflection coefficient required to achieve $T_{CAS,min}$:

$$\Gamma_{OPT,T_{CAS}}=\frac{T_{CAS,min}\left|1+\Gamma_{on}\right|^2C_1^*+4\mathrm{T}_0r_n\left|S_{21}\right|^2\Gamma_{on}}{T_{CAS,min}\left|1+\Gamma_{on}\right|^2P+\mathrm{T}_0\left|S_{21}\right|^2\left(4r_n-W\right)}. \tag{7.16}$$

Now, equipped with equations (7.4)–(7.16), it is possible to determine the noise parameters, T_{CAS}, $T_{CAS,min}$, and $\Gamma_{OPT,T_{CAS}}$ directly from the dc and RF network parameters extracted in Chapters 5 and 6, allowing for a detailed comparison of the noise performance of state-of-the-art SiGe HBTs.

Furthermore, as $T_{CAS,min}$ is a function of both frequency and collector-current density, we can define its global minima for a given frequency as

$$T_{CAS,min,J_C}\equiv\min_{J_C\in[0,\infty)}T_{CAS,min}\left(J_C\right). \tag{7.17}$$

Finally, if we define the current density required to achieve T_{CAS,min,J_C} as $J_{C,min,min}$ then we can define the following quantities:

$$R_{min,min}\equiv R_{OPT,T_{CAS}}\left(J_{C,min,min}\right), \tag{7.18}$$

$$X_{min,min}\equiv X_{OPT,T_{CAS}}\left(J_{C,min,min}\right), \tag{7.19}$$

$$Q_{Z_{min,min}}\equiv\frac{X_{OPT,T_{CAS}}\left(J_{C,min,min}\right)}{R_{OPT,T_{CAS}}\left(J_{C,min,min}\right)}, \tag{7.20}$$

and

$$N_{min,min}\equiv N\left(J_{C,min,min}\right). \tag{7.21}$$

The discussion and understanding of these parameters and their temperature dependence is the topic of the following section.

7.4 Experimental Results

With the information obtained in Chapters 5 and 6, the noise parameters, $T_{CAS,min}$, and $\Gamma_{OPT,T_{CAS}}$ were computed from 50 MHz-40 GHz as a function of bias point for each of the seven devices described in Chapter 4. Once $T_{CAS,min}$ was known as a function of bias, its global minimum value, T_{CAS,min,J_C} was located at each frequency by fitting a quadratic to interpolated data in an area of the $T_{CAS,min}(f, J_C)$ versus J_C curve that was found to be near the location of the minima. The collector-current density at which the minima occurs, $J_{C,min,min}$, was also recorded as a function of frequency. The fitting and interpolation routine is used in order to generate smooth curves of T_{CAS,min,J_C} versus frequency (rather than discrete steps). Once T_{CAS,min,J_C} was determined, the associated optimum source impedances were determined.

In addition, the noise resistance, R_n at the optimum bias is also recorded with the assumption that the gain at the optimum bias point is sufficient to assure that $T_{min} \approx T_{CAS,min}$. An example plot demonstrating the difference between T_{min} and $T_{CAS,min}$ appears in Fig. 7.2. The data illustrates the difference between the room temperature value of $T_{CAS,min}$ and T_{min} as a function of bias for a ST-G4 device at 40 GHz. While the two curves stray at low-biases, in the vicinity of the minima of $T_{CAS,min}$, the values of $T_{CAS,min}$ and T_{min} are fairly similar indicating that the assumption made above will not lead to large errors. Thus, the sensitivity factor $N_{min,min}$ can be approximated as $R_n\left(J_{C,min,min}\right) R_{min,min} / \left(R^2_{min,min} + X^2_{min,min}\right)$.

For reference, the complete set of data (T_{CAS,min,J_C}, $J_{C,min,min}$, $R_{min,min}$, $X_{min,min}$, $N_{min,min}$, and $Q_{Z_{minmin}}$) is presented for each device in Figs. 7.11–7.17 at the end of the chapter (pp. 139–145). In the following sections, the modeled noise performance of the devices will be compared and analyzed at 18, 50, 77, 200, and 300 K ambient temperatures.

7.4.1 The Optimized Cascaded Noise Temperature, T_{CAS,min,J_C}

The optimized cascaded noise temperature,

$$T_{CAS,min,J_C}(f) \equiv \min_{J_C \in [0,\infty)} T_{CAS,min}, \tag{7.22}$$

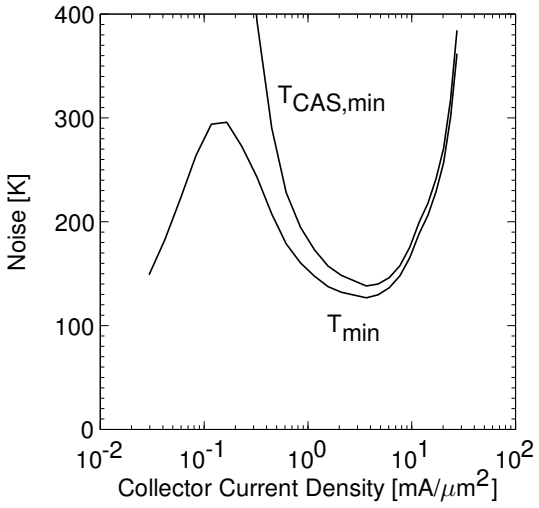


Figure 7.2: Comparison of T_{min} with $T_{CAS,min}$ for a ST-G4 device at 300K. The operating frequency is 40GHz.

is the global minima of $T_{CAS,min}$ as a function of bias[4], and is a very fundamental figure of merit for a low-noise amplifying device. To begin with, we will investigate how the optimized cascaded noise temperature compares among the different processes. This will be followed by a look into how this figure of merit behaves as a function of temperature. T_{CAS,min,J_C} was extracted for each of the devices at 300, 200, 77, 50, and 18 K, and the results appear in Fig. 7.3. Upon inspection of the data, several important observations can be made:

1) The noise performance of all the devices is greatly enhanced due to cooling. This is not a surprise as we saw in Chapters 5 and 6 that β_{DC}, f_t, and g_m all increase significantly with cooling.

2) In the low-GHz range, where the noise is dominated by β_{DC} and n_{cx}, the noise performance varies greatly from device to device. Furthermore, the modeled noise performance in this frequency range is fairly consistent[5] with the values calculated from dc measurements in Section 5.1.

3) At cryogenic temperatures, the noise performance of all of the devices converges above 10 GHz (with the exception of the JAZZ-G3 device).

4) In the low GHz frequency range, ST achieves far better noise performance than their competitors. This is due to the fact that their devices demonstrate an order of magnitude higher β_{DC}

[4] $T_{CAS,min}$ is the minimum value of T_{CAS} with respect to source impedance.
[5] The values in Section 5.1 tend to underestimate the noise by 10–15% due to the omission of the thermal noise due to the base and emitter resistances. This effect is less important at 18 K.

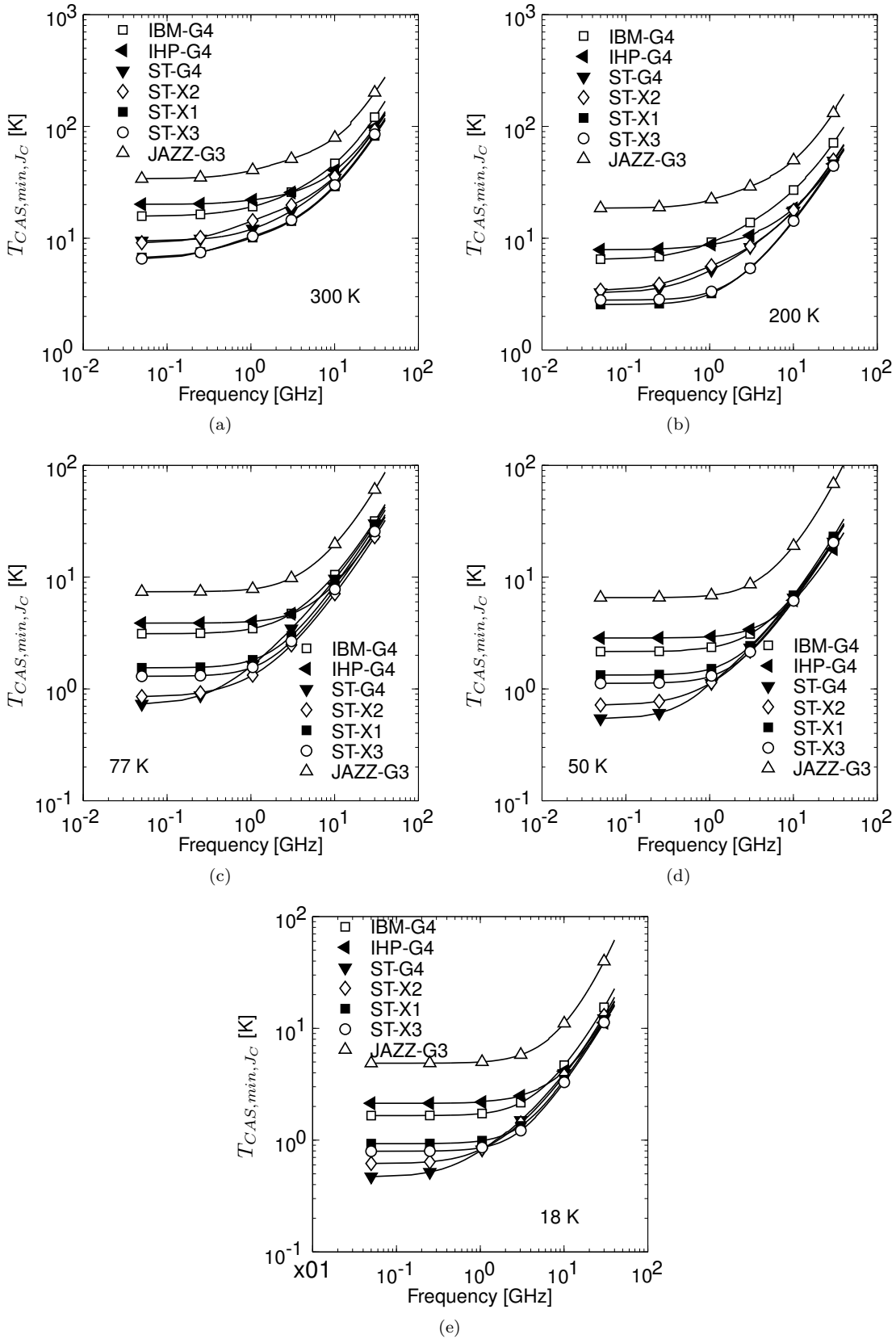


Figure 7.3: T_{CAS,min,J_C} as a function of frequency at (a) 300 K, (b) 200 K, (c) 77 K, (d) 50 K, and (e) 18 K

than any of the other devices.

5) Below 2 GHz, the ST-G4 device demonstrates a minimum cascaded noise temperature of less than 1 K at 18 K physical temperature. Furthermore, at 1 GHz, T_{CAS,min,J_C} is 0.69 K, which is within 0.3 K of a state-of-the art InP HEMT having no gate leakage [39].

6) The JAZZ-G3 device is consistently worse than the other devices. This is expected as the JAZZ device is from an older technology node and has lower β_{DC} and f_t than the other devices.

In addition to the raw noise performance, it is important to understand how things change quantitatively with temperature. To do this, we can use an analogous method to that used in Section 5.1.3, by defining the figure of merit

$$\Delta T_{CAS,min,J_C}(T_a) \equiv \frac{T_{CAS,min,J_C}(T_a)}{T_{CAS,min,J_C}(300\ K)}. \tag{7.23}$$

In Fig 7.4, $\Delta T_{CAS,min,J_C}$ is plotted as a function of ambient temperature for the seven devices at 1, 5, 20, and 40 GHz. Several key points can be highlighted:

1) With cooling down to 200 K, the average value of $\Delta T_{CAS,min,J_C}$ is 2.5 at 1 GHz, 2.1 at 5 GHz, and 1.8 at 20 GHz and 40 GHz. Thus the use of a thermoelectric cooler can improve the noise performance of a SiGe HBT by a factor of about two.

2) With cooling down to 77 K, liquid nitrogen temperature, the average value of $\Delta T_{CAS,min,J_C}$ is 6.7 at 1 GHz, 5 at 5 GHz, 3.7 at 20 GHz, and 3.4 at 40 GHz. Thus the effect of cooling to 77 K is significantly greater on the low frequency end, where the enhancement is boosted by large improvements in β_{DC} with cooling. At the high-frequency end, the effect is roughly equal to the drop in ambient temperature, which is not surprising as the thermal contributions due to the losses have decreased.

3) With cooling down to 50 K, which can be reached using inexpensive sterling coolers, the average value of $\Delta T_{CAS,min,J_C}$ is 8.6 at 1 GHz, 6.5 at 5 GHz, 4.7 at 20 GHz, and 4.3 at 40 GHz. Once again, the improvement is much greater on the low-frequency end, where a large increase β_{DC} has helped significantly.

4) With cooling all the way to 18 K, the average value of $\Delta T_{CAS,min,J_C}$ is 11.9 at 1 GHz, 10.4 at 5 GHz, 7.8 at 20 GHz, and 6.9 at 40 GHz.

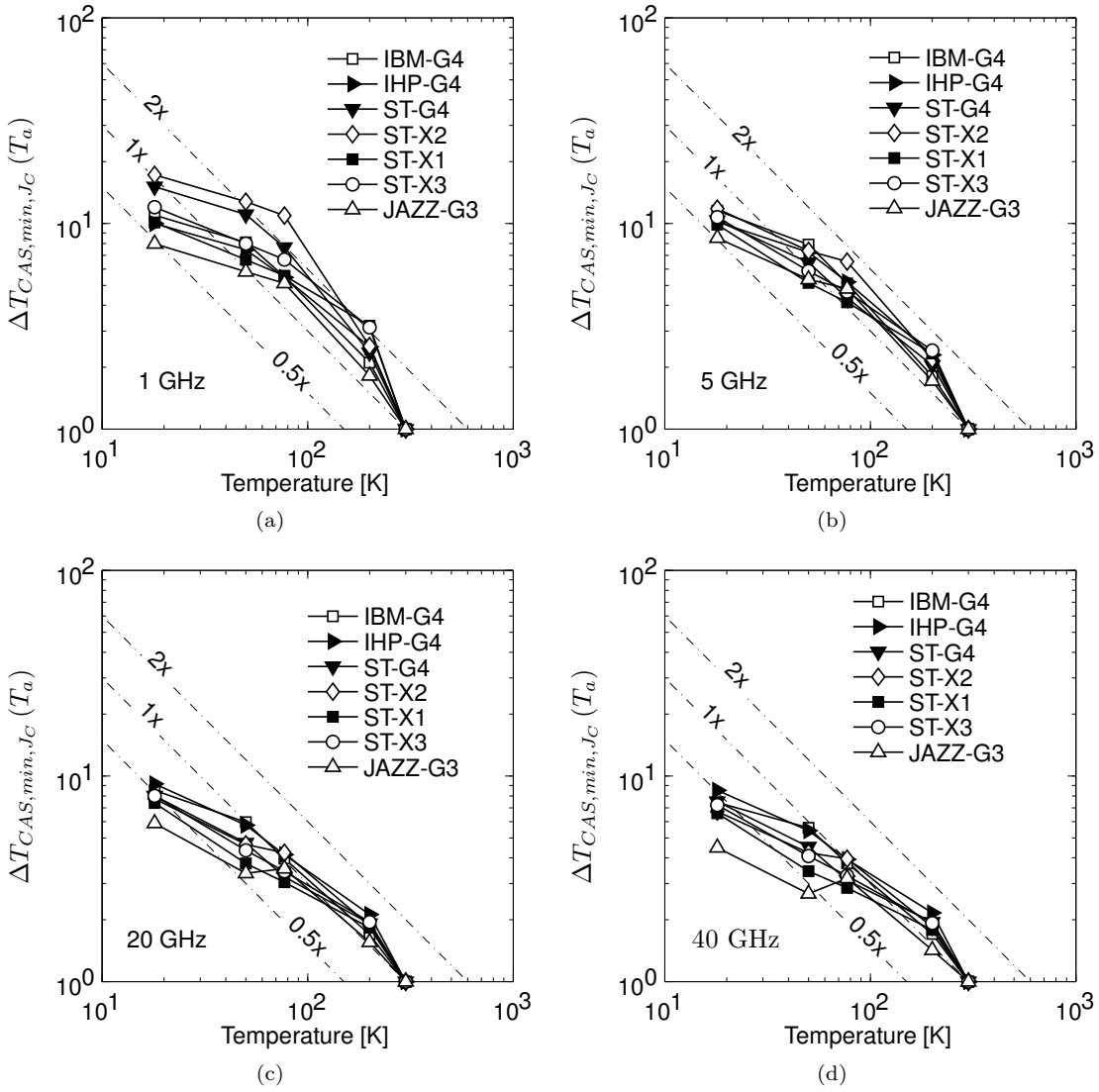


Figure 7.4: The figure of merit $\Delta T_{CAS,min,J_C}(T_a)$ plotted as a function of temperature at (a) 1 GHz, (b) 2 GHz, (c) 20 GHz, and (d) 40 GHz. Also included in the plots are contour lines marked "0.5X", "1X", and "2X." The interpretation of these contour lines is as follows: If the data is above the line marked "1X", this means that T_{CAS,min,J_C} has dropped by a factor greater than the factor of change in ambient temperature. Similarly, if the data is above the line marked "2X", this means that T_{CAS,min,J_C} has dropped by a factor greater than twice the factor of change in ambient temperature. The extension to the line marked "0.5X" is obvious.

5) The standard deviation of the values of $\Delta T_{CAS,min,J_C}$ is significantly smaller at higher frequencies than at lower frequencies. This can be attributed to the much wider variation in $\partial\beta_{DC}/\partial T_a$ than in $\partial f_t/\partial T_a$ and $\partial n_c/\partial T_a$.

7.4.2 Optimum Bias Point, $J_{C,min,min}$

The frequency dependence of the current density required to achieve the optimum cascaded noise temperature for a SiGe can be understood in terms of $T_{min}@J_{C,min,min}$, which is a very good approximation of T_{CAS,min,J_C}, so long as the frequency is low enough that the amplifier has at least 10 dB of available gain. The expression for the minimum noise temperature as a function of source impedance was originally given in Chapter 2 and, after some minor manipulations, is expressed here as

$$T_{min} \approx T_a \sqrt{\frac{1}{\beta_{DC,eff}} + \left(\frac{f}{f_{t,eff}}\right)^2}, \tag{7.24}$$

where $\beta_{DC,eff} = \beta_{DC} / \left(n_c^2 + 2g_m n_c \left(r_b + r_e\right)\right)$ and $f_{t,eff} = f_t / \sqrt{2g_m n_c \left(r_b + r_e\right)}$ represent effective degradations to the dc current gain and the unity-current gain cutoff frequency due to the base and emitter resistances. Referring to equation (7.24), there is a clear trade-off to be made between $\beta_{DC,eff}$ and $f_{t,eff}$, as the minimization of the noise with respect to bias point requires finding the optimum balance between the frequency independent term and the frequency dependent term.

To complicate matters, the collector-current density required for peak f_t in modern SiGe HBTs is on the order of 10 mA/μm^2 or higher, which is in the region where the collector-current ideality factor is quite large (i.e., the collector current slope is far lower than the ideal value predicted by the classical drift-diffusion equations). Thus, even if the collector-current density required to achieve $\beta_{DC,pk}$ coincides with that required for peak $f_{t,pk}$, the optimum bias point for noise in the low-frequency range will still be different than that in the high frequency range due to the detrimental effects of the steep rise in n_c above 1 mA/μm^2. Thus, there will always be a tradeoff between the low and high frequency noise in the selection of the bias point.

To assist in gaining an intuitive understanding of the nature of this trade-off and how it varies with temperature, the factors $1/\beta_{DC,eff}$ and $\left(f/f_{t,eff}\right)^2$ are shown in Fig. 7.5(a) for the frequency dependent optimum collector current density $J_{C,min,min}\left(f\right)$. In addition, the normalized values of $\beta_{DC,eff}$ and $f_{t,eff}$ are shown in Fig. 7.5(b), and the normalized values of β_{DC} and f_t are shown in Fig. 7.5(c). The information contained in these curves is quite revealing. First of all, as expected, in the sub-GHz range, the noise is primarily determined by the dc current gain. However, inspection of Figs. 7.5(b) and 7.5(c) reveals a very interesting feature of the cryogenic device operation in the 0.1-1 GHz range; while the device is biased near $\beta_{DC,pk}$, $\beta_{DC,eff}$ is significantly lower, and effectively cancels out the increase in β_{DC} which has occurred with cooling[6].

[6]The value of $\beta_{DC}@J_{C,min,min}$(150 MHz) at 18 K was found to be 40 times larger than at 300 K. However, at the same bias point, the 18 K value of $\beta_{DC,eff}$ is reduced from β_{DC} by a factor of 38. Fortunately, the temperature has dropped substantially, resulting in an overall improvement of over 16.

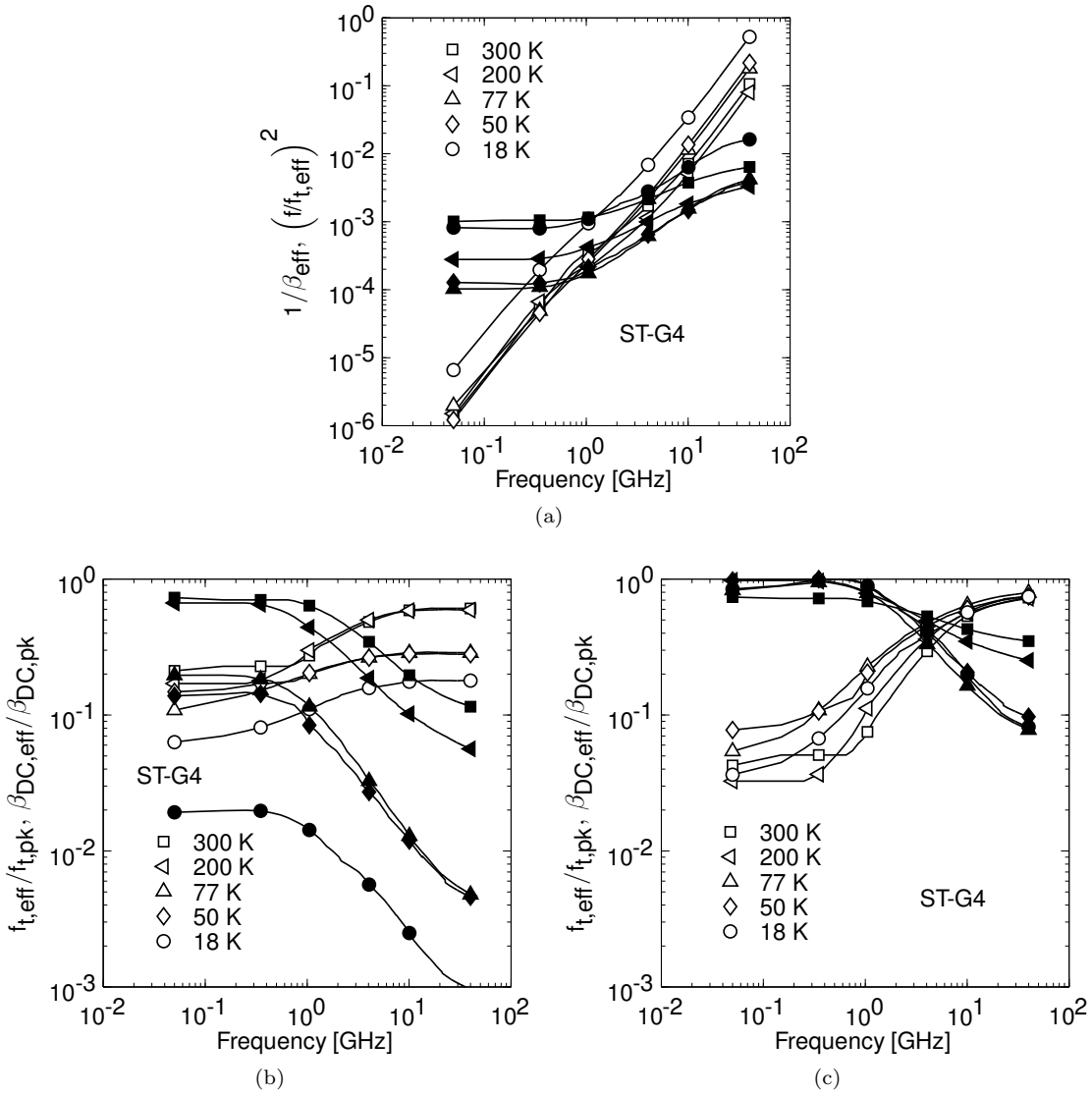


Figure 7.5: Illustration of the tradeoff between β_{DC} and f_t in the minimization of $T_{CAS,min}$ with respect to J_C. The solid markers are on the β_{DC} traces whereas the white markers are on the f_t traces.

In terms of the high-frequency noise performance, equally valuable information can be obtained from these curves. For instance, referring to Fig. 7.5(c), it can be seen that above 40 GHz, $J_{C,min,min}(f)$ moves towards the current-density required to achieve $f_{t,pk}$. However, inspection of Fig. 7.5(b) reveals that beyond 10 GHz, the increase in f_t results in diminishing returns in terms of noise performance as the normalized $f_{t,eff}$ curves flatten out. However, the reason that the optimum bias point continues to move towards peak f_t is that the criterion for bias selection is the minimization of $T_{CAS,min}$ as opposed to T_{min}. Thus, the increase in $J_{C,min,min}$ is simply occurring in order to improve the available gain. In addition to the saturation of $f_{t,eff}$, it is also notable that

its limiting value at cryogenic temperatures is reduced significantly due to the rise in the collector current ideality factor.

The optimum bias point required to achieve T_{CAS,min,J_C} for each of the seven devices is plotted as a function of frequency in Figs. 7.11(b)-7.17(b). In general, these curves follow what one would expect based upon the discussion above, with the bias at low temperatures being determined such that $\beta_{DC,eff}$ is optimized at low-frequencies and that $f_{t,eff}$ and G_a are optimized at high-frequencies. As a final note, it can be observed that $J_{C,min,min}$ (40 GHz) decreases significantly[7] for all the devices as the temperature is lowered from 300 to 18 K.

7.4.3 Optimum Source Impedance, $Z_{min,min}$

While closed form expressions were provided in Section 2.2 for Z_{OPT}, no closed form expressions exist for $J_{C,min,min}$. Thus, it is difficult to ascertain the exact behavior of $Z_{min,min}$ without resorting to numerical techniques, and the discussion regarding $Z_{min,min}$ will be restricted to modeling results. The optimum generator impedance ($\Gamma_{min,min}$) of the majority of the devices[8] appears in Fig. 7.6, and the real and imaginary components of the impedances appear in Figs. 7.11(c)–7.17(d). It is evident from these data that the optimum source resistance for cascaded noise temperature match decreases significantly with cooling. This is very important as it implies that a device sized for an LNA designed for operation at 300 K is not correctly sized for a noise match at 18 K. Thus, a SiGe amplifier cannot be simultaneously optimized for low-noise operation at 300 and 18 K.

To gauge the magnitude of the decrease in $R_{min,min}$ with temperature, the ratios of the 18 and 77 K values to the room temperature value were computed at 10 GHz for each of the devices; the results appear in Fig. 7.7. At 18 K, an average decrease in $R_{min,min}$ of 50% from the room temperature value was observed, with quite low variation among the different devices (σ = 5%). Thus, on average, to achieve a given value of $R_{min,min}$ at 18 K one would need to use a device half the size of that which would be used at room temperature. However, at 77 K, the results are closer to the room temperature values; the average drop in $R_{min,min}$ is only 12%. Thus, as the effect is so much more pronounced at 18 K, it is believed that it is associated with the collector current ideality factor[9].

In addition to $R_{min,min}$, we are also interested in the behavior of $X_{min,min}$ as a function of temperature as it indicates what kind of reactance will have to be tuned out in order to achieve

[7] The average decrease was 40%.

[8] The ST-X3 device was left out due to the fact that its associated result was indistinguishable from the result form the ST-X1 device.

[9] It should be noted that this conclusion is consistent with equation (2.35) in Section 2.2.

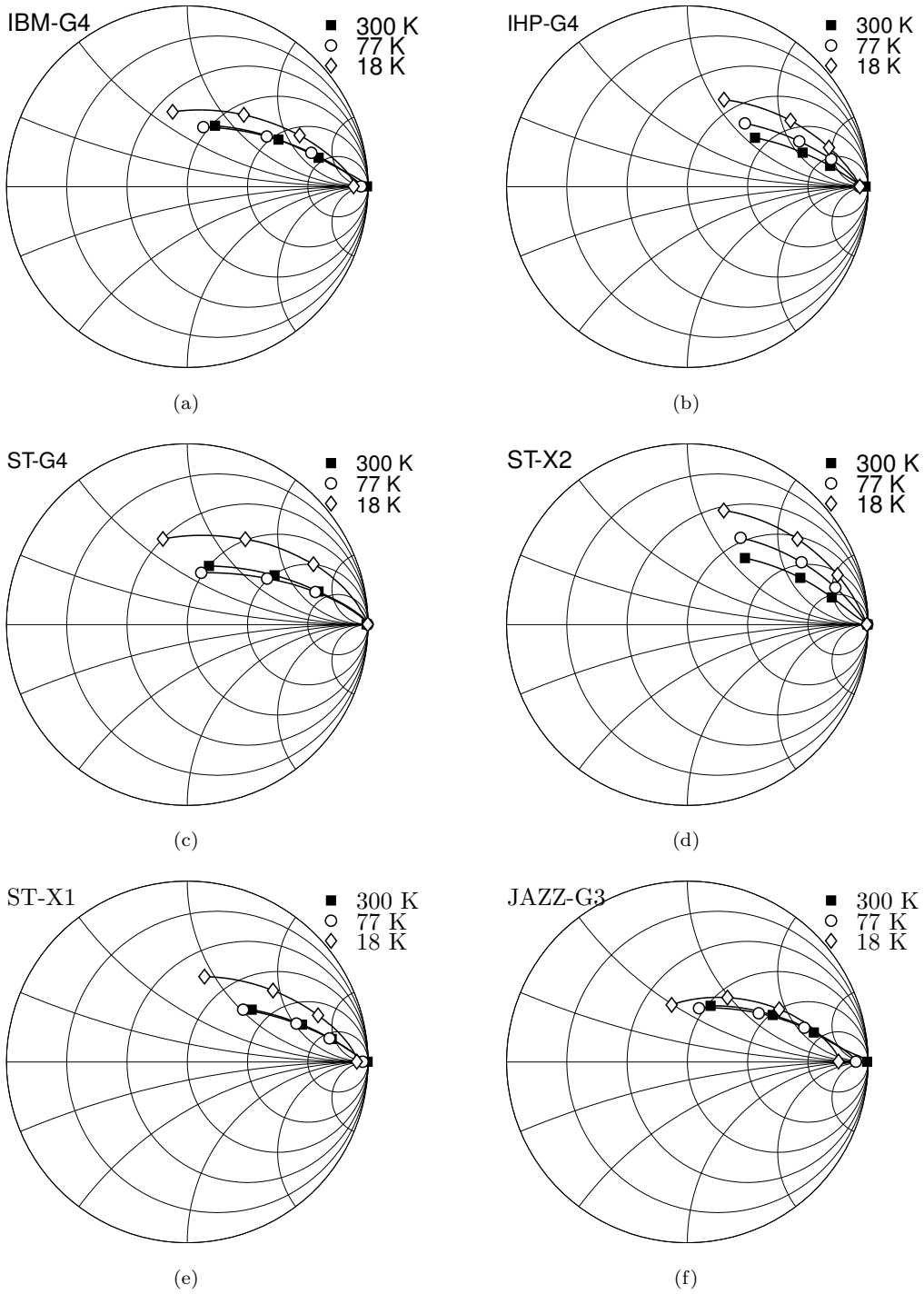


Figure 7.6: $\Gamma_{min,min}$ for the each of the studied devices at 300, 77, and 18 K. The results for the ST-X3 device are indistinguishable from the results for the ST-X1 device. Therefore, only the ST-X1 device is included in the figure

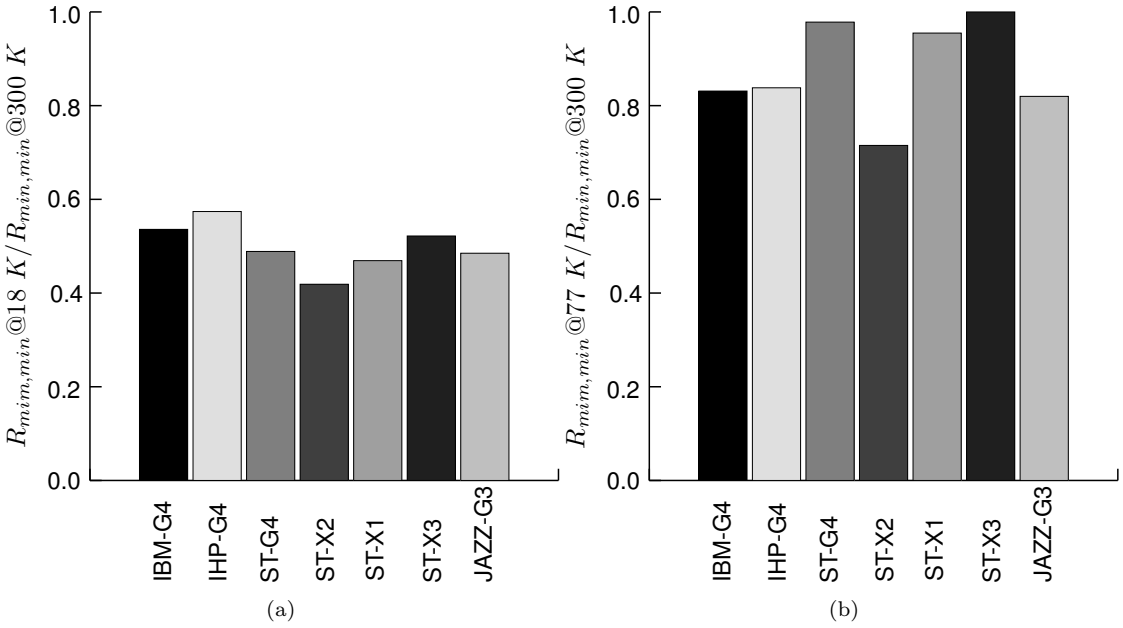


Figure 7.7: The ratio of change in $R_{min,min}$ with cooling. (a) T_a=18 K and (b) T_a=77 K

a noise match. Referring to Figs. 7.11(d)–7.17(d), we can see that $X_{min,min}$ is only sensitive to temperature below about 5 GHz. To understand this result, we can begin by reviewing the equation for the optimum source reactance for noise match:

$$X_{OPT} \approx \frac{1}{g_m} \frac{f/f_t}{1/\beta_{DC} + (f/f_t)^2}. \tag{7.25}$$

At low-frequencies, X_{OPT} depends upon the inverse of the transconductance. As not much transconductance is needed to provide considerable gain below 1 GHz, it is likely that the variation in the low frequency value of $X_{min,min}$ is related to low transconductances. To test this hypothesis, the 50 MHz values of $X_{min,min}$ taken at 200 K and below were normalized to the room temperature value and plotted against the normalized transconductance and a linear fit was performed. The results appear in Fig. 7.8 and seem to indicate that the variation in low-frequency $X_{min,min}$ is in fact correlated to variation in g_m.

7.4.4 The Sensitivity Factor, $N_{min,min}$

The sensitivity factor is defined in terms of the optimum noise match, not the match for optimum cascaded noise temperature. Thus, in the discussion to follow, it is assumed that the gain of the network is high enough to justify assuming $Z_{min,min} = Z_{OPT}$. The sensitivity factors were calculated

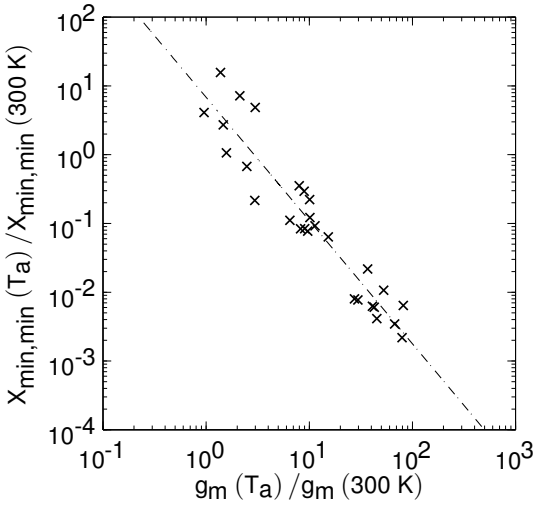


Figure 7.8: Normalized $X_{min,min}$ versus normalized g_m

and appear in Figs. 7.11(e)–7.17(e). As discussed in [69], in order for a set of noise parameters to be physical, the following must hold: $1 < 4NT_0/T_{min} < 2$. Thus, if the quantity $4NT_0/T_{min}$ is close to unity, then the interpretation is that the network is insensitive to source match, whereas if the quantity is close to two, the opposite is true. Therefore, this ratio is quite important, and thus has been plotted in Fig. 7.9. As discussed in Section 2.2, the first-order calculation predicts that $N \approx T_{min}/2T_0$. Referring to the modeled result, we see that this is true at low frequencies, but loses validity at higher frequencies, where the sensitivity factor begins to roll-off. Nonetheless, it is a fairly reasonable estimation. Referring to Fig. 7.9, we also see that the quantity $4NT_0/T_{min}$ is only weakly dependent upon temperature.

7.4.5 Comparison with State-Of-The-Art InP

Indium-phosphide (InP) high-electron-mobility transistors (HEMTs) are the gold standard in terms of semiconductors used in extremely low-noise cryogenically cooled low-noise amplifiers [6]. In this section, a comparison will be made between the cryogenic noise performance of state-of-the-art InP HEMTs with that of the SiGe HBTs modeled in this report. The InP devices serving as a reference in this comparison are from the Northrup Grumman Space Technology (NGST) 0.1μm InP HEMT process line. These are the same devices used in many of the world's most advanced cryogenic LNAs, such as those reported in [6] and [10]. The model used in this comparison was developed by N. Wadefalk and is for a $0.1 \times 4 \times 50\mu\text{m}^2$ HEMT operating at 15 K [148]. The topology of the model is that proposed by Pospieszalski [25]. This is the same model that was used in the design of the

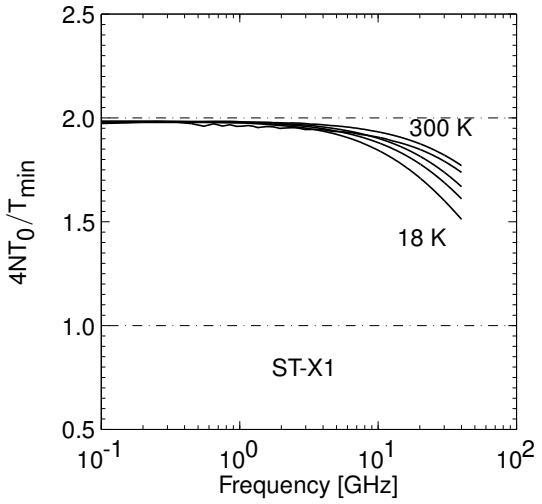


Figure 7.9: $4N_{min,min}T_0/T_{min}$ as a function of frequency. This ratio must lie in the range of 1–2.

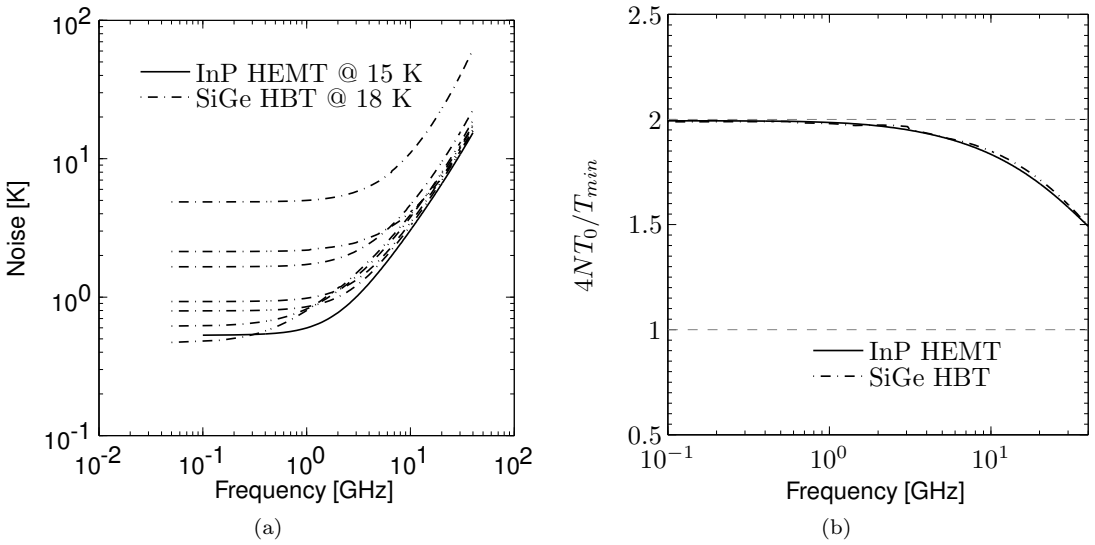


Figure 7.10: Comparison of SiGe HBT and InP HEMT noise at 18 K. (a) The minimum noise temperature. For the case of the SiGe curves, the minimum cascaded noise temperature has been plotted. (b) The ratio of $4NT_0/T_{min}$. This ratio must lie in the range of 1 and 2. It is assumed that the HEMT has 50 nA gate leakage current. In addition, it is assumed that the SiGe device is at the optimum bias at each frequency whereas the InP device is plotted at a fixed bias.

amplifier reported in [10].

A comparison of the minimum noise of the SiGe devices with that of an InP HEMT with 50 nA of gate leakage current appears in Fig. 7.10(a). While the InP device is clearly superior to the SiGe devices in terms of noise performance, the difference is not all that substantial. For instance, the InP device has a minimum noise temperature of around 15 K at 40 GHz whereas the better

SiGe devices have T_{CAS,min,J_C} values of around 20 K at this frequency. Furthermore, this is a pessimistic comparison as the noise of the SiGe devices has been overestimated due to the omission of correlation between the shot-noise sources in the SiGe noise model. In addition, the sensitivity factors of a typical SiGe device and the InP device are plotted in Fig. 7.10(b). Clearly, the sensitivity factor is quite similar for the two breeds of devices.

However, what is interesting to note is that at 40 GHz the transconductance of an optimally sized SiGe device is approximately 50% higher than that of the InP device[10]. Thus, the optimum SiGe device provides considerably more available gain than the optimum InP device at 40 GHz. Furthermore, the size of the SiGe devices are considerably smaller; the ideal InP device at 40 GHz is 0.1×60 mA/μm^2 whereas the optimum ST-G4 device is 0.13×3mA/μm^2 (the optimum sizing of all of the SiGe devices is in this range). This in combination with the fact that SiGe devices generally come in a process in which CMOS FETs are also available helps to build a compelling argument for using SiGe HBTs over InP HEMTs.

7.5 Summary

In this chapter, a comprehensive treatment of the cryogenic noise performance of SiGe HBTs has been presented. The discussion began with an introduction to noise parameters and the concept of minimum cascaded noise temperature. Next, a procedure by which one can compute noise performance from dc and small-signal parameters was presented. After the presentation of background information was complete, the cryogenic noise performance of a variety of devices was discussed in detail. In the next chapter, a brief summary of all that has been learned about how the properties of SiGe HBTs depend upon temperature will be presented, and then, in Chapter 9 we will forge ahead with some real world applications of SiGe HBTs.

[10]Interestingly enough, this is true for all of the devices modeled in this work.

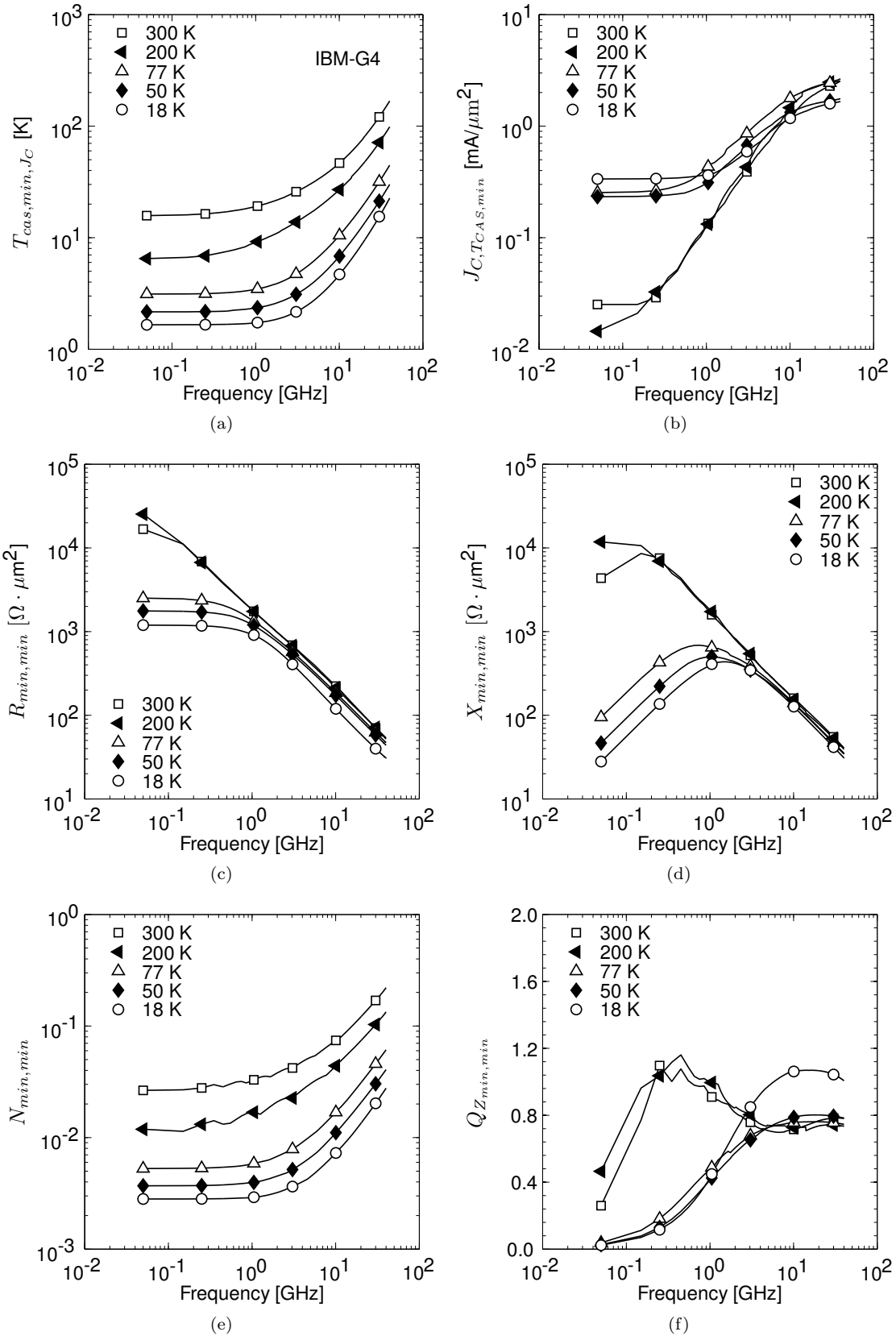


Figure 7.11: IBM-G3 noise properties

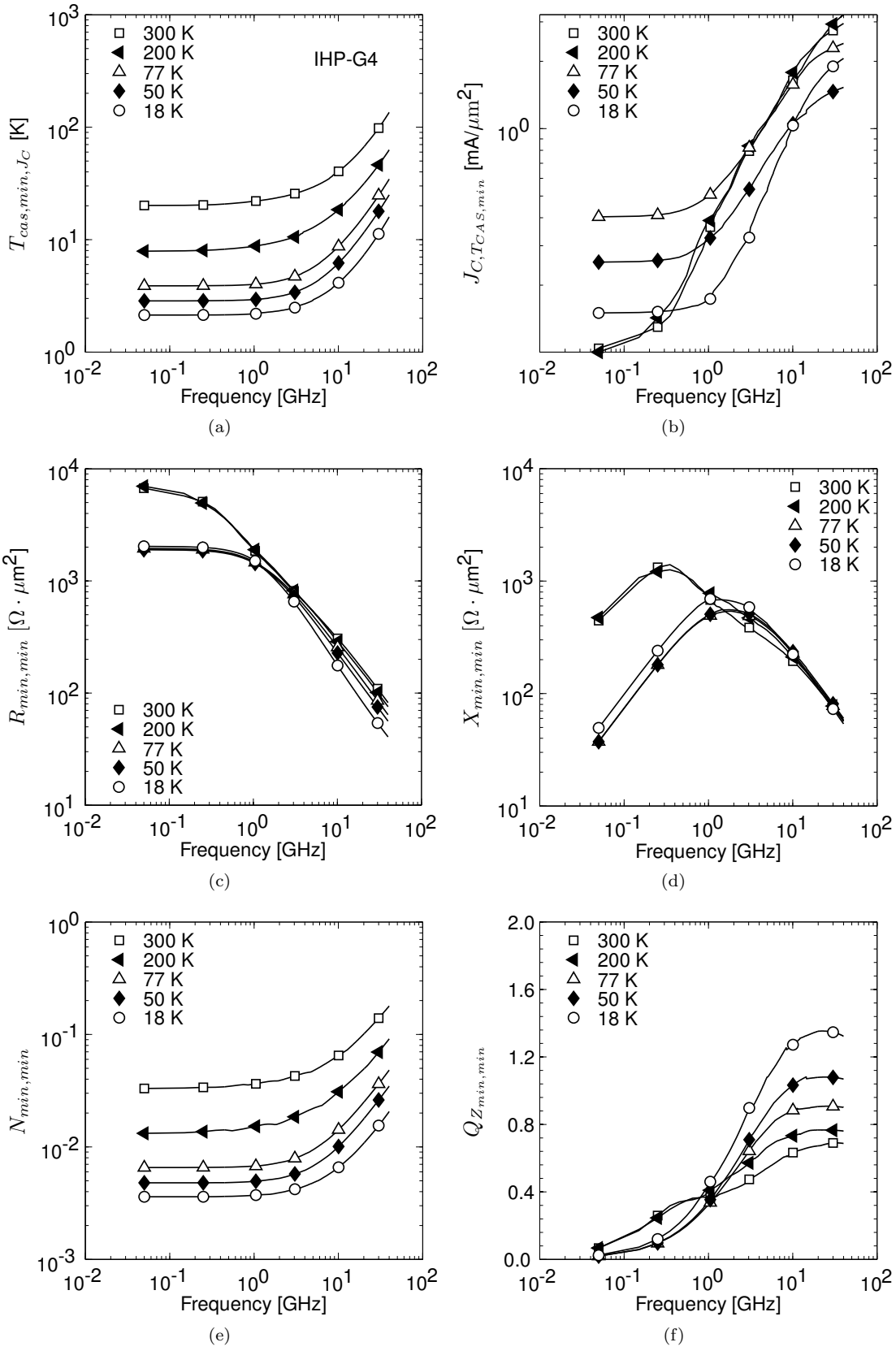


Figure 7.12: IHP-G4 noise properties

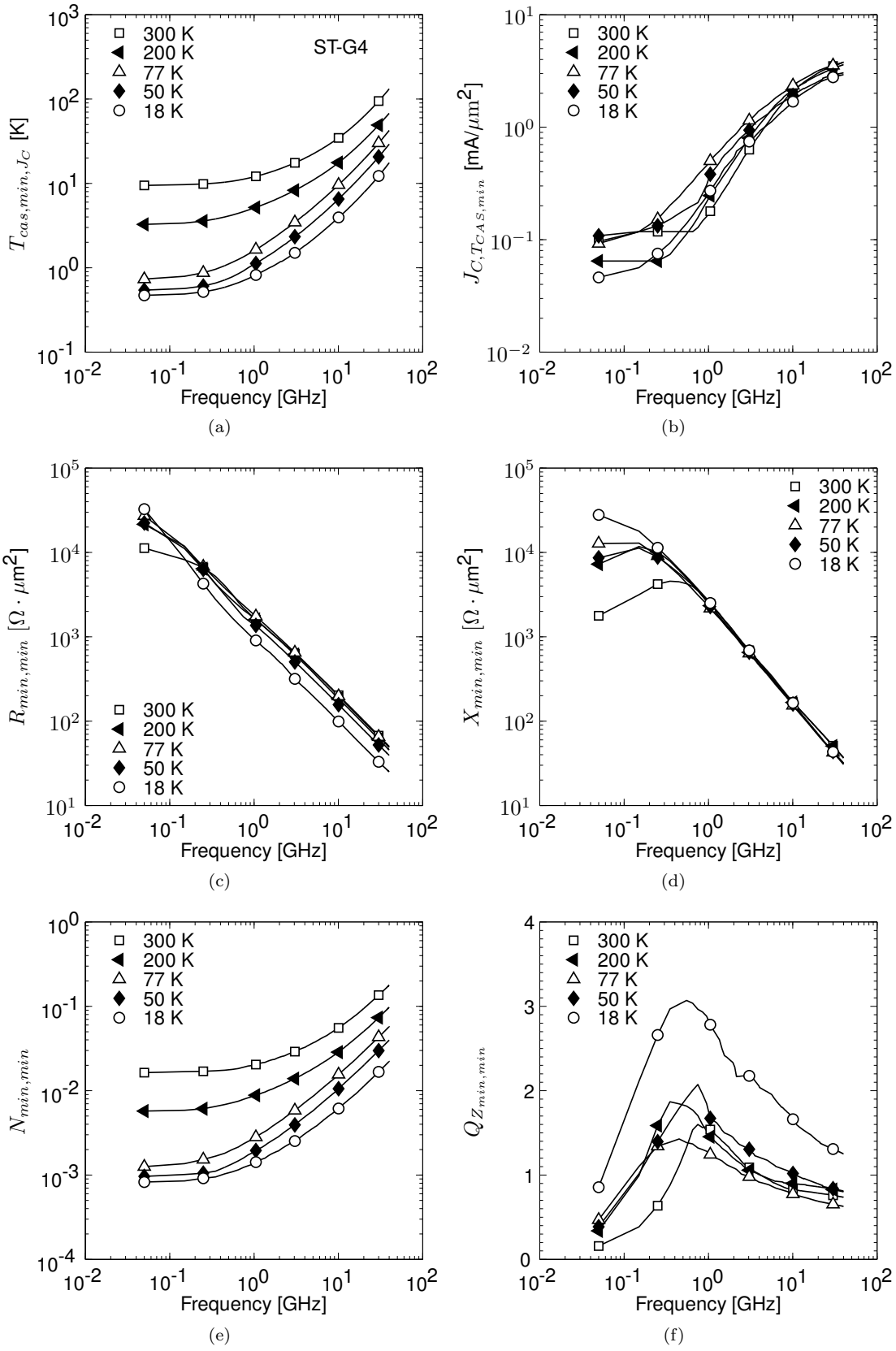


Figure 7.13: ST-G4 noise properties

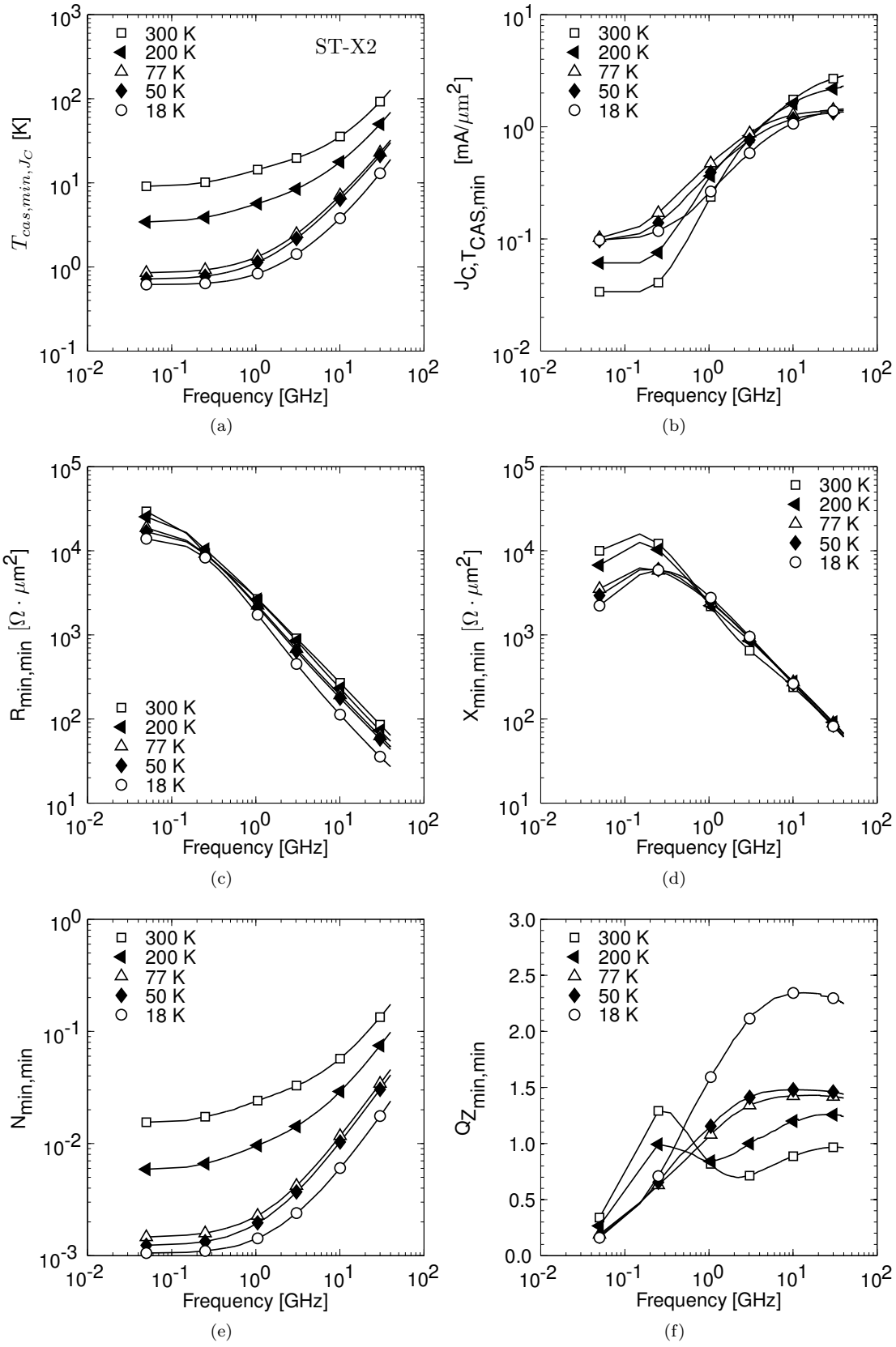


Figure 7.14: ST-X2 noise properties

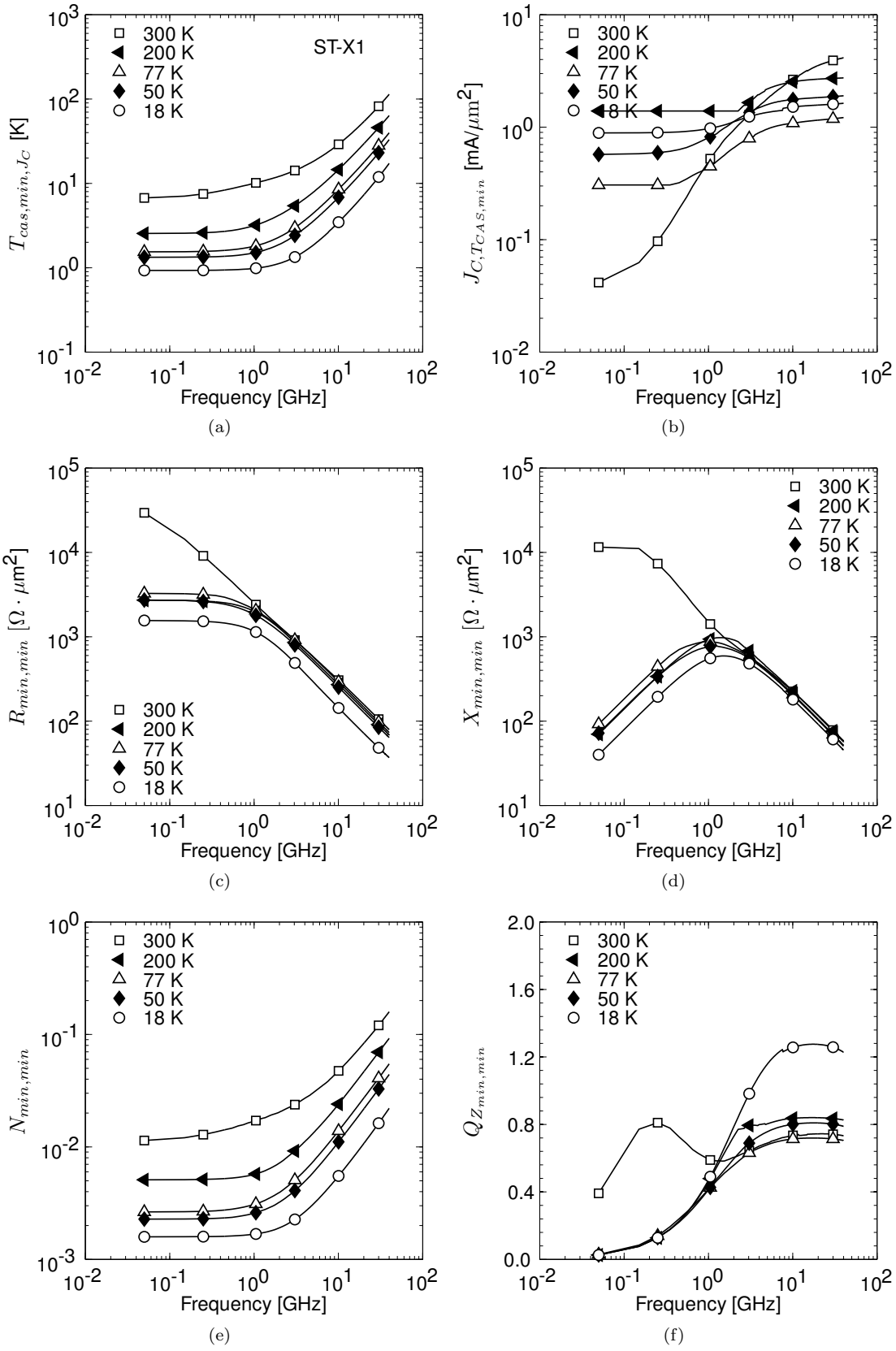


Figure 7.15: ST-X1 noise properties

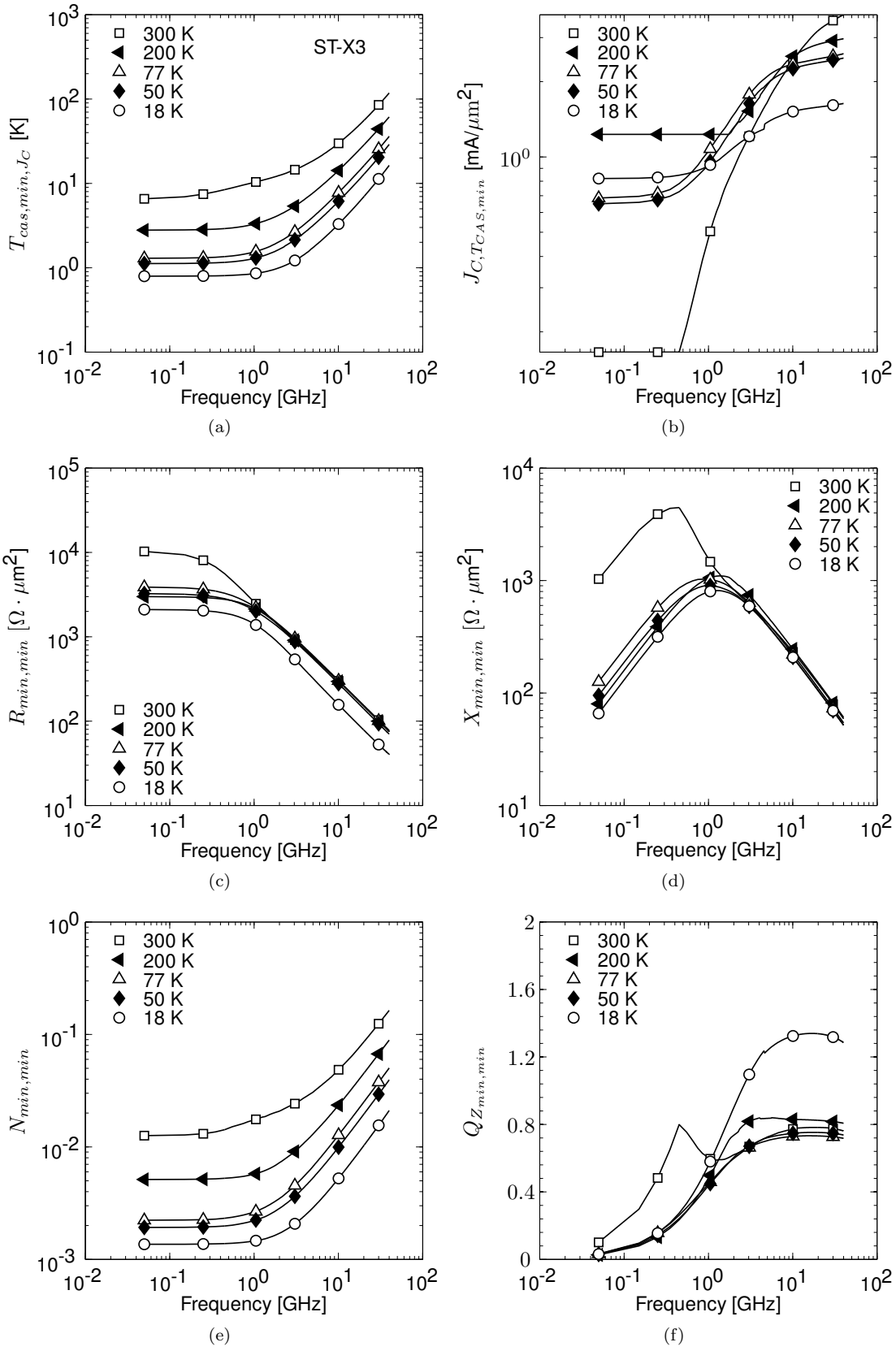


Figure 7.16: ST-X3 noise properties

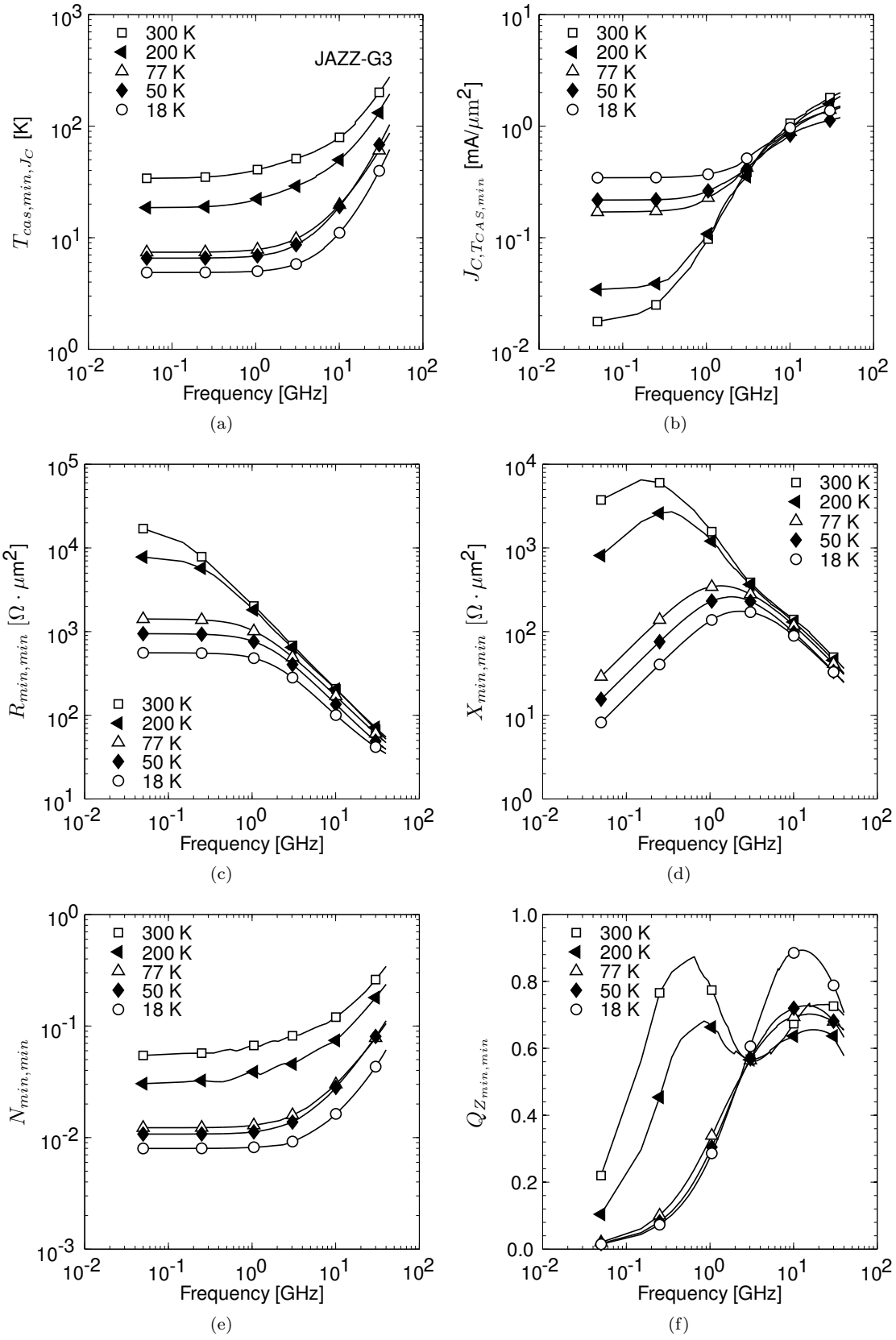


Figure 7.17: JAZZ-G3 noise properties

Chapter 8

Summary of Changes that Occur in SiGe HBTs with Cooling

In the previous three chapters, a large amount of data has been presented and analyzed and the purpose of this chapter is to provide a succinct recapitulation of the key changes that occur in SiGe HBTs with cooling. This information is provided in tabular format below.

Table 8.1: Summary of SiGe HBT temperature behavior

DC Parameters		
Parameter	Temperature Dependence	Comments
n_c	$\sim 220\%$ increase	For cooling from 300 to 18 K. High variability among the different processes ($\sigma = 68\%$). Increase presumably due to ballistic transport phenomena.
J_{B0}	$7.83 \times 10^{18} e^{-E_g(T_{eff,h})/kT_{eff}} J_{B0}$ (300 K)	Experimentally observed. The explanation is still an open issue.
n_b	$\sim 390\%$ increase	For cooling from 300 to 18 K. High variability among the different processes ($\sigma = 80\%$). Physical cause for increase still an open issue.
$\beta_{DC,pk}$	$20 - 3400\%$ increase	For cooling from 300 to 18 K. Tremendous variability between processes due to exponential dependence of Ge content on neutral base side of B-E SCR.
RF Figures of Merit		
Parameter	Temperature Dependence	Comments
$f_{t,pk}$	$\sim 50\%$ increase	For cooling from 300 to 18 K. Consistent from one process to another ($\sigma = 6\%$). Not much change below 77 K, presumably due to the fact that peak f_t occurs at a very high collector current density where considerable self heating is likely.

$f_{max,pk}$	$\sim 35\%$ increase	High level of variation among processes ($\sigma = 15\%$)

Small-Signal Model Parameters

Parameter	Temperature Dependence	Comments
r_b	$\sim 26\%$ decrease	For cooling from 300 to 18 K. J_C =1mA/μm^2. Varies strongly from foundry to foundry ($\sigma = 20\%$).
r_e	$\sim 25\%$ increase	For cooling from 300 to 18 K. Varies strongly from foundry to foundry ($\sigma = 55\%$). Increase may have to do with emitter annealing process.
r_c	$\sim$50% decrease	For cooling from 300 to 18 K. Varies moderately from foundry to foundry ($\sigma = 10\%$). Does not seem to depend upon style of sub-collector.
g_m	$\sim 200\%$ increase	J_C =1 mA/μm^2. For cooling from 300 to 18 K. Varies significantly from foundry to foundry ($\sigma = 65\%$).
g_{be}	$\sim 84\%$ decrease	J_C=1 mA/μm^2. For cooling from 300 to 18 K. Fairly consistent ($\sigma = 9\%$).
τ_d	-	
C_{BE}	$\sim 68\%$ increase	For cooling from 300 to 18 K. For J_C =1mA/μm^2. Mainly attributed to change in diffusion capacitance due to increase in g_m. Moderate variation from process to process ($\sigma = 19\%$)
C_{CB}	$\sim$10% decrease	For cooling from 300 to 18 K. Attributed to a decrease in the concentration of ionized impurities in the collector.
C_{CS}	$\sim 44\%$ decrease	For cooling from 300 to 18 K. V_{CS}=1 V. Attributed to substrate freeze-out. It is believed that the residual capacitance is due to de-embedding errors.

Noise Parameters

Parameter	Temperature Dependence	Comments
$T_{cas,min}$	$\sim 91\%$ decrease	For cooling from 300 to 18 K. 1.5 GHz. Small variation from foundry to foundry ($\sigma = 2.5\%$)
	$\sim 83\%$ decrease	For cooling from 300 to 18 K. Average value from 1-40 GHz. Small variation from foundry to foundry ($\sigma = 3.8\%$)
$J_{C,min,min}$	$\sim 25\%$ decrease	For cooling from 300 to 18 K. For noise optimization at 40 GHz.
R_{opt}	$\sim 50\%$ decrease	For cooling from 300 to 18 K. Average value from 1-40 GHz. Moderate variation from foundry to foundry ($\sigma = 10\%$).

X_{opt}	$\sim 18\%$ decrease	For cooling from 300 to 18 K. Average value from 1-40 GHz. Significant variation from foundry to foundry ($\sigma = 16\%$)
$Q_{Z_{OPT}}$	$\sim 84\%$ increase	For cooling from 300 to 18 K. Frequency is 10 GHz. Inconsistent across foundries ($\sigma = 44\%$).
N	$\sim 89\%$ decrease	For cooling from 300 to 18 K. Frequency is 10 GHz. Very consistent across foundries ($\sigma = 1.4\%$).

Part III

Applications

Chapter 9

Cryogenic Low-Noise Amplifiers

In the preceding sections of this dissertation, a framework has been built upon which to understand and predict the performance of SiGe HBTs at cryogenic temperatures. In this section, this framework is utilized to facilitate the design of low-noise amplifiers with state-of-the-art noise performance. Several amplifiers are presented with topologies ranging from discrete transistor amplifiers to differential integrated-circuit feedback amplifiers. For each of these cases, simulation is compared with measurement in order to validate the noise modeling procedure described previously.

9.1 Small-Signal Compact Noise Models

One difficulty in the design of integrated circuits intended to operate at cryogenic temperatures is that transistor models in this temperature range are not available from the foundry. While the small-signal models developed in Chapters 6 and 7 are certainly a good place to start in the development of simulation tools aimed at cryogenic applications, the fact that each bias point has a separate model is certainly a cumbersome feature that would ideally be avoided.

Referring back to Fig. 7.1, it can be seen that the model being used to describe the small-signal and noise performance of the devices is actually quite simple; at a given bias point, 9 parameters are used to describe the small-signal terminal current-voltage relationships and only two additional dc currents are needed to determine the magnitude of the shot-noise sources. Furthermore, because of a lack of redundancy in the model, we expect that the component values can be uniquely determined leading to a smooth variation in extracted parameters as a function of bias. This expectation has been corroborated by the data presented earlier. Thus, it is possible to fit smooth curves to the small-signal model parameters as a function of collector current density.

Using the data extracted in Chapters 5-6, parameterized small-signal models were generated for

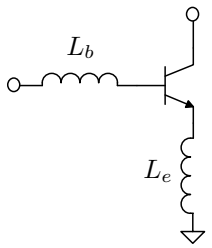


Figure 9.1: Basic LNA with emitter degeneration

several of the devices by fitting smooth curves to each of the small-signal component values as a function of collector current density. The models were then entered into AWR's Microwave Office (MWO) simulation environment and scalability was added to the models. The resulting small-signal models are parameterized in terms of collector current density and device area, and are very powerful in the design of an amplifier as they allow one to quickly evaluate how changes in device size and bias point affect the response of a circuit.

9.2 Inductively Degenerated Amplifiers

A very common technique used to achieve simultaneous input and noise match is to use inductive emitter-degeneration as shown in Fig. 9.1. It can be shown that the small-signal input impedance of an inductively degenerated transistor is

$$Z_{in} = \left(j\omega L_e + \frac{g_m L_e}{C_{BE}} - \frac{j}{\omega C_{BE}} \right) \frac{1 + j\omega C_{CB} Z_L}{(1 + j\omega C_{CB} Z_L) + \frac{C_{CB}}{C_{BE}} (1 + g_m Z_L + j\omega g_m L_e - \omega^2 L_e C_{BE})}$$

$$\approx j\omega \widetilde{L}_e + \frac{g_m L_e}{\widetilde{C}_{IN,m}} - \frac{j}{\omega \widetilde{C}_{IN,m}}, \tag{9.1}$$

where Z_L is the load impedance, $\widetilde{C}_{IN,m} = C_{BE} + C_{CB}(1 - A_v)$ is input capacitance accounting for the Miller effect [149], and $\widetilde{L}_e = L_e/(1 - A_v)$ is a reduction in the equivalent value of L_e due to the Miller effect. In practice, a cascoded input stage is commonly used to isolate the input from the output, thereby reducing the Miller effect tremendously. In this case, $\widetilde{C}_{IN,m} \approx C_{BE}$, and the design procedure is greatly simplified. However, the discussion here will be kept general so as to apply to amplifiers without cascoded input stages.

To begin with, let us consider equation (9.1), in which it can be seen that the introduction of L_e to the circuit has two effects: 1) the input reactance is increased by $\omega \widetilde{L}_e$; and 2) a real part of the input impedance, with value $g_m L_e/\widetilde{C}_{IN,m}$, is generated. In addition, it can be shown that the

addition of a lossless inductance to the input loop does not have much effect upon the noise [30].

The procedure for designing a fully noise-optimized and power-matched LNA using emitter degeneration is as follows:

1) If device models are unavailable, generate small-signal noise models over a wide range of bias conditions following the methodology presented in Chapters 5–7.

2) Determine the collector current density, $J_{C,OPT}$, which corresponds to the global minima of T_{CAS} at the desired upper frequency of operation. This is the current at which the input transistor will be biased.

3) Compute the optimum source resistance for a unit-area transistor and scale the device size such that the optimum source resistance for the scaled device is near the desired generator impedance at the upper frequency of operation. The noise optimization is chosen for the upper frequency rather than the center of the band since the sensitivity parameter N rises rapidly with frequency.

4) Determine the collector–base capacitance and transconductance at the bias point and use these values to compute the emitter inductance required for a power match somewhere near the center of the band. The power-match frequency is chosen to be below the noise match frequency so as to improve the bandwidth over which reasonable noise and power match is achieved.

5) If designing an amplifier with a cascoded first stage, continue on to the next step. If designing an amplifier with a single transistor as the first stage, determine the desired gain and choose a load resistance value. Be sure that sufficient gain is realized in order to reduce the noise contribution from subsequent stages to a reasonable level.

6) Compute the series inductance needed for noise match at the upper frequency of operation.

7) Design subsequent stages to flatten out the gain and provide reasonable output return loss. If necessary, build a pad into the output to improve S_{22}. If using a single transistor as the input stage, ensure that the loading of the first stage is such that the effective input capacitance is not changed[1].

[1] i.e., either use small devices at the output or a cascoded input stage to improve the input–output isolation

9.2.1 Example Design: a 0.7–3.0 GHz Cryogenic LNA with Very Low-Noise Temperature

Upon measuring and modeling the ST BiP-X1 devices described in Chapters 4-7 it was realized that the performance of these devices was expected to be very good. Therefore, a discrete transistor amplifier was designed using the procedure described above. In addition to providing a means to evaluate the ST BiP-X1 device, the design and testing of a discrete amplifier also provides a framework in which to verify the noise modeling procedure developed in this work. The design process and measurements results will be detailed below.

9.2.1.1 Design Process

In this section, the design process of a discrete-transistor two-stage low-noise amplifier is described. The amplifier is to operate at 18 K physical temperature, use ST BiP-X1 devices, and cover the 0.7–3 GHz frequency range. The design procedure will follow the plan described in the previous section and will be carried out using custom scalable compact noise models of the type described in Section 9.1. As a first step, the collector current density needs to be determined. To accomplish this task, $T_{CAS,min}$ at 2.9GHz was plotted as a function of collector current density and the global minima was located. Referring to Fig. 9.2(a), the optimum collector current density is approximately 0.75 mA/μm^2. At this bias current, the value of $T_{CAS,min}$ is $\sim$ 1.45 K. Following the determination of the bias point, the device size is determined by choosing the size that produces $R_{OPT,T_{CAS,min}} = 50\ \Omega$ at 2.9 GHz. Referring to Fig. 9.2(b), we see that the optimum source resistance at 2.9 GHz for this line of devices is approximately 550 $\Omega \cdot \mu$m^2, corresponding an ideal device size of 11 μm^2. Thus, we will use a device with four $0.17 \times 14.79 \mu$m^2 fingers and a total device area just over 10 μm^2. Thus, the collector current should be about 7.5 mA.

With the bias point and device size selected, the small-signal parameter values can now be determined. The calculated values of the input transistor small-signal component values appear in Table 9.1. Next, a load resistance value of 50 Ω is selected to give an open-loop voltage gain greater

Table 9.1: Small-signal model component values for the input device at the optimum bias point.

C_{BE}	C_{CB}	C_{CS}	g_m	g_{be}	r_b	r_c	r_e
680 fF	120 fF	14 fF	1.0 S	22μS	1.3Ω	0.48Ω	0.32Ω

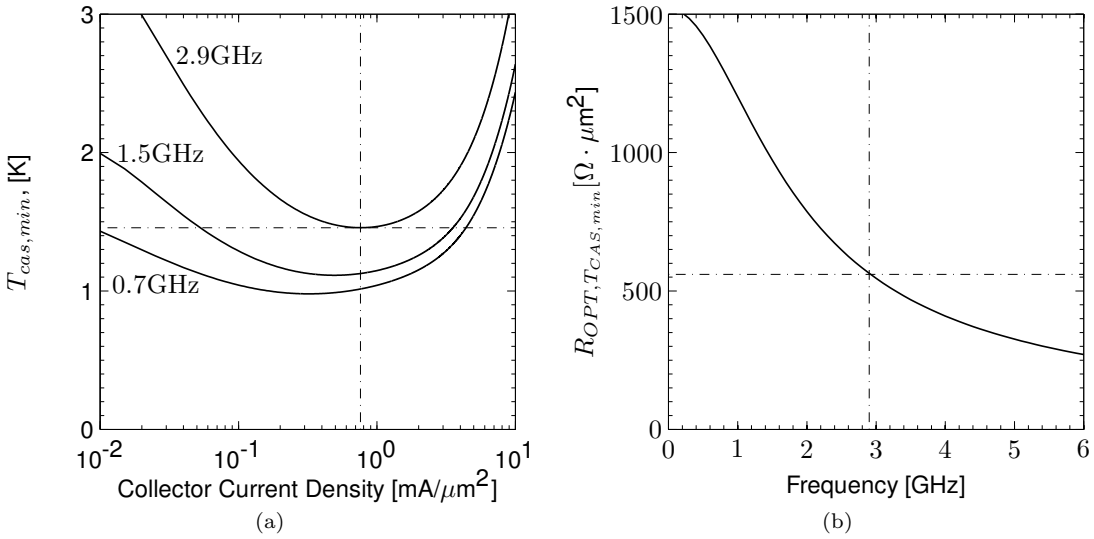


Figure 9.2: LNA device size optimization: selection of (a) collector current and (b) device size

than 30 dB. The resulting equivalent input capacitance next computed as

$$\widetilde{C}_{IN} = C_{BE} + C_{CB}\left(1 + \frac{g_m R_L}{1 + g_m R_e}\right) = 5.3\ pF. \tag{9.2}$$

With knowledge of $\widetilde{C}_{IN}$, we are now ready to complete the design of the input stage by determining the required values of L_e and L_b. The degeneration inductance, L_e is computed directly from equation (9.1) as $L_e = Z_0\widetilde{C}_{IN}/g_m \approx 270$ pH. Finally, L_b is chosen as 2.1 nH to provide noise match at 3 GHz, while providing power match near the center of the band.

With the topology of the first stage in place, the rest of the amplifier was designed and a schematic diagram of the final circuit appears in Fig. 9.3. The amplifier consists of the low-noise input stage described above, followed by a capacitively coupled output buffer stage. The output stage helps improve the performance of the amplifier by increasing the input–output isolation and providing additional gain. The device selected for the output buffer is a ST BiP-X1 device with four $0.17 \times 5.61\mu\text{m}^2$ fingers. This device is among the smallest of the sample devices that were available and was chosen so as to minimize capacitive loading on the input stage. The bias current for the output buffer was set to be approximately 1 mA, which biases the device in a range where it can provide gain without presenting too large of a capacitive load to the input stage.

In order to flatten the gain response of the amplifier, the stages are coupled via 2.7 pF single-layer ceramic capacitors fabricated by Dielectric-Labs (DLI). As it turns out, it is important to be judicious in selection of the material used in ceramic capacitors when working at cryogenic temperatures.

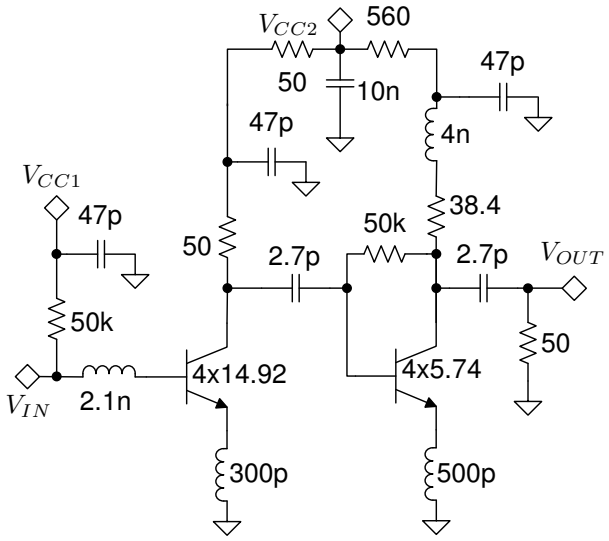


Figure 9.3: Schematic drawing of discrete-transistor cryogenic amplifier. The devices are from the experimental ST BiPX1 process.

In order to determine which materials were acceptable, a wide variety of DLI capacitors were characterized as a function of temperature and the temperature coefficients were extracted. A list of the dielectric constants of the materials used in the tested capacitors appears in Table 9.2. The results of this experiment appear in Fig. 9.4 and clearly show that the majority of the capacitors change significantly with cooling. Furthermore, referring to Fig. 9.4 and Table 9.2, it is clear that there is a correlation between the room temperature dielectric constant of the capacitors and the change that occurs with cooling. Fortunately, the CG material was found to be adequately temperature stable.. In addition to the ceramic capacitors, a metal–insulator–semiconductor (MIS) capacitor was also characterized, and its temperature stability was found to be excellent. Thus, the capacitors used in the amplifier are of the MIS and CG variety.

The amplifier was assembled in an inexpensive package described in [43]. Photographs of the assembled module appears in 9.5. The transistors were mounted in via holes and the 300 pH and 500 pH degeneration inductors were realized as bond wires to the metallic ground plane of the pcb. A close-up photograph of one of the input transistor appears in Fig. 9.5(b), and illustrates the grounding scheme. The 2.1 nH inductor at the input of the amplifier was also realized using a bondwire. Measured results are presented below.

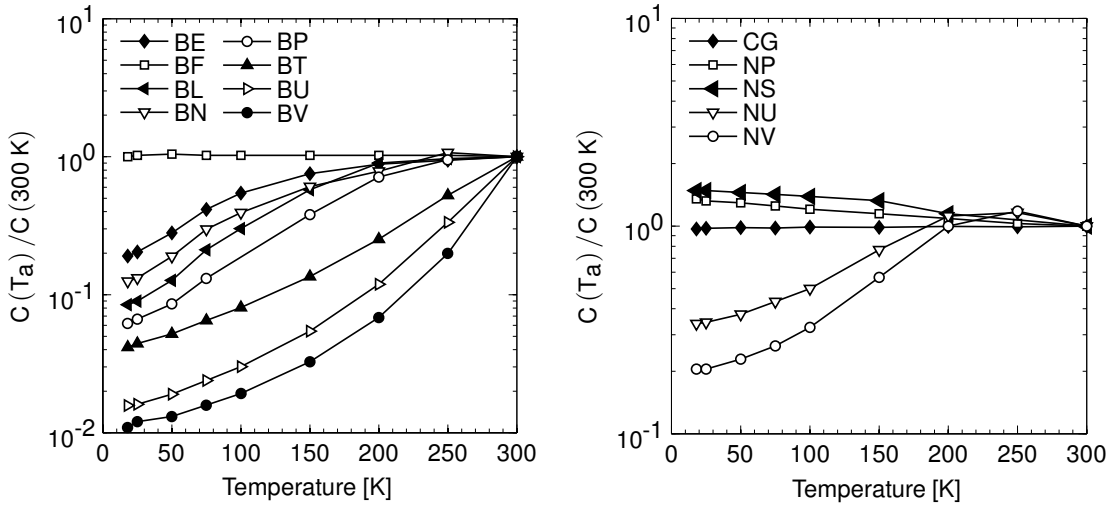


Figure 9.4: Change in capacitance of DLI single-layer ceramic capacitors with cooling

Table 9.2: Dielectric constants for different materials used in DLI single-layer capacitors [150]

Di. Code	BE	BF	BJ	BL	BN	BT	BU	BV	CG	NP	NS	NU	NV
ϵ_r	1250	445	3300	2000	4500	4200	8500	13500	70	85	300	600	900

9.2.1.2 Measurement Results

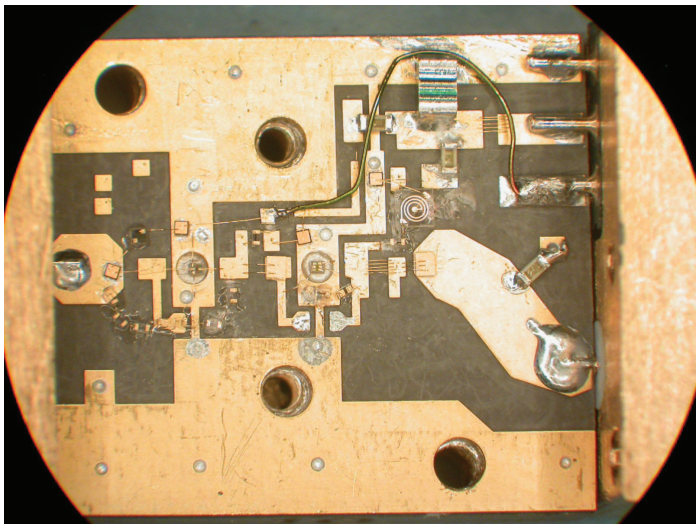
Upon assembly, the amplifier was cooled to 18 K where the gain and noise were measured using the cold attenuator method described in Section 1.2.3. The measured noise appears in Fig. 9.6 along with simulation and inspection reveals several key results:

1) The modeled noise is very consistent with the measured noise. This fact is important because it provides strong support for the noise-modeling procedure presented in Chapter 7.

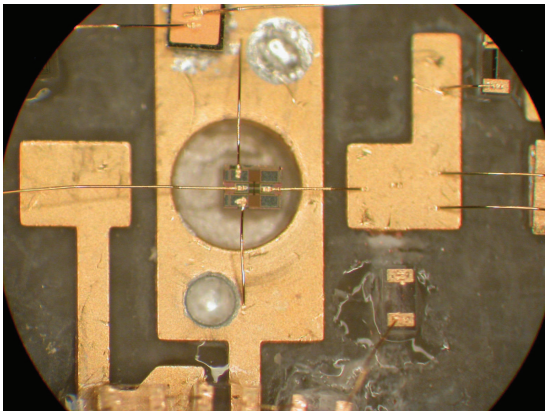
2) At 2.9 GHz, the measured noise of ∼1.75 K is within 0.3 K of T_{CAS,min,J_C}, the predicted global optimum of the device.

3) The noise temperature is less than 2 K from 1–3 GHz. It is believed that this is the best noise result to date for a LNA employing devices fabricated in a Si technology. Furthermore, this result is comparable to the state-of-the-art result for InP of ∼ 1.5 K from 2-4 GHz [9].

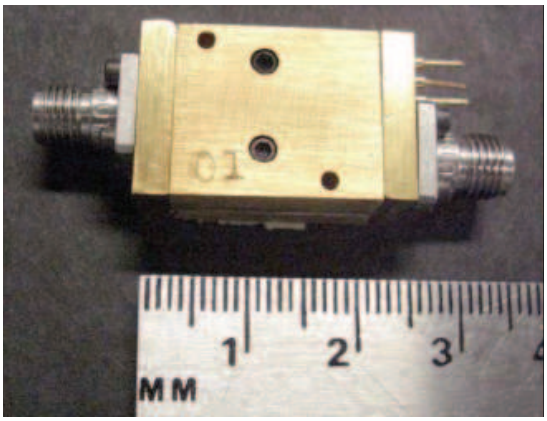
After measuring the noise, the cryostat was heated to room temperature so that the amplifier could be connected to an auxiliary channel and cryogenic S-parameter measurements could be taken.



(a)



(b)



(c)

Figure 9.5: (a) The assembled discrete amplifier. The input is on the left and the output is on the right. DC power comes in via a set of pins in the upper right-hand corner. (b) Closeup of the input device. The transistor is mounted inside of a via hole. The base connection is to the left, whereas the collector connection is to the right. The emitter degeneration inductance is realized using a pair of bondwires connecting from the top and bottom edges of the chip to the chassis. (c) The packaged amplifier with the lid installed

Prior to installing the amplifier, an SOLT calibration was completed with the reference plane located at the end of the internal cables connecting to the amplifier. Once the calibration was completed, the amplifier was installed and the cryostat was cooled back down to 18 K, where the measurement was completed. The results of the S-parameter measurements are presented along with simulated data in Fig. 9.7. The agreement is quite good for all results except for S_{22}, in which a strong disagreement is seen even at low-frequencies. It is believed that this disagreement may have to do with the 50 Ω termination resistor on the output of the amplifier and is not an indication of a transistor modeling

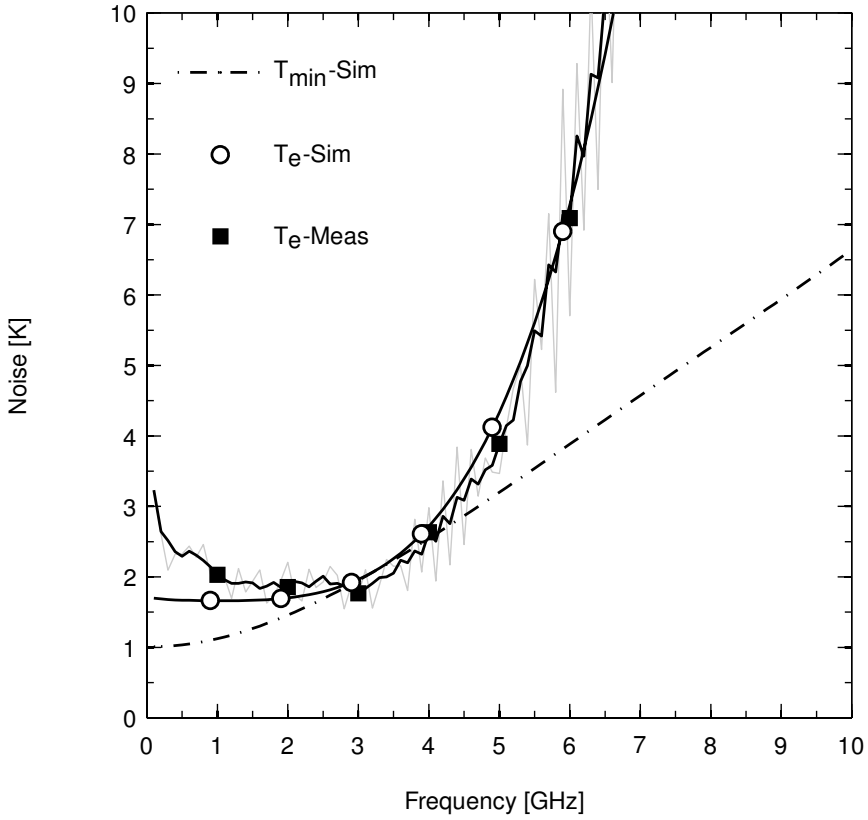


Figure 9.6: Measured and modeled noise for discrete ST amplifier at 18 K physical temperature. The raw data is plotted in grey with a smoothed overlay plotted in black. The agreement between the measured trace and the simulation offers tremendous support for the modeling presented in Chapters 5–7.

mistake. As the main point of the amplifier design was to prove that the modeling of the noise properties of the HBTs is correct, the disagreement in S_{22} was deemed unimportant and was not further investigated.

9.3 Extremely Broadband LNAs Employing Resistive Feedback

There are several applications in which very low-noise input-matched amplifiers in the 0–5 GHz frequency range are required. For example, THz receiver systems that employ hot-electron-bolometer (HEB) mixers in their front ends suffer from stability issues when the low-noise amplifier directly following the mixer is poorly matched. The physical explanation for this problem is believed to be

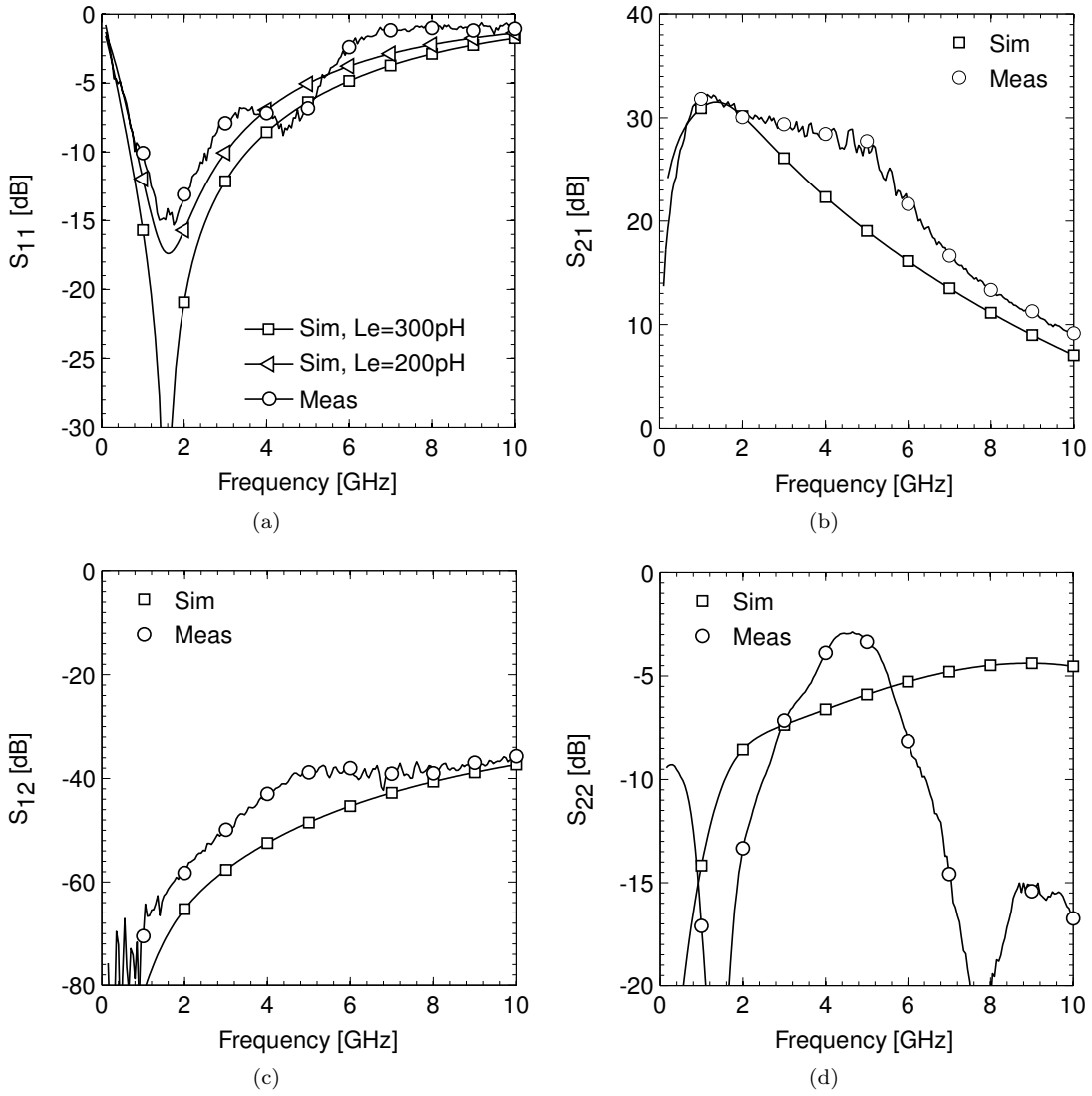


Figure 9.7: Measured and simulated S-parameters for discrete amplifier. The ambient temperature is 18 K for both measurement and simulation

that reflections into HEB mixer cause a modulation of the bias current. This leads to problems as HEB mixers are superconducting devices biased at a critical current density, above which the material is resistive and below which it is superconducting. Thus, if the bias current is modulated due to reflections, electrothermal-feedback will occur and the device performance will suffer [151]. Typical HEB mixers can be used with IF bandwidths spanning 0–3 GHz. However, due to the issue of reflections, the bandwidth is usually limited to an octave by an isolator which is inserted between the mixer and the LNA [44]. Thus, very-low noise amplifiers with impedance match are desired for integration with HEB mixers.

While inductively degenerated III-V HEMT amplifiers with excellent input match and noise performance have been demonstrated in the 1–10 GHz range [10], there have been no results in which these devices have been applied to multi-octave extremely low-noise amplifiers covering frequencies in the 0.1–1 GHz range. This is mainly due to the fact that the input impedance of FETs below 1 GHz is very high and, because of the high gain of the devices and resulting Miller capacitances, unreasonably large source inductor values are required in order to generate a real part of the input impedance at low frequencies. Furthermore, as the optimum source resistance of a FET goes as f_t/f [6], very large devices are required in order to obtain values of R_{OPT} in the 50 Ω range. Unfortunately, due to the very large available gain of these devices, this often results in odd-mode oscillations at very high frequencies[2]. Finally, many HEMTs suffer from gate leakage currents, which translates into an increase in T_{min} at low frequencies. Thus, this is an area where there is a large potential for SiGe LNAs.

It is well known that one can achieve an input match through the use of resistive feedback. However, at room temperature, resistive feedback is generally considered to be an impractical option for use in very-low noise applications as the thermal noise from the feedback resistor contributes an unacceptable amount of noise. However, as the feedback resistor is cooled from 300 K to 15 K, its thermal noise reduces by a factor of 20. Thus, it makes sense to look into the feasibility of using a feedback resistor to provide impedance match without destroying the noise performance of the amplifier. For the case of lossless feedback, feedback has no impact on noise figure. However, as the gain is reduced, the noise measure is increased [132]. In the next section , the impact of lossy feedback on the noise parameters of a two-port will be investigated.

9.3.1 The Effect of Lossy Feedback on Noise Performance

When designing a feedback amplifier analytically, a typical strategy is to compute the open-loop characteristics and then determine the feedback component required to achieve a specific design goal. For instance, the closed-loop gain and input resistance of the shunt–shunt feedback configuration shown in Fig. 9.8 can be computed explicitly from the open-loop characteristics as [132, 152]

$$A_{V,CL} \approx -\frac{A_{V,OL}}{1 + A_{V,OL}R_f/R_S} = A_{V,OL} \cdot \theta_{A_V}\left(A_{V,OL}, R_f, R_S\right) \tag{9.3}$$

$$R_{IN,CL} \approx \frac{R_f}{1 + A_{V,OL}} = R_f \cdot \theta_{R_{IN}}\left(A_{V,OL}\right), \tag{9.4}$$

[2]Odd-mode oscillation issues have been observed to be a major problem in state-of-the-art HEMT devices with 400 μm of gate periphery or higher.

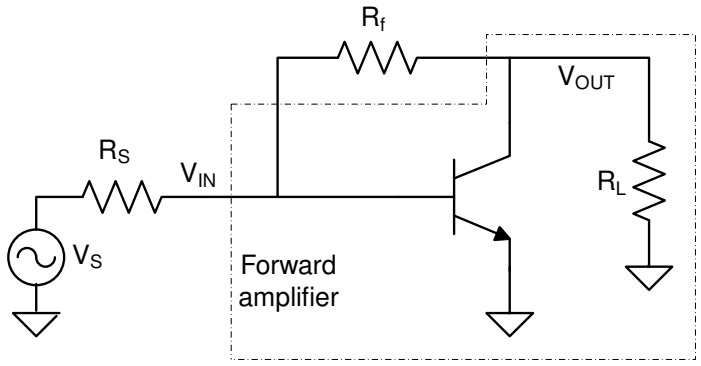


Figure 9.8: Simple shunt–shunt feedback amplifier

where $A_{V,OL}$ is open-loop gain and θ_{A_V} and $\theta_{R_{IN}}$ are scaling parameters relating the closed-loop gain and input resistance to the open-loop parameters and feedback resistance. There is a clear advantage to this approach, in that a scaling factor can be applied to basic theory regarding the open-loop circuit in order to gain an intuitive feel of the closed-loop performance. Thus, from a design point of view, it is important to develop the same sort of framework regarding the noise parameters of feedback circuits. For instance, having a feeling as to the effect that adding lossy feedback has on the minimum noise temperature or optimum source impedance of a amplifier in terms of the open-loop noise parameters is invaluable, as it facilitates the efficient design of optimized circuits.

The impact that lossy feedback will have on an amplifier at cryogenic temperatures can be determined by analyzing the generic circuit shown in Fig. 9.9(b). In this work, the open-loop

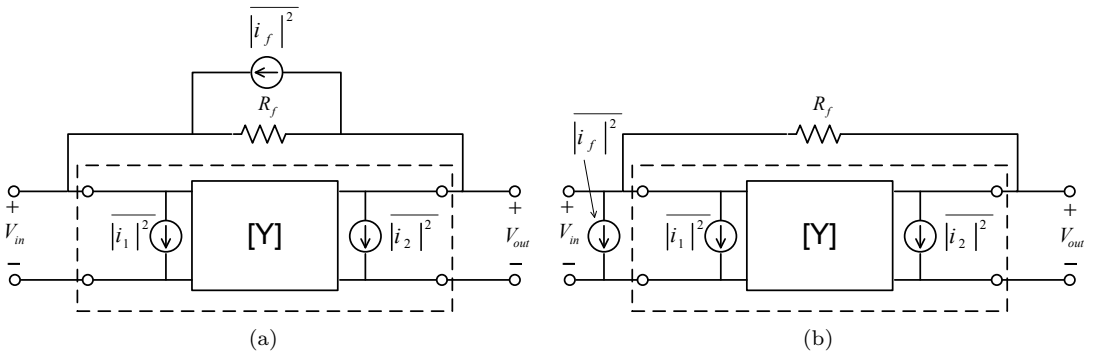


Figure 9.9: (a). Generic two-port network with shunt resistive feedback applied. The two-port network is represented in terms of Y-parameters and the noise is represented by an equivalent input and output current source. The feedback network is located outside the dotted line. (b). Simplified equivalent circuit in which the current noise due to the feedback resistor has been moved to the input. This simplification involves ignoring a fully correlated current noise source at the output and is valid so long as the close loop circuit has high gain.

network is represented in terms of Y-parameters and has external noise-generators[3] $\overline{|i_1|^2}$ and $\overline{|i_2|^2}$. The feedback network consists of a resistor with its noise represented in terms of a thermal current source with power spectral density given as $\overline{|i_f|^2} = 4kT_a g_f \Delta f$, where $g_f = 1/R_f$ is the feedback conductance.

It can be shown[4] that at frequencies well below the unity-current gain cutoff frequency of the two-port, the minimum noise and optimum source resistance of the closed-loop circuit can be approximated as[5]

$$T_{min} \approx \frac{1}{2k}\sqrt{\frac{\overline{|i_1'|^2}\,\overline{|i_2|^2}}{\overline{|Y_{21} - g_f|^2}}} \tag{9.5}$$

and

$$R_{OPT} \approx \frac{2kT_{min}}{\overline{|i_1'|^2}}, \tag{9.6}$$

where the quantity $\overline{|i_1'|^2} = \overline{|i_1|^2} + 4kT_a g_f$ represents a temperature dependent increase in the input current noise power spectral density due to the addition of the feedback resistor to the circuit. Furthermore, if the input transistor is a SiGe HBT with high dc current gain, $\overline{|i_1'|^2} \approx 2qI_{C1}/\beta_{DC1} + 4kT_a g_f$, where I_{C1} and β_{DC1} are the dc collector current and dc current gain of the input transistor. Under these conditions, it can be shown that the following happens to the noise parameters as a consequence of the resistive feedback:

T_{min} The minimum noise of the open-loop amplifier is multiplied, due to g_f, by a factor of $\theta_{NP} \approx \sqrt{1 + 2\beta_{DC1} kT_a g_f / qI_{C1}} = \sqrt{1 + 2\beta_{DC1} g_f / g_{m,ideal}}$. In Chapters 5 and 7, it was observed that at a fixed bias point, β_{DC} rises as roughly $1/T_a$ and T_{min} drops as roughly $1/T_a$. Thus, to first order, θ_{NP} is independent of temperature and the noise added by the lossy feedback will decrease proportionally to T_a. Finally, in order to avoid degradation of the noise due to the inclusion of R_f in the circuit, it is necessary that $\beta_{DC1} g_f \ll g_{m,ideal}/2$.

R_{OPT} The optimum source resistance is multiplied by a factor $1/\theta_{NP}$. Thus, the optimum source resistance will decrease.

N The sensitivity factor N is multiplied by θ_{NP}. This means that the noise performance of the amplifier becomes more sensitive to source noise mismatch once the loop is closed. It should

[3] The internal network is noiseless and the noise parameters are completely described by the two external current generators and their complex correlation coefficient, $\overline{i_1 i_2^*}/\overline{|i_1|\,|i_2|}$.

[4] See Appendix E.2 for the derivation

[5] It has been assumed in the following that $1/Y_{11} \gg R_S$, $\overline{|i_1|^2} \ll \overline{|i_2|^2}$, and $\overline{|i_1|^2 |i_2|^2} \gg \left(4kT_a g_f\right)^2$. These assumptions are valid for SiGe HBTs with high β_{DC} so long as R_f is much larger than the generator resistance

be noted that, under the approximations stated above, the figure-of-merit $4NT_0/T_{min}$ is unaffected by lossy feedback.

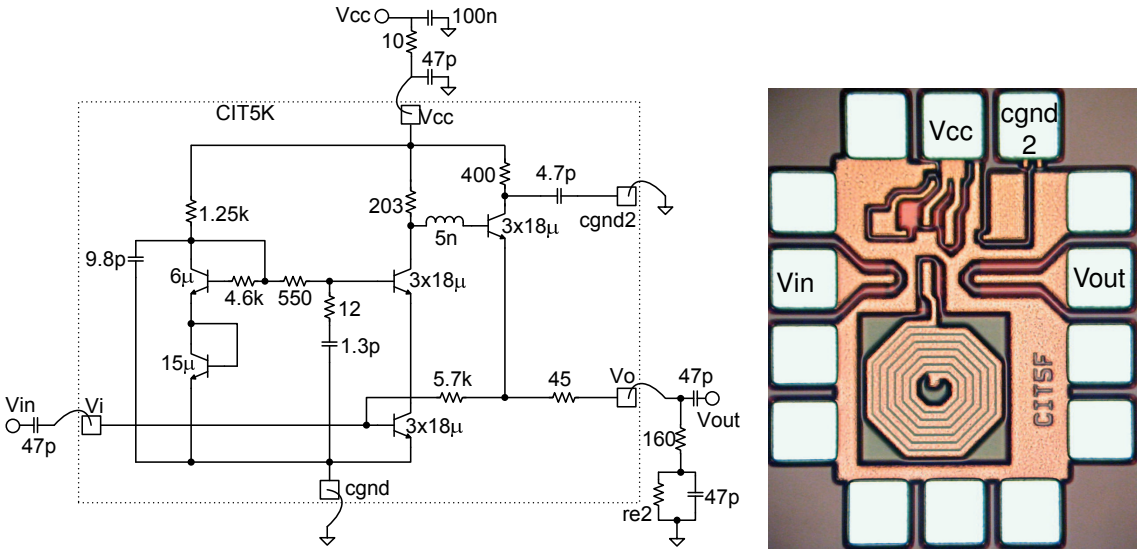
Thus, it can be seen that by connecting the feedback resistor to a point with high open-loop gain, its value can be made large (i.e., g_f small) and the desired R_{in} can be obtained with a degradation factor close to unity.

In this section, the impact of lossy feedback has been analyzed for an amplifier operating well below f_t. The provided equations give insight into how the noise properties of a feedback amplifier are related to those of the same topology without feedback. However, in order to arrive at these simple equations, several assumptions were made which will lose validity as the frequency of operation gets high enough that the input capacitance is non-negligible. Fortunately, equations exist in the literature which allow one to calculate the effect of feedback applied to any two-port [153]. However, due to the involved nature of the calculations, the resulting formulas are not intuitive, and will not be discussed here. Instead, we will move on and take a look at some very broadband amplifier designs based upon this topology.

9.3.2 Example Design: A 0.1–5 GHz MMIC LNA

As a first attempt to design an integrated circuit using the models developed in Chapters 5-7, a 0.1–5 GHz LNA design was targeted for the IBM BiCMOS8HP process. In order to facilitate a rapid design process, compact noise models of the type described in Section 9.1 were developed in AWR's Microwave Office (MWO) for a $0.12 \times 3 \times 18\mu\text{m}^2$ HBT operating at 15 K physical temperature. The amplifier was then designed in MWO. Next, simulation was carried out at 300 K in Cadence Virtuoso using foundry supplied models. After verifying that the simulation results at 15 K and 300 K agreed, a layout was generated and parasitic capacitances were extracted. Finally, the parasitic capacitances were back-annotated into the MWO simulation environment and the parasitics were tuned out. Extracted resistance values were not included in the cryogenic simulations as the conductance of metals increases significantly at cryogenic temperatures.

A schematic diagram of the designed amplifier is shown in Fig. 9.10(a) along with the external components required to operate the amplifier. The circuit consists of a cascode input stage driving an emitter follower buffer amplifier. The output of the emitter follower is fed back to the input via a 5.7 kΩ feedback resistor. The input transistor is sized such that the optimum source resistance in the 5 GHz range is approximately 50 Ω. In order to improve the gain bandwidth of the amplifier, a series inductive element is included on the base of the emitter follower stage.



(a) (b)

Figure 9.10: (a) Schematic diagram of 0.1-5 GHz MMIC LNA including external components. Dotted lines indicate the chip boundary. (b) Photograph of the fabricated IC

In order for the amplifier to work at both room temperature and at 15 K, it is necessary to include provisions to reduce the bias current in each of the transistors by a factor of roughly three when it is operated cryogenically. This reduction is required in order to achieve the same open-loop gain at 15 K as was achieved at room temperature, thereby preserving the frequency response of the amplifier. To accomplish this task, the biasing of the amplifier is configured such that a change in the element r_{e2} and a reduction in V_{CC} is all that is needed. The bias to the base of the cascoded transistor is provided through a current-mirror circuit and the base terminal is bypassed through a damped capacitor. The reason for including the resistance in series with the capacitor is to prevent a parasitic feedback path from forming due to the bondwire inductance from the cgnd terminal to chassis ground. This problem is especially troublesome in extremely low-noise LNAs as the gain tends to be on the order of 30dB. In addition, another parasitic feedback path was avoided by providing a separate ground pad for the bypass capacitor located at the collector of the emitter follower.

A die photograph of the fabricated amplifier appears in Fig. 9.10(b). The amplifier measures $0.5 \times 0.6\text{mm}^2$ including all bondpads, which is approximately one quarter the die area of a typical cryogenic low-noise amplifier fabricated in a III-V process[6]. For testing, the amplifier was mounted in an inexpensive package shown in Fig. 9.11 and consisting of a PC board sandwiched in between

[6]The standard dimension for III-V amplifiers designed in our lab is $2 \times 0.6\text{mm}^2$.

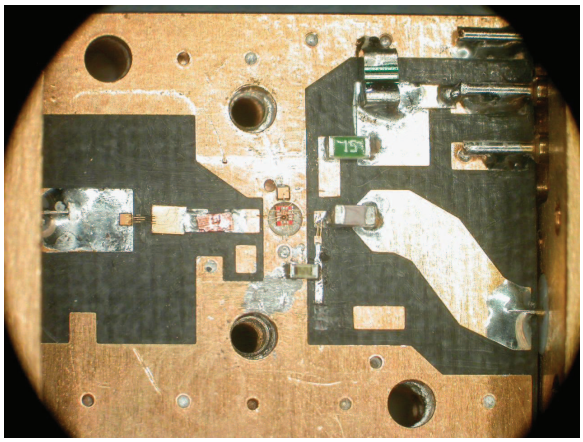


Figure 9.11: Photograph of packaged 0.1–5 GHz amplifier

Table 9.3: Measured performance metrics for 0.1–5 GHz MMIC LNA

Temp.	Freq. Range	T_e 1.5/3GHz	S_{21} 1.5/3GHz	S_{11} 1.5/3GHz	S_{22} 1.5/3GHz	OP1dB	OIP2	OIP3	Pdiss
K	GHz	K	dB	dB	dB	dBm	dBm	dBm	mW
15	0.1–5	4.3/3.7	31.7/29.9	-14.1/-14.7	-17.3/-18.2	-10.5	30.5	5.9	20
300	0.1–5	67/75	28.3/27.0	-8.4/-11.8	-16.0/-15.5	-	-	-	76

two gold plated brass carriers [43], with the IC mounted in a VIA hole to minimize bondwire lengths.

Measurements were made at both 15 K and 300 K. Room temperature gain and noise are plotted along with simulation results in Fig. 9.12(a) and the return loss is plotted in Fig. 9.12(b). At 300 K, the packaged amplifier has a T_e of less than 92 K (1.2 dB NF) out to 5 GHz with an average gain and T_e of 27.6 dB and 76 K (1.0 dB NF) over the 0.1–5 GHz range. In addition, inspection of Figs. 9.12(a) and (b) reveals very good agreement between simulation and measurement. The slight discrepancy in the gain can be explained by the failure to account for wiring resistances in the simulation.

Following room temperature measurements, cryogenic gain and noise measurements were carried out using the cold attenuator method described in [154] and the results are plotted in Fig. 9.12(c). The amplifier achieves a T_e of better than 5.4 K out to 5 GHz with an average value of T_e and gain of 4.3 K and 30.8 dB over the operating band. Following noise and gain measurements, the amplifier was cooled without an attenuator on the input and its S_{11} and linearity were measured and the results appear in Fig. 9.12(d) and Table 9.3. At 15 K, S_{11}, S_{21}, and T_e were all found to be in excellent agreement with the modeled result. The ripple in the return loss measurement is believed to be due to changes in the stainless steel cable connecting the amplifier to input of the

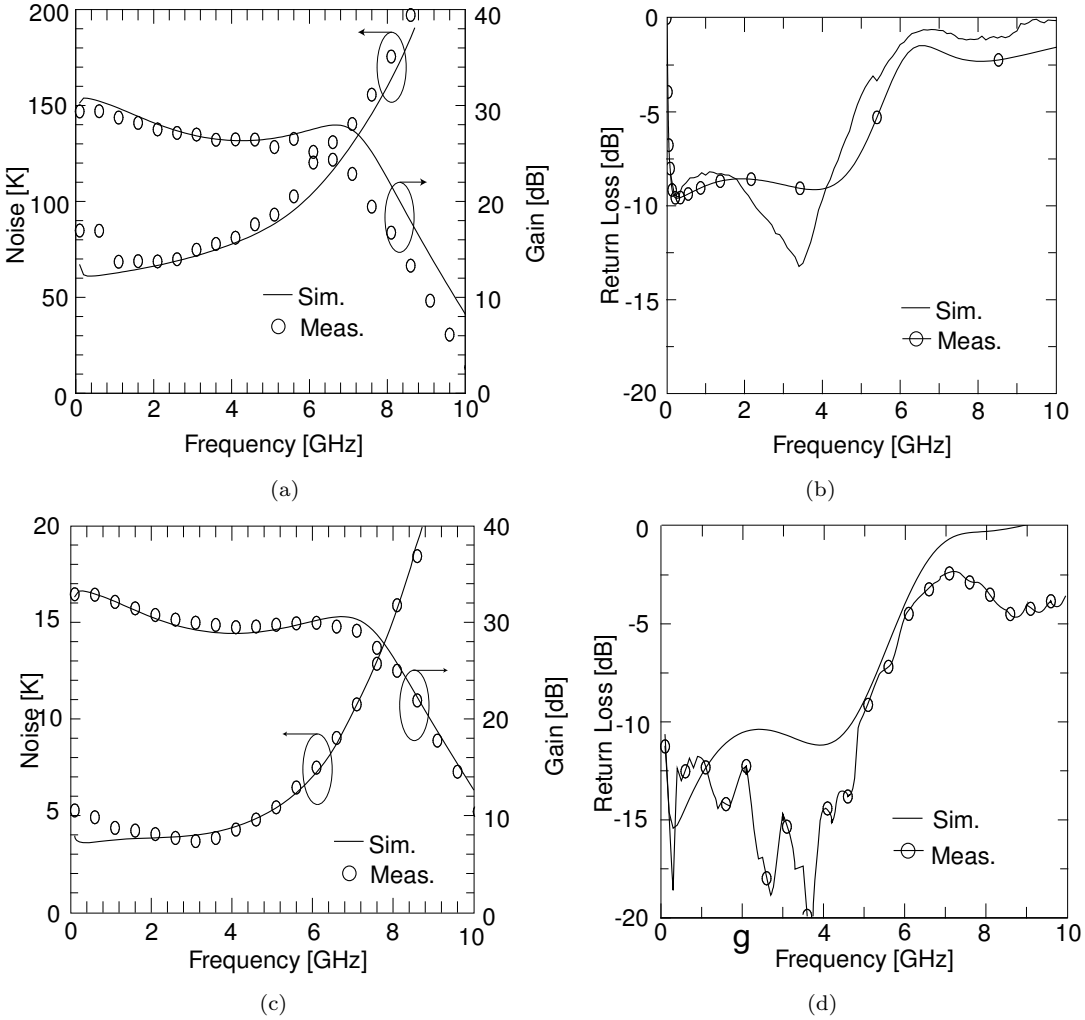


Figure 9.12: Gain and noise measured at (a) 300 K and (c) 15 K physical temperature. (b) Return loss measured at 300 K and (d) 15 K. The 15 K noise measurement setup has been calibrated to ±1 K accuracy. To account for packaging effects in the simulation, an input loss of 0.27 dB was assumed at 300 K and the inductance values of bondwires connecting to the amplifier were tuned as the length of each bondwire was not accurately known. The bondwire inductances were assumed to be independent of temperature and were tuned in both the 15 and 300 K simulations simultaneously.

dewar occurring with cooling as calibration of the VNA was carried out at 300 K.

9.3.3 Other Amplifier Designs

Several additional single ended LNAs have been developed in the IBM BiCMOS8HP process. Of particular interest is an extremely broadband LNA, which has not been fully characterized as of the writing of this dissertation. A schematic diagram of the amplifier appears in Fig. 9.13. The amplifier uses a topology similar to that of the 0.1–5 GHz amplifier described above with several

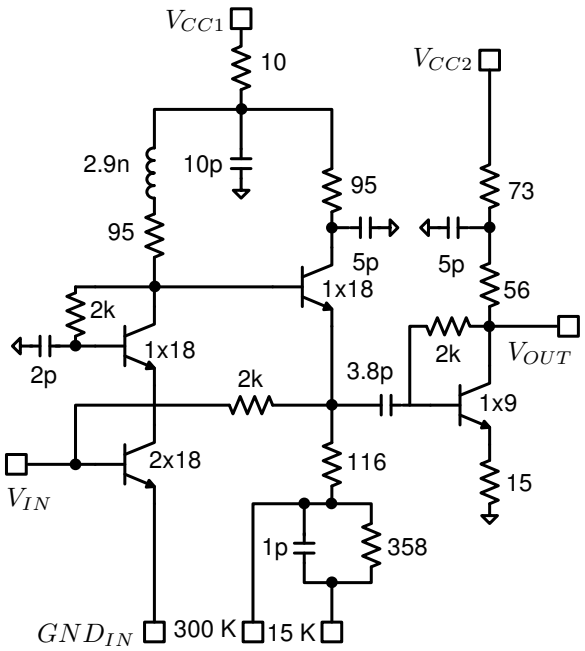


Figure 9.13: Schematic diagram for the 0.1–10GHz amplifier

notable topology improvements:

1) The bandwidth has been doubled (0.1–10 GHz).

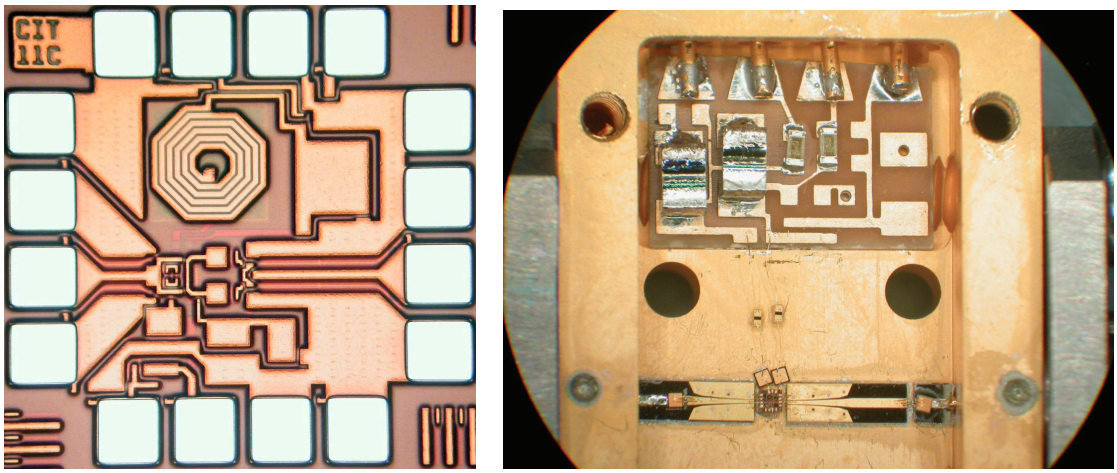
2) The input transistor has its own ground return path through the input CPW transmission line. As this structure is floating with respect to the global chip ground, this feature allows one to have a well defined emitter degeneration inductance and also avoids problems associated with feedback due to a common ground inductance.

3) The cascode transistor is self-biased using a 2 kΩ polysilicon resistor. This is advantageous in that it removes the requirement for a current mirror and reduces power dissipation.

4) The entire bias network is included on chip and the circuit can be configured to operate at either 300 or 15 K by shorting one of two pins to ground (e.g., to operate at 15 K, the pin labeled "15 K" should be tied to ground). This is useful as it reduces the number of components and interconnects.

5) An output buffer amplifier has been included in order to isolate the feedback amplifier from the load. This significantly reduces the sensitivity of the circuit to packaging.

The circuit was fabricated in the IBM BiCMOS8HP process and a photograph of the circuit appears in Fig. 9.14(a). The circuit including all bondpads occupies approximately 0.6×0.6 mm^2.



(a) (b)

Figure 9.14: (a) 0.1–10 GHz amplifier die photograph. The die area including bondpads is 0.6 × 0.6mm^2. (b) Packaged amplifier

The chip was mounted in a coaxial fixture in order to make preliminary room temperature measurements. A photo of the packaged amplifier appears in Fig. 9.14(b). The input signals are brought in using microstrip to co-planar waveguide transitions. The dc lines are bypassed by 47 pF MIS capacitors near the chip with larger capacitors on a dc bias board. 10 Ω series resistors are placed between the two bypass capacitors to prevent them from resonating with each other.

The S-parameters of the amplifier were measured at room temperature and the preliminary results are shown in Fig. 9.15. The disagreement between the room temperature measurements and simulations are not yet fully understood, but are believed to have to do with the packaging of the chip, as the response is quite sensitive to the location of the ground bondwires. Further work is needed to determine the exact reason that the circuit is not behaving as expected and to find out whether the problem is intrinsic to the chip or in the way it is packaged.

9.4 Differential LNAs

While the majority of applications require that a signal which is referenced to ground be sensed, there are some cases in which the signal to be sensed is defined differentially between a pair of wires. Furthermore, many of these differential systems do not have 100 Ω output impedance. For example, quasi-self complimentary antennas can be made to work over extremely wide bandwidths, but have a frequency independent differential input impedance of 270 Ω. Thus, to use a quasi-self complementary feed, either a transformer is required to adapt the 270 Ω differential mode to a 50 Ω

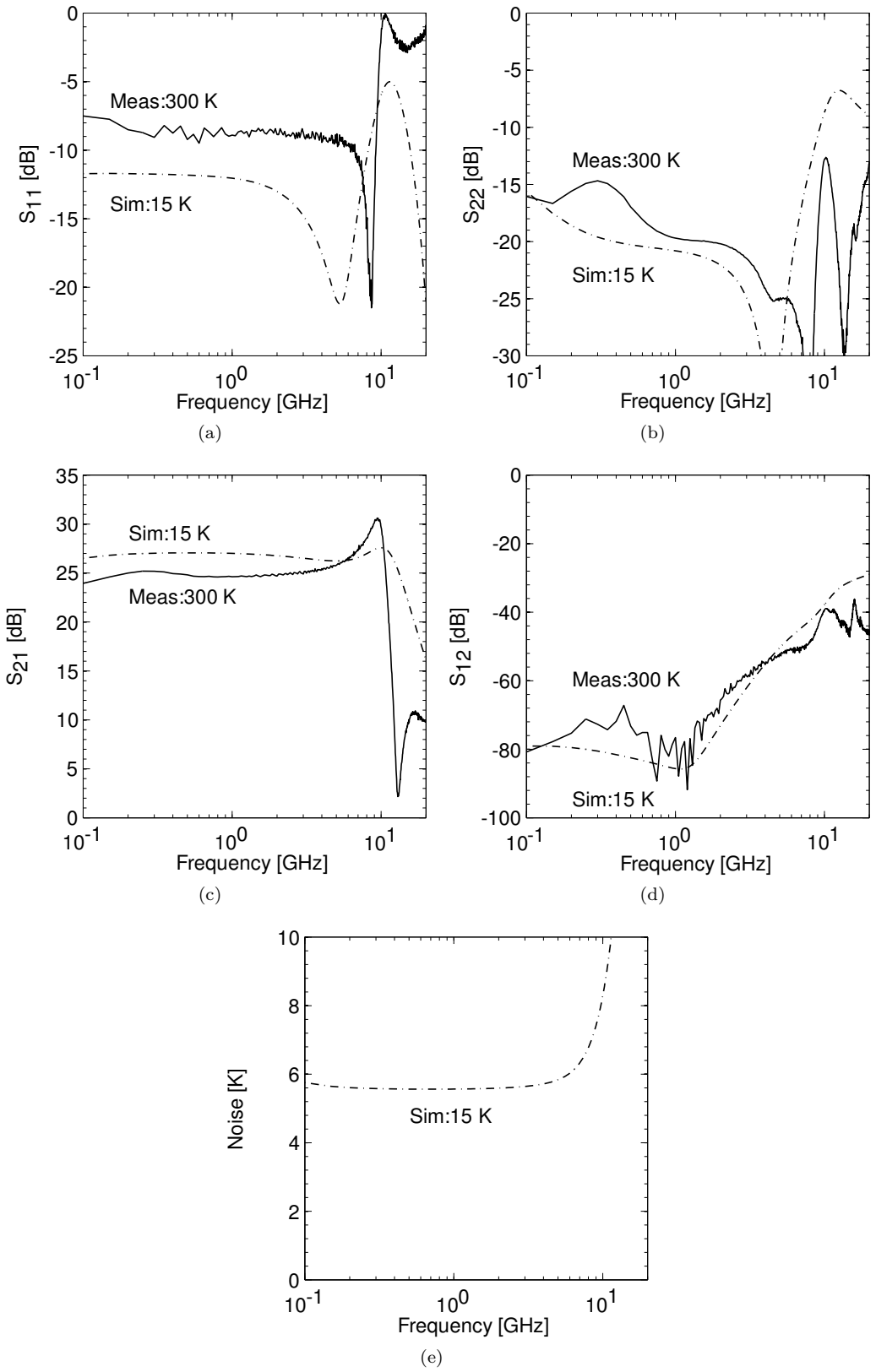


Figure 9.15: 0.1–10GHz amplifier response. The measurements were taken at 300 K whereas the modeling was done at 15 K.

Table 9.4: Bias currents for 0.1–10 GHz differential 270 Ω cryogenic LNA

Temperature K	V_{CC1} V	V_{CC2} V	I_{CC1} mA	I_{CC2} mA
15	2.00	2.35	1.5	9.5
40	2.05	2.50	2.0	14.0
300	2.60	3.90	6.9	35.3

single ended signal, or an amplifier that is designed to sense the high-impedance differential signal directly must be used. When evaluating the two options, it becomes apparent that the latter is preferable due to the fact that the transformer will introduce loss (which in turn introduces noise), whereas a single ended amplifier can be converted to a differential structure with no change in the noise characteristics[7]. In this section, the design of differential LNAs intended to interface to decade bandwidth feeds with 270 Ω output impedance will be presented. These amplifiers have been designed using the same custom compact-noise models for the IBM BiCMOS8HP devices that were used to design the single ended integrated circuit amplifiers discussed above.

9.4.1 Example Design: A 0.1–10 GHz Differential LNA with 270 Ω Input Impedance

In this section, the simulation results and preliminary measurements of a 0.1–10GHz differential LNA designed[8] for operation at 15 K and matched to a 270 Ω differential-mode source impedance and a 100 Ω differential load impedance is presented. A schematic diagram of the amplifier is shown in Fig. 9.16. The amplifier consists of a cascade of two differential gain stages with resistive differential-mode feedback applied in order to generate an input match. Inductive peaking is used in the load of the first stage in order to improve the bandwidth of the amplifier. In order to avoid the use of a cascode topology, capacitive neutralization is applied using dummy transistors with floating emitters. The neutralization capacitors are denoted as C_n on the schematic. The amplifier requires two reference currents and two reference voltages; the expected V_{CC} voltages and currents are listed in Table 9.4. Although the amplifier was designed for operation at 15 K, simulations have been also been carried out at 40 and 300 K.

The circuit was fabricated in the IBM BiCMOS8HP process and a die photograph is shown in Fig. 9.17. The amplifier performance was measured on wafer at 300 K using an Agilent 4-port vector

[7] This assumes an infinite common-mode rejection ratio (CMRR). In practice, there will always be a small degradation to the noise due to a finite CMRR.

[8] As explained in Chapter 7, the fact that the amplifier is designed for operation at 15 K means that the input device will be to small for noise match at higher temperatures.

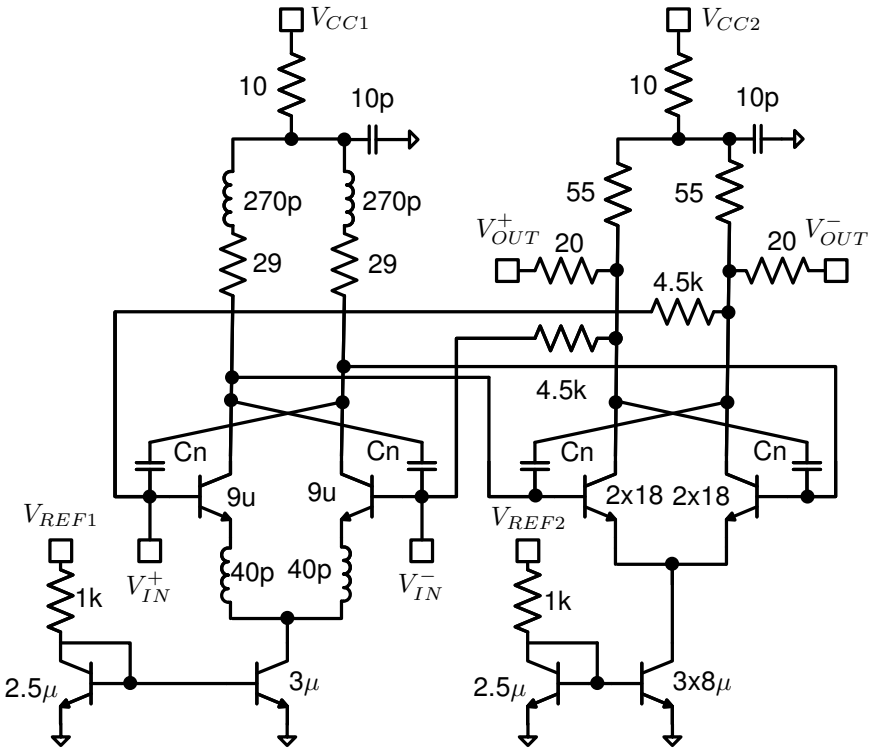


Figure 9.16: Schematic diagram of differential 0.1–10 GHz 270 Ω cryogenic LNA.

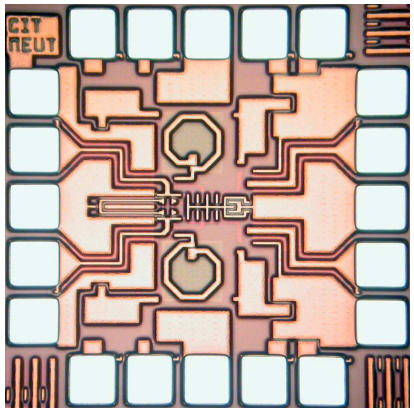


Figure 9.17: Die photograph of 0.1–10GHz differential 270 Ω LNA. The chip dimensions are 0.7 × 0.7 mm^2.

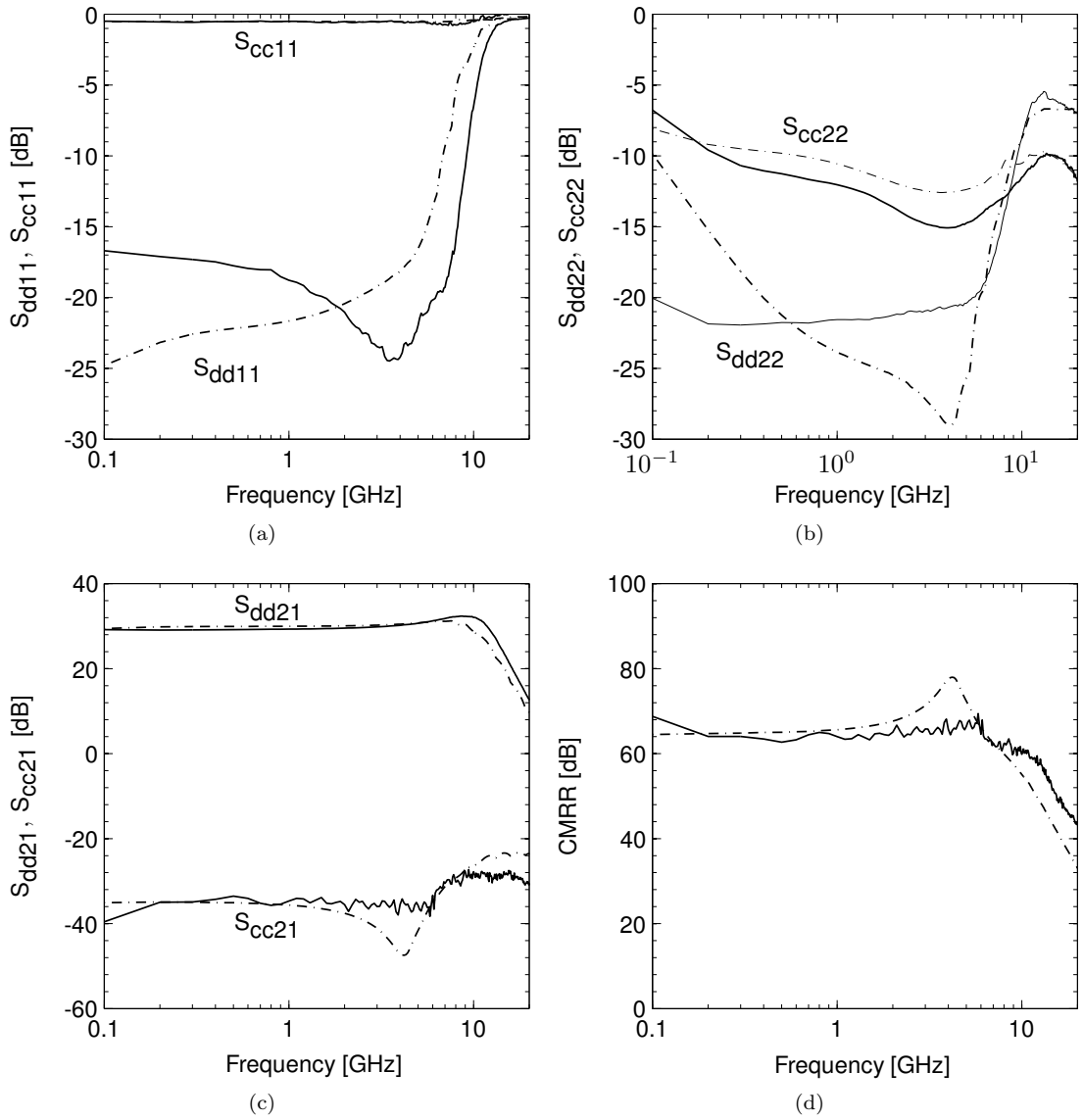
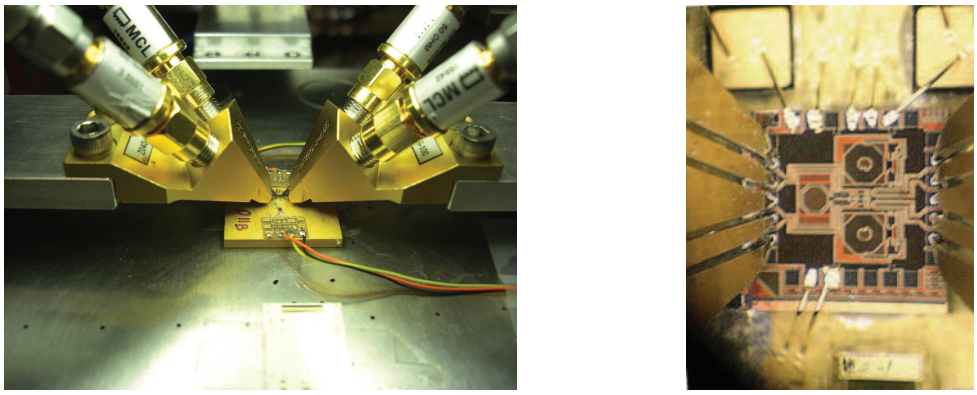


Figure 9.18: Measured and simulated mixed-mode S-parameters for 0.1–10 GHz 270 Ω differential LNA at 300 K physical temperature. Solid lines represent the measured data and dashed lines represent the simulation. The measurements are normalized to 270 Ω differential at the input port and 100 Ω differential at the output port. The agreement with the room-temperature simulation is quite good.

network analyzer and the results are compared to simulation in Fig. 9.18. A photograph of the test setup appears in Fig. 9.19. The agreement between the measurement and simulation is very good. The common-mode rejection ratio was found to be greater than ∼60 dB over the entire operating range and the input return loss was measured to be better than 15 dB over most of the operating band. The excellent CMRR was achieved by careful circuit design and the use of common-centroid layout techniques. Due to the very good agreement between room temperature measurements and



(a) (b)

Figure 9.19: Photographs of the differential amplifier test setup

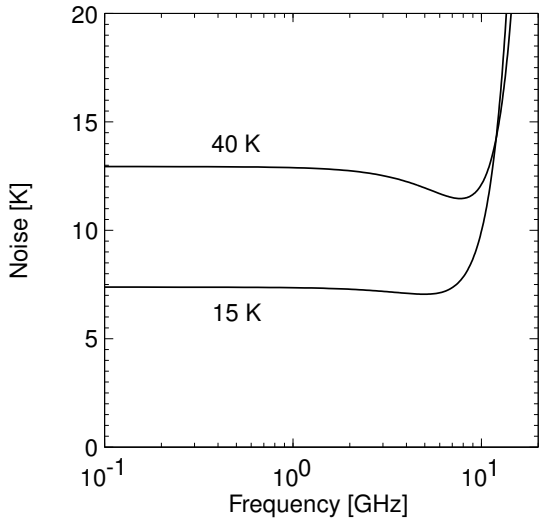


Figure 9.20: Expected noise for 0.1–10 GHz 270 Ω differential LNA operating at 15 and 40 K

simulations, the chip is believed to be quite promising. While the noise is yet to be measured, the simulated differential mode noise is shown for operation at 15 and 40 K in Fig. 9.20. The noise is expected to be in the 7 K range at 15 K physical temperature with a considerable degradation expected at 40 K physical temperature due to the change in optimum device size. Future steps include mounting the dies in coaxial fixtures and measuring the cryogenic noise performance of the amplifier.

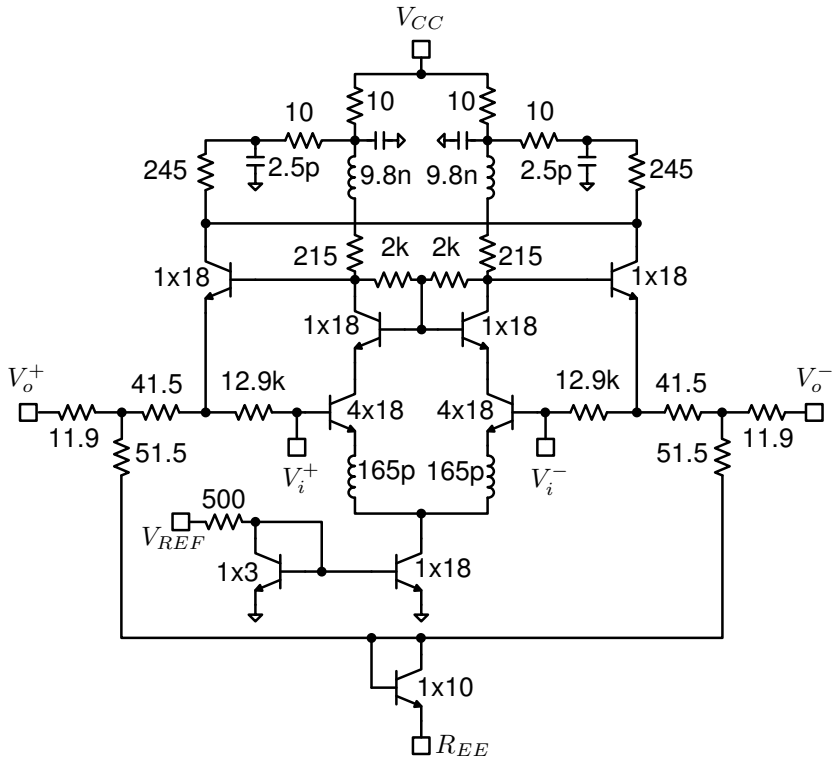


Figure 9.21: Schematic diagram for 0.5–4 GHz 270 Ω differential-input, 100 Ω differential-output cryogenic LNA

9.4.2 Example Design: 0.5–4 GHz Differential LNA with 270 Ω Input Impedance

A 0.5–4 GHz 270 Ω differential-input, 100 Ω differential-output, cryogenic LNA that has been optimized for operation at 40 K has been designed and fabricated in the IBM BiCMOS8HP process. The schematic diagram of the amplifier appears in Fig. 9.21. The topology is quite similar to that used in the single ended design presented in Section 9.3.2, with the appropriate modifications made to convert the structure to a differential form. In order to bias the amplifier, a resistor must be connected from the pad labeled R_{EE} to chassis ground. The values of R_{EE} required to operate the amplifier at 300, 40, and 18 K are given along with the associated bias voltage and current in Table 9.5.

The fabricated amplifiers have been received and preliminary on-wafer measurements have been made at 300 K. A die photograph appears in Fig. 9.23 and the room temperature measurement results appear in Fig. 9.22 along with 15 and 40 K simulation results. The simulated response is very promising; the amplifier is expected to have better than 5 K noise up to 3 GHz when operating

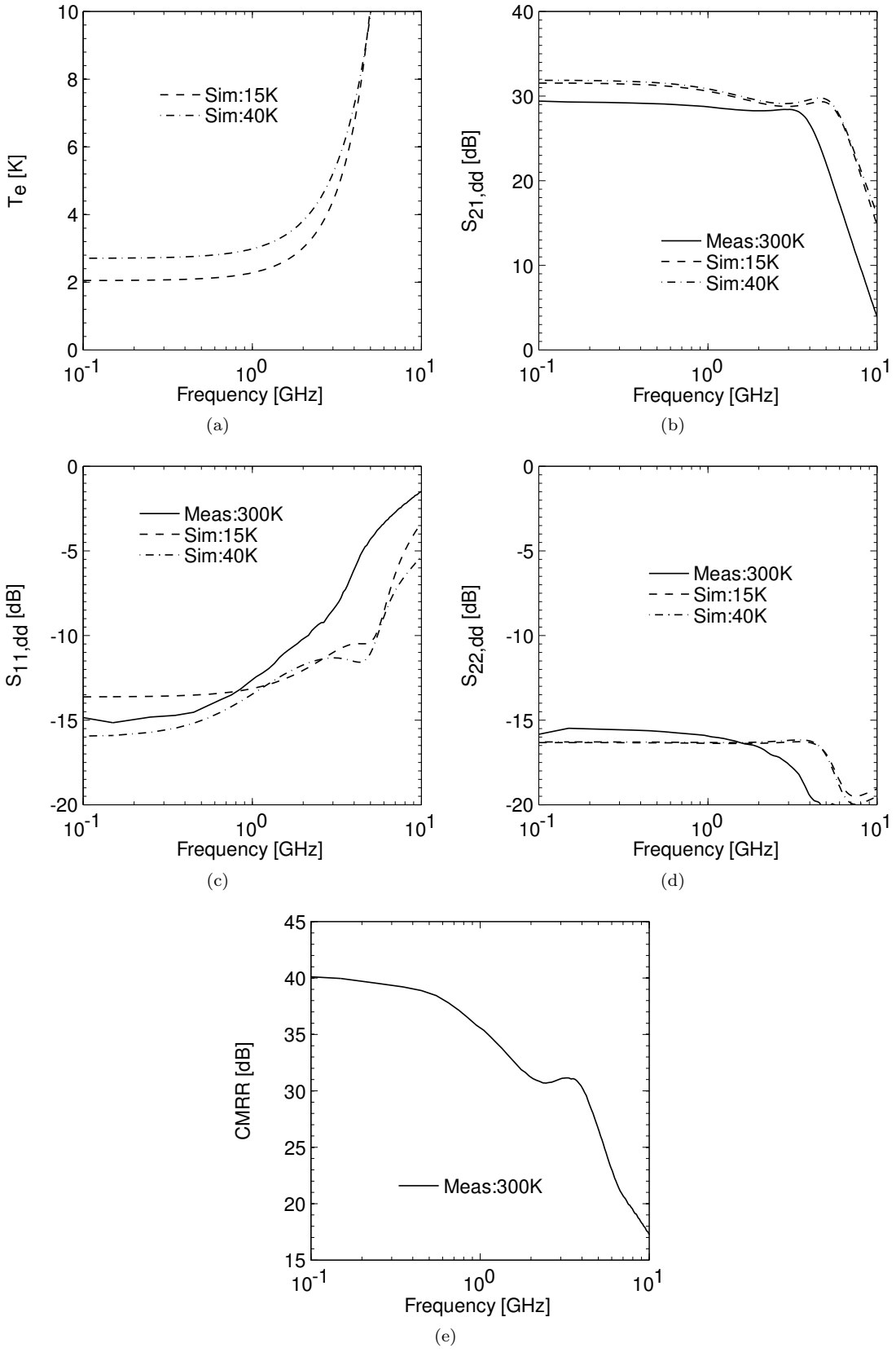


Figure 9.22: Simulated and measured response for 0.5–4 GHz 270 Ω differential-input LNA

Table 9.5: Bias point and value of R_{EE} for 0.5–4 GHz LNA

Temperature K	R_{EE} Ω	V_{CC} V	I_{CC} mA
15	180	3.60	10.4
40	140	3.75	11.8
300	0	5.50	38.0

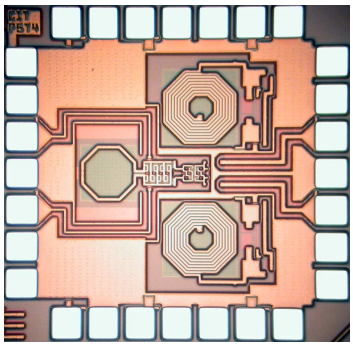


Figure 9.23: Die photograph for 0.5–4 GHz differential LNA. The chip measures approximately 0.95×0.95 mm^2.

at 40 K. The measured S-parameters agree with the expectation at low frequencies, however the bandwidth of the received amplifier is less than expected. It is believed that this is due to the fact that the simulation is for cryogenic operation whereas the measurement was made at room temperature, but further simulations and measurements are needed to clarify this issue. The measured common-mode rejection is greater than 30 dB over the entire range of operation. While the noise performance was not measured, the simulated results are quite promising, with a noise temperature at 40 K physical temperature of better than 5 K expected at frequencies below 3 GHz. Further work is needed in order to package the amplifier and measure its cryogenic noise performance.

9.5 Summary

In this chapter, the models developed in Chapters 5–7 have been applied to the design of several state-of-the-art low-noise amplifiers. The design and measurement of a discrete amplifier employing inductive degeneration has been presented and the excellent agreement between the measured and simulated noise offers strong support for the noise modeling procedure presented in Part II of the dissertation. Furthermore, the noise performance of the discrete amplifier is on par with the best InP amplifiers reported to date. Thus, this result is quite important in validating the selection of

SiGe devices over their III-V counterparts for applications in which performance is the key criterion by which a technology is to be selected.

In addition to the inductively degenerated discrete LNA, the use of resistive feedback to design extremely broadband integrated circuit LNAs has been explored both theoretically and experimentally. As an example, a 0.1–5 GHz LNA has been designed and measured. The experimental results have been found to agree very well with simulation, and the low noise (<5 K) in conjunction with the good return loss (>10 dB) over this extremely wide, 50:1, bandwidth offers further evidence of the utility of resistive feedback at cryogenic temperatures. Finally, at the end of the chapter, some newly fabricated high-impedance differential amplifiers were presented and their preliminary room temperature measurements and expected cryogenic performance were discussed.

Chapter 10

Conclusions

In this dissertation, the theory, modeling, and applications of state-of-the-art SiGe heterojunction bipolar transistors operating at cryogenic temperatures have been investigated. In the first part of the thesis, the physical operating principles of SiGe HBTs were reviewed and a framework was developed in order to facilitate an understanding of the physical reasons behind the changes in the devices as they are cooled. In the second part of the dissertation, the dc, RF, and noise properties of a variety of state-of-the-art SiGe HBTs were studied at temperatures ranging from 18–300 K. The study was conducted in a systematic manner and the temperature dependence of each of the key dc, small-signal, and noise parameters were extracted and analyzed. Furthermore, dc, small-signal, and noise models were developed. The results of this study are important, as they provide a great deal of practical information which can aid in developing an understanding of the physical changes that occur when SiGe devices are operated at cryogenic temperatures. The section concluded in Chapter 8 with a concise summary of many of the important trends that were identified in this part of the dissertation.

In the final part of the thesis, the models developed in the earlier Chapters were applied to the design of state-of-the-art SiGe low-noise amplifiers, in both discrete and MMIC form. Not only did the study of these "proof of concept" amplifiers help to verify the correctness of the theoretical noise modeling through the very good agreement that was observed between the measured and modeled noise performance, but their experimentally observed extremely-low, sub 2 K, noise temperatures also helped to demonstrate that SiGe amplifiers can compete with state-of-the-art InP amplifiers in the frequency range below 5 GHz. Thus, in this dissertation, it has been shown that commercial SiGe HBTs are a very strong contender for extremely low-noise cryogenic applications.

10.1 Suggestions for Future Work

In the last section of the introduction, it was stated that the literature regarding the small-signal and noise performance of SiGe HBTs at temperatures below 77 K is quite limited. While certain aspects of their cryogenic operation have been studied in detail in this dissertation, it is the author's belief that the work reported here is just the tip of the iceberg and that there is still an enormous amount of room to make meaningful contributions to the field. Firstly, it was seen in this work that several of the small- and large-signal model parameters vary significantly from foundry to foundry. Thus, a study in which the device fabrication is varied in a controlled manner would aid greatly in understanding what factors influence the changes in small-signal model parameters. For instance, in order to try to determine if the rise in emitter resistance that was observed below ∼50 K is due to carrier freeze-out effects in the intrinsic emitter, an experiment could be carried out in which otherwise identically constructed devices were processed with different emitter annealing durations.

Secondly, in Section 5.2.3, it was seen that both the electrons and holes behave as if they are at an elevated temperature when the devices are cooled to cryogenic temperatures. While the increase in effective electron temperature is generally explained in the literature as being related to non-equilibrium (i.e. ballistic) transport through the very thin base region, this explanation does not apply to the holes which are back injected into the emitter and consistently act as if they are at a warmer temperature than the presumably ballistic electrons. Thus, a device level investigation would help to explain why the base current behaves in this manner.

Another very important issue that has not been addressed in the experimental portion of this work is the correlation of the shot-noise sources in the noise model. As the correlation term was ignored in the noise modeling work, the noise at high frequencies has been overestimated. However, despite the fact that the correlation has been disregarded, the noise performance of SiGe devices at cryogenic temperatures was found to be only slightly worse than that of state-of-the-art 0.1 μm InP transistors. Thus, the determination of the correlation delay term, a frequency independent constant, would be very interesting as it would allow for a less pessimistic comparison of SiGe with InP at cryogenic temperatures. The correlation term could be determined by mounting a device that has near 50Ω optimum source resistance in a module and measuring the 50 Ω noise using the cold attenuator method. As the delay term is a single number, it should be relatively easily determined using statistical methods provided the transistor embedding network[1] and its associated losses are accurately known.

[1] i.e. the network connecting from the plane of the base contact to the end of the bondwire connecting to a 50Ω input line.

In summary, the noise and small-signal properties of SiGe HBTs have been investigated in detail, beginning with the basic theoretical physics governing the cryogenic operation of SiGe HBTs and continuing all the way through the application of the devices in state-of-the-art extremely low-noise amplifiers. The results obtained in the work are very promising and demonstrate that it is now possible to successfully use commercial SiGe HBTs in applications which were once restricted to III-V HEMTs. However, despite the successes reported here, there is still ample room for basic research in how the device fabrication relates to the dc, RF, and noise performance of SiGe HBTs below 77 K, as well as applied research in the application of SiGe HBTs to extremely low-noise amplifiers in the frequency range above 10 GHz, where the shot-noise correlation is expected to have a significant impact.

Bibliography

[1] S. Weinreb, "Low-noise cooled GASFET amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 28, pp. 1041–1054, Oct. 1980.

[2] F. Schwierz and C. Schippel, "Performance trends of Si-based RF transistors," *Electron Devices and Solid-State Circuits, 2005 IEEE Conference on*, pp. 299–304, Dec. 2005.

[3] R. Lai, X. Mei, W. Deal, W. Yoshida, Y. Kim, P. Liu, J. Lee, J. Uyeda, V. Radisic, M. Lange, T. Gaier, L. Samoska, and A. Fung, "Sub 50 nm InP HEMT device with fmax greater than 1 THz," *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pp. 609–611, Dec. 2007.

[4] I. Post, M. Akbar, G. Curello, S. Gannavaram, W. Hafez, U. Jalan, K. Komeyii, J. Lin, N. Lindert, J. Park, J. Rizk, G. Sacks, C. Tsai, D. Yeh, P. Bai, and C.-H. Jan, "A 65nm CMOS SOC technology featuring strained silicon transistors for RF applications," *Electron Devices Meeting, 2006. IEDM '06. International*, pp. 1–3, Dec. 2006.

[5] Z. Griffith, E. Lind, M. Rodwell, X.-M. Fang, D. Loubychev, Y. Wu, J. Fastenau, and A. Liu, "Sub-300 nm InGaAs/InP type-i DHBTs with a 150 nm collector, 30 nm base demonstrating 755 GHz fmax and 416 GHz ft," *Indium Phosphide and Related Materials, 2007. IPRM '07. IEEE 19th International Conference on*, pp. 403–406, May 2007.

[6] M. Pospieszalski, "Extremely low-noise amplification with cryogenic FETs and HFETs:1970–2004," *IEEE Microwave Magazine*, vol. 6, pp. 62–75, Sept. 2005.

[7] N. Jarosik, "Measurements of the low-frequency-gain fluctuations of a 30-GHz high-electron-mobility-transistor cryogenic amplifier," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 44, no. 2, pp. 193–197, Feb 1996.

[8] J. C. Bardin and S. Weinreb, "Experemental modeling and noise of SiGe HBTs," in *Proc. IEEE IMS*, Atlanta, GA, June 2008, pp. 459–462.

[9] A. Mellberg, N. Wadefalk, I. Angelov, E. Choumas, E. Kollberg, N. Rorsman, P. Starski, J. Stenarson, and H. Zirath, "Cryogenic 2-4 GHz ultra low noise amplifier," *Microwave Symposium Digest, 2004 IEEE MTT-S International*, vol. 1, pp. 161–163 Vol.1, June 2004.

[10] J. Pandian, L. Baker, G. Cortes, P. Goldsmith, A. Deshpande, R. Ganesan, J. Hagen, L. Locke, N. Wadefalk, and S. Weinreb, "Low-noise 6-8 GHz receiver," *Microwave Magazine, IEEE*, vol. 7, no. 6, pp. 74–84, Dec. 2006.

[11] R. Grundbacher, R. Lai, M. Barsky, R. Tsai, T. Gaier, S. Weinreb, D. Dawson, J. Bautista, J. Davis, N. Erickson, T. Block, and A. Oki, "0.1 m InP HEMT devices and MMICs for cryogenic low noise amplifiers from X-band to W-band," *Indium Phosphide and Related Materials Conference, 2002. IPRM. 14th*, pp. 455–458, 2002.

[12] Y.-L. Tang, N. Wadefalk, M. Morgan, and S. Weinreb, "Full Ka-band high performance InP MMIC LNA module," *Microwave Symposium Digest, 2006. IEEE MTT-S International*, pp. 81–84, June 2006.

[13] J. B. Johnson, "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, no. 1, p. 97, Jul 1928.

[14] H. Nyquist, "Thermal agitation of electric charge in conductors," *Phys. Rev.*, vol. 32, no. 1, pp. 110–113, Jul 1928.

[15] J. Engberg and T. Larsen, *Noise Theory of Linear and Nonlinear Circuits*. New York: John Wiley and Sons, 1995.

[16] C. Claeys and E. Simoen, "Radiation effects and low-frequency noise in silicon technologies," in *Low Temperature Electronics: Physics, Devices, Circuits, and Applications*, E. Gutierrez-D., M. Deen, and C. Claeys, Eds. San Diego: Academic Press, 2001, pp. 385–488.

[17] S. Wedge, "Computer-aided design of low noise microwave circuits," Ph.D. dissertation, California Institute of Technology, Pasadena, California, 1991.

[18] R. Meys, "Noise properties of linear microwave devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-26, pp. 34–37, Jan. 1978.

[19] R. Lane, "The determination of device noise parameters," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1461–1462, Aug. 1969.

[20] M. Mitama and H. Katoh, "An improved computational method for noise parameter measurement," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 27, no. 6, pp. 612–615, Jun 1979.

[21] S. Wedge and D. Rutledge, "Noise waves and passive linear multiports," *Microwave and Guided Wave Letters, IEEE*, vol. 1, no. 5, pp. 117–119, May 1991.

[22] H. Bosma, "On the theory of linear noisy systems," Ph.D. dissertation, Technische Universiteit Eindhoven, Eindhoven, NL, 1967.

[23] S. Delcourt, G. Dambrine, N. Bourzgui, S. Lepilliet, C. Laporte, J.-P. Fraysse, and M. Maignan, "On-wafer high frequency noise power measurements under cryogenic conditions: a new de-embedding approach," *Microwave Conference, 2004. 34th European*, vol. 2, pp. 913–916, Oct. 2004.

[24] G. Dambrine, H. Happy, F. Danneville, and A. Cappy, "A new method for on wafer noise measurement," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 41, no. 3, pp. 375–381, Mar 1993.

[25] M. Pospieszalski, "Modeling of noise parameters of MESFETs and MODFETs and their frequency and temperature dependence," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 37, no. 9, pp. 1340–1350, Sep 1989.

[26] "Agilent N4000A, N4001A, N4002A SNS series noise sources 10 MHz to 26.5 GHz product overview," Agilent, Tech. Rep., July 2008. [Online]. Available: http://cp.literature.agilent.com/litweb/pdf/5988-0081EN.pdf

[27] J. Cressler, J. Comfort, E. Crabbe, G. Patton, J. Stork, J.-C. Sun, and B. Meyerson, "On the profile design and optimization of epitaxial Si- and SiGe-base bipolar technology for 77 K applications. I. transistor DC design considerations," *Electron Devices, IEEE Transactions on*, vol. 40, no. 3, pp. 525–541, Mar 1993.

[28] J. Cressler, E. Crabbe, J. Comfort, J. Stork, and J.-C. Sun, "On the profile design and optimization of epitaxial Si- and SiGe-base bipolar technology for 77 K applications. II. circuit performance issues," *Electron Devices, IEEE Transactions on*, vol. 40, no. 3, pp. 542–556, Mar 1993.

[29] D. M. Richey, A. J. Joseph, J. D. Cressler, and R. C. Jaeger, "Evidence for non-equilibrium base transport in Si and SiGe bipolar transistors at cryogenic temperatures," *Solid-State Electronics*, vol. 39, no. 6, pp. 785 – 789, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/B6TY5-3VTNT9C-11/2/c38be78fc3511c79fc7bc30460e9fb2c

[30] J. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistors*. Boston, MA: Artech House, Mar. 2003.

[31] J. Yuan, R. Krithivasan, J. Cressler, M. Khater, D. Ahlgren, and A. Joseph, "On the frequency limits of SiGe HBTs for terahertz applications," *Bipolar/BiCMOS Circuits and Technology Meeting, 2007. BCTM '07. IEEE*, pp. 22–25, 30 2007-Oct. 2 2007.

[32] P. Chevalier, N. Zerounian, B. Barbalat, F. Aniel, and A. Chantre, "On the use of cryogenic measurements to investigate the potential of Si/SiGe:C HBTs for terahertz operation," *Bipolar/BiCMOS Circuits and Technology Meeting, 2007. BCTM '07. IEEE*, pp. 26–29, 30 2007-Oct. 2 2007.

[33] E. Ramirez-Garcia, N. Zerounian, P. Crozat, M. Enciso-Aguilar, P. Chevalier, A. Chantre, and F. Aniel, "SiGe heterojunction bipolar transisor issues towards high cryogenic performances," *Cryogenics*, 2009.

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

[35] R. Krithivasan, Y. Lu, J. Cressler, J.-S. Rieh, M. Khater, D. Ahlgren, and G. Freeman, "Half-terahertz operation of SiGe HBTs," *Electron Device Letters, IEEE*, vol. 27, no. 7, pp. 567–569, July 2006.

[36] J.-L. Olvera-Cervantes, J. Cressler, J.-L. Medina-Monroy, T. Thrivikraman, B. Banerjee, and J. Laskar, "A new analytical method for robust extraction of the small-signal equivalent circuit for SiGe HBTs operating at cryogenic temperatures," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 56, no. 3, pp. 568–574, March 2008.

[37] S. Pruvost, S. Delcourt, I. Telliez, M. Laurens, N. Bourzgui, F. Danneville, A. Monroy, and G. Dambrine, "Microwave and noise performance of SiGe BiCMOS HBT under cryogenic temperatures," *IEEE Electron Device Letters*, vol. 26, pp. 105–108, Feb. 2005.

[38] B. Banerjee, S. Venkataraman, C.-H. Lee, and J. Laskar, "Broadband noise modeling of SiGe HBT under cryogenic temperatures," in *Proc. IEEE RFIC*, Honolulu, HI, June 2007, pp. 765–768.

[39] S. Weinreb, J. C. Bardin, and H. Mani, "Design of cryogenic SiGe low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, pp. 2306–2312, Nov. 2007.

[40] J. C. Bardin, H. Mani, and S. Weinreb, "Silicon germanium (SiGe) low-noise amplifiers for radio astronomy," in *URSI General Assembly*, Chicago, IL, Aug. 2008, poster presentation. [Online]. Available: http://www.its.caltech.edu/~jbardin/papers/URSI_Sige_POSTER.pdf

[41] J. C. Bardin and S. Weinreb, "A 0.5–20GHz quadrature downconverter," in *Proc. IEEE BCTM*, Monterrey, CA, Oct. 2008, pp. 186–189.

[42] ——, "A 0.1-5 GHz cryogenic SiGe MMIC LNA," *Microwave and Wireless Components Letters, IEEE*, vol. 19, no. 6, pp. 407–409, June 2009.

[43] S. Weinreb, J. Bardin, H. Mani, and G. Jones, "Matched wideband low-noise amplifiers for radio astronomy," *Review of Scientific Instruments*, vol. 80, no. 4, p. 044702, 2009. [Online]. Available: http://link.aip.org/link/?RSI/80/044702/1

[44] P. Pütz, F. Schomacker, M. Schultz, M. Justen, F. Schmülling, N. Honingh, K. Jacobs, J. Bardin, G. Jones, H. Mani, and S. Weinreb, "System performance of NbTiN THz SHEB waveguide mixers and cryogenic SiGe LNA," in *at The International Space Teraherz Meeting*, Charlottseville, VA, 2009, oral presentation.

[45] J. C. Bardin and S. Weinreb, "On the prospect of commercial SiGe devices for cryogenic low-noise applications," *In preparation for IEEE Trans. Elecron Devices.*

[46] ——, "The systematic design of a 0.8–2.5GHz cryogenic LNA using 0.17μm SiGe HBTs," *In preparation for IEEE Trans. Microw. Tech.*, 2009, in preparation.

[47] W. Shockley, "Circuit element utilizing semiconductive material," United States Patent 2,569,347, Sept. 1951.

[48] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceedings of the IRE*, vol. 45, no. 11, pp. 1535–1537, Nov. 1957.

[49] D. Harame and B. Meyerson, "The early history of IBM's SiGe mixed signal technology," *Electron Devices, IEEE Transactions on*, vol. 48, no. 11, pp. 2555–2567, Nov 2001.

[50] "Physics 2000," 2009. [Online]. Available: http://nobelprize.org/nobel_prizes/physics/laureates/2000/index.html

[51] H. Kroemer, "Nobel lecture: Quasielectric fields and band offsets: teaching electrons new tricks," *Reviews of Modern Physics*, vol. 73, pp. 783–793, July 2001.

[52] B. Streetman and S. Banerjee, *Solid State Electronic Devices*, 5th ed. Prentice Hall, 2000.

[53] U. Mishra, "Ece 132 course notes," University of CA, Santa Barbara. [Online]. Available: http://my.ece.ucsb.edu/mishra/ece132.htm\#notes132

[54] P. Ashburn, *SiGe Heterojunction Bipolar Transistors*. Hoboken, NJ: John Wiley and Sons, 2003.

[55] S. Sze, *Physics of Semiconductor Devices*. New York, NY: John Wiley and Sons, 1981.

[56] I. Anteney, G. Lippert, P. Ashburn, H. Osten, B. Heinemann, and G. Parker, "Electrical determination of bandgap narrowing and parasitic energy barriers in SiGe and SiGe:C heterojunction bipolar transistors," in *Proc. 1997 Workshop on High Performance Electron Devices for Microwave and Optoelectronic Applications*, 1997, pp. 55–60.

[57] E. Johnson, "Physical limitations on frequency and power parameters of transistors," *IRE International Convention Record*, vol. 13, pp. 27–34, Mar 1965.

[58] H. Kroemer, "Quasi-electric and quasi-magnetic fields in nonuniform semiconductors," *RCA Review*, vol. 18, pp. 332–342, 1957.

[59] T. Meister, H. Schäfer, W. Perndl, and J. Böch, "Industry examples at the state-of-the-art: Infineon," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 283–293.

[60] P. Chevalier, "High-speed SiGe BiCMOS technologies for ethernet communications from 10Gb/s to 100Gb/s," Monterey, CA, 2008, IEEE Bipolar-Circuits Technology Meeting Short Course Notes.

[61] M. Chrzanowska-Jeske and R. Jaeger, "BILOW-simulation of low-temperature bipolar device behavior," *Electron Devices, IEEE Transactions on*, vol. 36, no. 8, pp. 1475–1488, Aug 1989.

[62] P. Kemph, "Industry examples at the state-of-the-art: JAZZ," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 283–293.

[63] B. A. Floyd, "SiGe BiCMOS transceivers for millimeter-wave," Boston, MA, 2007, IEEE Bipolar-Circuits Technology Meeting Short Course Notes.

[64] J. Pierce, "Physical sources of noise," *Proceedings of the IRE*, vol. 44, no. 5, pp. 601–608, May 1956.

[65] A. van der Ziel, "Theory of shot noise in junction diodes and junction transistors," *Proceedings of the IRE*, vol. 43, no. 11, pp. 1639–1646, Nov. 1955.

[66] G. Niu, J. D. Cressler, S. Zhang, W. E. Ansley, C. S. Webster, and D. L. Harame, "A unified approach to RF and microwave noise parameter modeling in bipolar transistors," *IEEE Transactions on Electron Devices*, pp. 2568–2574, Nov. 2001.

[67] M. Rudolph, *Introduction to Modeling HBTs*. Norwood, MA: Artech House Inc., 2006.

[68] J. Lange, "Noise characterization of linear twoports in terms of invariant parameters," *Solid-State Circuits, IEEE Journal of*, vol. 2, no. 2, pp. 37–40, Jun 1967.

[69] M. Pospieszalski, "On certain noise properties of field-effect and bipolar transistors," *Microwaves, Radar & Wireless Communications, 2006. MIKON 2006. International Conference on*, pp. 1127–1130, May 2006.

[70] W. Webster, "On the variation of junction-transistor current-amplification factor with emitter current," *Proceedings of the IRE*, vol. 42, no. 6, pp. 914–920, June 1954.

[71] J. Kirk, C.T., "A theory of transistor cutoff frequency fT falloff at high current densities," *Electron Devices, IRE Transactions on*, vol. 9, no. 2, pp. 164–174, March 1962.

[72] W. Bludau, A. Onton, and W. Heinke, "Temperature dependence of the band gap of silicon," *Journal of Applied Physics*, vol. 45, no. 4, pp. 1846–1848, 1974. [Online]. Available: http://link.aip.org/link/?JAP/45/1846/1

[73] E. A. Gutierrez-D., *Physics of Semiconductor Devices at Cryogenic Temperatures*. New York: Academic Press, 2001, ch. Low Temperature Electronics: Physics, Devices, Circuits, and Applications, pp. 1–103.

[74] M. A. Green, "Intrinsic concentration, effective densities of states, and effective mass in silicon," *Journal of Applied Physics*, vol. 67, pp. 2944–2954, Mar. 1990.

[75] A. B. Sproul and M. A. Green, "Intrinsic carrier concentration and minority-carrier mobility of silicon from 77 to 300 K," *Journal of Applied Physics*, vol. 73, no. 3, pp. 1214–1225, 1993. [Online]. Available: http://link.aip.org/link/?JAP/73/1214/1

[76] J. Bardeen and W. Shockley, "Deformation potentials and mobilities in non-polar crystals," *Phys. Rev.*, vol. 80, no. 1, pp. 72–80, Oct 1950.

[77] D. Caughey and R. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2192–2193, Dec. 1967.

[78] S. Li and R. Thurber, "The dopant density and temperature dependence of electron mobility and resistivity in n-type silicon," *Solid-State Electronics*, pp. 609–616, 1977.

[79] J. Dorkel and P. Leturcq, "Carrier mobilities in silicon semi-empirically related to temperature, doping, and injection level," *Solid-State Electronics*, vol. 24, pp. 821–825, 1981.

[80] "Electrical resistivity of pure metals," in *CRC Handbook of Chemistry and Physics*, 85th ed., D. R. Lide, Ed. Boca Raton, FL: CRC Press, 2004, pp. 12–43–12–44.

[81] R. M. Warner and J. N. Fordemwalt, Eds., *Integrated Circuits: Design Principles and Fabrication*, ser. Motorola Series in Solid-State Electronics. New York: McGraw-Hill Book Company, 1965.

[82] E. Conwell, "Properties of silicon and germanium: II," *Proceedings of the IRE*, vol. 46, no. 6, pp. 1281–1300, June 1958.

[83] A. Jonscher, "Semiconductors at cryogenic temperatures," *Proc. IEEE*, pp. 1092–1104, Oct. 1964.

[84] N. Mott, "On the transition to metallic conduction in semiconductors," *Canadian Journal of Physics*, vol. 34, pp. 1356–1368, 1956.

[85] P. P. Altermatt, A. Schenk, B. Schmithüsen, and G. Heiser, "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," *Journal of Applied Physics*, vol. 100, no. 11, p. 113715, 2006. [Online]. Available: http://link.aip.org/link/?JAP/100/113715/1

[86] P. P. Altermatt, A. Schenk, and G. Heiser, "A simulation model for the density of states and for incomplete ionization in crystalline silicon. I. establishing the model in Si:P," *Journal of Applied Physics*, vol. 100, no. 11, p. 113714, 2006. [Online]. Available: http://link.aip.org/link/?JAP/100/113714/1

[87] P. Chevalier, Private communication, 2008.

[88] M. Law, E. Solley, M. Liang, and D. Burk, "Self-consistent model of minority-carrier lifetime, diffusion length, and mobility," *Electron Device Letters, IEEE*, vol. 12, no. 8, pp. 401–403, Aug 1991.

[89] J. Dziewior and W. Schmid, "Auger coefficients for highly doped and highly excited silicon," *Applied Physics Letters*, vol. 31, no. 5, pp. 346–348, 1977. [Online]. Available: http://link.aip.org/link/?APL/31/346/1

[90] S. Sokolic and S. Amon, "Modelling base transport properties of npn SiGe HBT," *Semiconductor Conference, 1996., International*, vol. 1, pp. 263–266 vol.1, Oct 1996.

[91] D. L. Harame, "Device structures and BiCMOS integration," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 215–231.

[92] M. Khater, "Present status and future directions for SiGe HBT technology," *Proceedings of the 2006 Lester Eastman Conference on High Performance Devices*, p. 13, Aug. 2006.

[93] B. Barbalat, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, A. Lachater, F. Saguin, N. Zerounian, F. Aniel, and A. Chantre, "The effect of carbon on neutral base recombination in high-speed SiGeC heterojunction bipolar transistors," *Semiconductor Science Technology*, vol. 22, pp. S99–S102, 2007.

[94] H. J. Osten, "Carbon doping of SiGe," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 157–169.

[95] B. Orner, Q. Liu, B. Rainey, A. Stricker, P. Geiss, P. Gray, M. Zierak, M. Gordon, D. Collins, V. Ramachandran, W. Hodge, C. Willets, A. Joseph, J. Dunn, J.-S. Rieh, S.-J. Jeng, E. Eld, G. Freeman, and D. Ahlgren, "A 0.13 um BiCMOS technology featuring a 200/280 GHz (fT/fmax) SiGe HBT," *Bipolar/BiCMOS Circuits and Technology Meeting, 2003. Proceedings of the*, pp. 203–206, Sept. 2003.

[96] *BiCMOS 8HP Design Manual*. IBM Microelectronics, 2008.

[97] H. Riicker, B. Heinemann, R. Barth, J. Bauer, D. Blum, D. Bolze, J. Drews, G. Fischer, A. Fox, O. Fursenko, T. Grabolla, U. Haak, W. Hoppner, D. Knoll, K. Kopke, B. Kuck, A. Mai, S. Marschmeyer, T. Morgenstern, H. Richter, P. Schley, D. Schmidt, K. Schulz, B. Tillack, G. Weidner, W. Winkler, D. Wolansky, H.-E. Wulf, and Y. Yamamototo, "SiGe BiCMOS technology with 3.0 ps gate delay," *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pp. 651–654, Dec. 2007.

[98] P. Hurwitz, *Electrical Parameters of the SBC18 Process Family.* JAZZ Semiconductor, 2008.

[99] W. van Noort, A. Rodriguez, H. Sun, F. Zaato, N. Zhang, T. Nesheiwat, F. Neuilly, J. Melai, and E. Hijzen, "BiCMOS technology improvements for microwave application," *Bipolar/BiCMOS Circuits and Technology Meeting, 2008. BCTM 2008. IEEE*, pp. 93–96, Oct. 2008.

[100] B. McConnel, "IBM announces next generation silicon germanium technology," World Wide Web electronic publication, Aug. 2005. [Online]. Available: http://www-03.ibm.com/press/us/en/pressrelease/7819.wss

[101] P. Chevalier, C. Fellous, L. Rubaldo, D. Dutartre, M. Laurens, T. Jagueneau, F. Leverd, S. Bord, C. Richard, D. Lenoble, J. Bonnouvrier, M. Marty, A. Perrotin, D. Gloria, F. Saguin, B. Barbalat, R. Beerkens, N. Zerounian, F. Aniel, and A. Chantre, "230 GHz self-aligned SiGeC HBT for 90 nm BiCMOS technology," *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting*, pp. 225–228, Sept. 2004.

[102] A. J. Joseph and J. S. Dunn, "Industry examples at the state-of-the-art: IBM," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 265–282.

[103] B. Jagannathan, M. Khater, F. Pagette, J.-S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. Greenberg, R. Groves, S. Jeng, J. Johnson, E. Mengistu, K. Schonenberg, C. Schnabel, P. Smith, A. Stricker, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, "Self-aligned SiGe NPN transistors with 285 GHz fMAX and 207 GHz fT in a manufacturable technology," *Electron Device Letters, IEEE*, vol. 23, no. 5, pp. 258–260, May 2002.

[104] P. J. Geiss, A. J. Joseph, R. Krishnasamy, and X. Liu, "SiGe heterojunction bipolar transistor HBT and method of fabrication," US Patent Application 0124882, May 2008.

[105] M. Dahlstrom, K. Walter, S. Von Bruns, R. Malladi, K. Newton, and A. Joseph, "Influence of the Ge profile on VBE and current gain mismatch in advanced SiGe BICMOS NPN HBT with 200 GHz fT," *Bipolar/BiCMOS Circuits and Technology Meeting, 2006*, pp. 1–4, Oct. 2006.

[106] F. Pourchon, C. Raya, N. Derrier, P. Chevalier, D. Gloria, S. Pruvost, and D. Celi, "From measurement to intrinsic device characteristics: Test structures and parasitic determination," *Bipolar/BiCMOS Circuits and Technology Meeting, 2008. BCTM 2008. IEEE*, pp. 232–239, Oct. 2008.

[107] H. Rucker, B. Heinemann, R. Barth, D. Bolze, J. Drews, U. Haak, W. Hoppner, D. Knoll, K. Kopke, S. Marschmeyer, H. Richter, P. Schley, D. Schmidt, R. Scholz, B. Tillack, W. Winkler, H.-E. Wulf, and Y. Yamamoto, "SiGe:C BiCMOS technology with 3.6 ps gate delay," *Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International*, pp. 5.3.1–5.3.4, Dec. 2003.

[108] D. Knoll, "Industry examples at the-state-of-the-art: IHP," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 321–341.

[109] M. Racanelli, K. Schuegraf, A. Kalburge, A. Kar-Roy, B. Shen, C. Hu, D. Chapek, D. Howard, D. Quon, F. Wang, G. U'ren, L. Lao, H. Tu, J. Zheng, J. Zhang, K. Bell, K. Yin, P. Joshi, S. Akhtar, S. Vo, T. Lee, W. Shi, and P. Kempf, "Ultra high speed SiGe NPN for advanced BiCMOS technology," *Electron Devices Meeting, 2001. IEDM Technical Digest. International*, pp. 15.3.1–15.3.4, 2001.

[110] P. Deixler, A. Rodriguez, W. De Boer, H. Sun, R. Colclaser, D. Bower, N. Bell, A. Yao, R. Brock, Y. Bouttement, G. Hurkx, L. Tiemeijer, J. Paasschens, H. Huizing, D. Hartskeerl, P. Agrarwal, P. Magnee, E. Aksen, and J. Slotboom, "QUBiC4X: An fT/fmax = 130/140GHz SiGe:C-BiCMOS manufacturing technology with elite passives for emerging microwave applications," *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting*, pp. 233–236, Sept. 2004.

[111] R. Colcaser and P. Deixler, "Industry examples at the state-of-the-art: Philips," in *Silicon Heterostructure Handbook*, J. Cressler, Ed. Boca Raton: CRC Press, 2006, pp. 371–386.

[112] "Zprobe data-sheet," 2004. [Online]. Available: http://www.suss.com/download?id=730

[113] J. Fernandez, "Cryogenic HEMT characterization," Anaheim, CA, 1999, course notes, Cryogenic Wafer Probing of Microwave Devices, IEEE-MTTS IMS.

[114] R. Venkatasubramanian, "Cascade cryogenic thermoelectric cooler for cryogenic and room temperature applications," United States Patent 6,662,570, Dec. 2003.

[115] "Cryotel mt measured performance," World Wide Web electronic publication. [Online]. Available: http://www.sunpower.com/lib/sitefiles/pdf/cryo_MT.pdf

[116] S. Weisberg, *Applied Linear Regression*. Hoboken, NJ: Wiley-Interscience, 2005.

[117] A. Appaswamy, J. Cressler, and G. Niu, "A novel base current phenomenon in SiGe HBTs operating in inverse mode," *Solid State Device Research Conference, 2007. ESSDERC 2007. 37th European*, pp. 350–353, Sept. 2007.

[118] S. Mason, "Power gain in feedback amplifier," *Transactions of the IRE Professional Group on Circuit Theory*, vol. CT-1, no. 2, pp. 20–25, June 1954.

[119] B. Woods, H. Mantooth, and J. Cressler, "SiGe HBT compact modeling for extreme temperatures," *Semiconductor Device Research Symposium, 2007 International*, pp. 1–2, Dec. 2007.

[120] P. Tasker and M. Fernandez-Barciela, "HBT small signal T and Pi model extraction using a simple, robust and fully analytical procedure," *Microwave Symposium Digest, 2002 IEEE MTT-S International*, vol. 3, pp. 2129–2132, 2002.

[121] K. Xia, G. Niu, D. Sheridan, and W. Ansley, "Ratio based direct extraction of small-signal parameters for SiGe HBTs," *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting*, pp. 144–147, Sept. 2004.

[122] K. Lee, K. Choi, S.-H. Kook, D.-H. Cho, K.-W. Park, and B. Kim, "Direct parameter extraction of SiGe HBTs for the VBIC bipolar compact model," *Electron Devices, IEEE Transactions on*, vol. 52, no. 3, pp. 375–384, Mar 2005.

[123] T.-R. Yang, J. M.-L. Tsai, C.-L. Ho, and R. Hu, "SiGe HBT's small-signal pi modeling," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 55, no. 7, pp. 1417–1424, July 2007.

[124] M. Dvorak and C. Bolognesi, "On the accuracy of direct extraction of the heterojunction-bipolar-transistor equivalent-circuit model parameters Cp, CBC, and RE," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, no. 6, pp. 1640–1649, June 2003.

[125] T. Ning and D. Tang, "Method for determining the emitter and base series resistances of bipolar transistors," *Electron Devices, IEEE Transactions on*, vol. 31, no. 4, pp. 409–412, Apr 1984.

[126] W. Sansen and R. Meyer, "Characterization and measurement of the base and emitter resistances of bipolar transistors," *Solid-State Circuits, IEEE Journal of*, vol. 7, no. 6, pp. 492–498, Dec 1972.

[127] C. Raya, F. Pourchon, T. Zimmer, D. Celi, and P. Chevalier, "Scalable approach for HBT's base resistance calculation," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 21, no. 2, pp. 186–194, May 2008.

[128] L. Giacoletto, "Measurement of emitter and collector series resistances," *Electron Devices, IEEE Transactions on*, vol. 19, no. 5, pp. 692–693, May 1972.

[129] T. Chen, T. Krakowski, A. Strachan, Y. Liu, A. Sadovnikov, and J. Babcock, "Simultaneous extraction of collector and substrate series resistance by simple dc measurement," *Bipolar/BiCMOS Circuits and Technology Meeting, 2006*, pp. 1–4, Oct. 2006.

[130] R. Gabl and M. Reisch, "Emitter series resistance from open-collector measurements-influence of the collector region and the parasitic pnp transistor," *Electron Devices, IEEE Transactions on*, vol. 45, no. 12, pp. 2457–2465, Dec 1998.

[131] K. Morizuka, O. Hidaka, and H. Mochizuki, "Precise extraction of emitter resistance from an improved floating collector measurement," *Electron Devices, IEEE Transactions on*, vol. 42, no. 2, pp. 266–273, Feb 1995.

[132] P. R. Grey and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed. New York: John Wiley and Sons, 1993.

[133] R. Anholt, *Electrical and Thermal Characterization of MESFETs, HEMTs, and HBTs*. Boston, MA: Artech House, 1995.

[134] K. Xia, "Improved RF noise modeling for silicon-germanium heterojunction bipolar transistors," Ph.D. dissertation, Auburn University, Auburn, AL, 2006.

[135] W. Kloosterman, J. Paasschens, and D. Klaassen, "Improved extraction of base and emitter resistance from small signal high frequency admittance measurements," *Bipolar/BiCMOS Circuits and Technology Meeting, 1999. Proceedings of the 1999*, pp. 93–96, 1999.

[136] Y. Gobert, P. Tasker, and K. Bachem, "A physical, yet simple, small-signal equivalent circuit for the heterojunction bipolar transistor," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 45, no. 1, pp. 149–153, Jan 1997.

[137] E. A. Mechtly, "Properties of materials," in *Reference Data For Engineers: Radio Electronics, Computer, and Communications*, 9th ed., W. M. Middleton, Ed. Boston: Newnes, 2002, pp. 4–1–4–33.

[138] W. Clark, B. El-Kareh, R. Pires, S. Titcomb, and R. Anderson, "Low temperature CMOS-a brief review," *Components, Hybrids, and Manufacturing Technology, IEEE Transactions on*, vol. 15, no. 3, pp. 397–404, Jun 1992.

[139] H. Friis, "Noise figures of radio receivers," *Proceedings of the IRE*, vol. 32, pp. 419–422, July 1944.

[140] H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*. New York, NY: John Wiley & Sons, Inc., 1959.

[141] H. Fukui, "Available power gain, noise figure, and noise measure of two-ports and their graphical representations," *IEEE Trans. Circuit Theory*, vol. CT-13, pp. 137–142, June 1966.

[142] H. Hillbrand and P. Russer, "An efficient method for computer aided noise analysis of linear amplifier networks," *Circuits and Systems, IEEE Transactions on*, vol. 23, no. 4, pp. 235–238, Apr 1976.

[143] D. Frickey, "Conversions between S, Z, Y, H, ABCD, and T parameters which are valid for complex source and load impedances," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 42, no. 2, pp. 205–211, Feb 1994.

[144] H. Haus, W. Atkinson, G. Branch, W. Davenport, W. Fonger, W. Harris, S. Harrison, W. McLeod, E. Stodola, and T. Talpey, "Representation of noise in linear twoports," *Proceedings of the IRE*, vol. 48, pp. 69–74, Jan. 1960.

[145] R. Tucker, "Low-noise design of microwave transistor amplifiers (short papers)," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 23, no. 8, pp. 697–700, Aug 1975.

[146] C. Poole and D. Paul, "Optimum noise measure terminations for microwave transistor amplifiers (short paper)," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 33, no. 11, pp. 1254–1257, Nov 1985.

[147] J.-C. Liu, S.-S. Bor, and P.-C. Lu, "Comments on 'optimum noise measure terminations for microwave transistor amplifiers' by C.R. Poole and D.K. Paul [and reply]," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 41, no. 11, pp. 2042–2043, Nov 1993.

[148] N. Wadefalk and S. Weinreb, "Very low noise amplifiers for very large arrays," 2005, iEEE Microwave Symposium Short Course.

[149] P. Leroux and M. Steyaert, *LNA-ESD co-design for fully integrated CMOS wireless receivers*. Dordrecht, NL: Springer Press, 2005.

[150] Dielectric Labs, "Single-layer and broadband blocking capacitors." [Online]. Available: {www.dilabs.com/pdfs/SLC\%20Catalog.pdf}

[151] J. W. Kooi, J. J. A. Baselmans, M. Hajenius, J. R. Gao, T. M. Klapwijk, P. Dieleman, A. Baryshev, and G. de Lange, "IF impedance and mixer gain of NbN hot electron bolometers," *Journal of Applied Physics*, vol. 101, no. 4, p. 044511, 2007. [Online]. Available: http://link.aip.org/link/?JAP/101/044511/1

[152] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New Dehli: Tata McGraw-Hill, 2002.

[153] S. Iversen, "The effect of feedback on noise figure," *Proceeding of the IEEE*, vol. 63, pp. 540–5420, Mar. 1975.

[154] J. E. Fernandez, "A noise measurement system using a cryogenic attenuator," The Jet Propulsion Laboratory, California Institute of Technology, TMO Progress Report 42-135F, Nov. 1998. [Online]. Available: http://ipnpr.jpl.nasa.gov/progress_report/42-135/135F.pdf

[155] J. Moll and I. Ross, "The dependence of transistor parameters on the distribution of base layer resistivity," *Proceedings of the IRE*, vol. 44, no. 1, pp. 72–78, Jan. 1956.

[156] E. Heasell, "The derivation of a relationship for the charge-control base lifetime in terms of the base and collector small-signal conductances," *Electron Devices, IEEE Transactions on*, vol. 28, pp. 595–596, May 1981.

[157] H. Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region," *Solid-State Electronics*, vol. 28, pp. 1101–1103, Nov. 1985.

[158] S. A. Wartenberg, *RF Measurements of Die and Packages*. Boston, MA: Artech House, 2002.

[159] M. Myslinski, W. Wiatr, and D. Schreurs, "A three-step procedure utilizing only two test structures for deembedding transistor from on-wafer s-parameter measurements," *Microwaves, Radar and Wireless Communications, 2004. MIKON-2004. 15th International Conference on*, vol. 2, pp. 674–677 Vol.2, May 2004.

[160] P. Penfield, "Wave representation of amplifier noise," *IRE Transactions on Circuit Theory*, pp. 84–86, Mar. 1962.

Appendix A

BJT Derivations

A.1 Terminal Currents in Si Transistors

In this section, the collector and base current densities are derived under the assumption that the base width is very small.

A.1.1 Base Current Density, J_B

The base current density can be calculated by computing the hole current flowing from the p-doped region into the n-doped region in a p-n diode. The first step in this calculation is to derive an expression for the ratio of hole charge on each side of the depletion region in terms of the built in voltage, V_0. Next, this expression will be generalized to include the effect of bias voltages. Finally, the excess diffusion current density will be calculated. For all calculations, it is assumed that the emitter width is greater than a diffusion length, that the junction is abrupt, and that the doping levels are constants within each region. The expression for the hole current is [55]:

$$J_p(x) = q\mu_p p(x)\,\mathcal{E}(x) - qD_p\frac{\partial p(x)}{\partial x}, \tag{A.1}$$

where $p(x)$ is the hole concentration in the semiconductor and x is the position along the device[1].

At equilibrium, the hole current density must be zero. In the regions outside the depletion region, this has obvious implications— the doping is constant so $\partial p(x)/\partial x = 0$, forcing the electric-field, $\mathcal{E}(x)$, to be non-existent. However, within the depletion region, $\partial p(x)/\partial x \neq 0$ meaning there must be a diffusion current flowing. Thus, a *built-in voltage* which produces a compensating drift current has to exist. This built-in voltage can be found by solving a partial-differential equation derived

[1] It is assumed that the device is uniform in the yz plane so the analysis of a 1-d device is sufficient. This assumption is equivalent to neglecting end effects.

from equation (A.1) under equilibrium conditions:

$$-\frac{1}{V_T}\frac{\partial V(x)}{\partial x} = \frac{1}{p(x)}\frac{\partial p(x)}{\partial x}, \tag{A.2}$$

where $V_T = kT/q$ is the thermal voltage. After performing integration, the expression for the built-in voltage is obtained:

$$V_0 = V_n - V_f = V_T \ln\left(\frac{p_{p0}}{p_{n0}}\right) \tag{A.3}$$

or

$$\frac{p_{p0}}{p_{n0}} = e^{V_0/V_T}. \tag{A.4}$$

where p_{p0} and p_{n0} are the equilibrium hole concentrations in the p and n regions.

Thus far, all calculations have assumed equilibrium conditions. However, the concentrations at the edge of the depletion region can be related to an applied bias through a modification of equation (A.4) [52]:

$$\frac{p_n(-x_{n0})}{p_p(x_{p0})} = e^{(V_{BE}-V_0)/V_T}, \tag{A.5}$$

where V_{BE} is the applied voltage with polarity defined from the p-region to the n-region. Assuming that the majority charge density is relatively independent of bias, the excess hole charge on the n-doped side of the depletion region can be written as

$$\Delta p_n(-x_{n0}) = p_{n0}\left(e^{V_{BE}/V_T} - 1\right). \tag{A.6}$$

From equation (A.1), we see that the excess diffusion current depends on the position dependent charge concentration. For a long emitter, the average injected hole will recombine in within a hole diffusion length, L_p [55]. Therefore, the position dependent excess minority charge distribution in the emitter is given as:

$$\Delta p_n(x) = p_n(-x_{n0})\, e^{(x+x_{n0})/L_p} = p_n(-x_{n0})\, e^{-(x')/L_p}, \tag{A.7}$$

where, to simplify calculations, a new coordinate $x' = -(x + x_{n0})$ has been defined such that the emitter area is in the region $x' > 0$. Now, the excess hole current can be directly calculated at the edge of the depletion region:

$$J_p(x' = 0) = -qD_n\frac{\partial p(x')}{\partial x} = p_n\frac{kT_a\mu_p}{L_p}e^{V_{BE}/V_T} \tag{A.8}$$

Realizing that, in the absence of recombination in the base, $J_B = J_p$ and $p_n = n_{io}^2/N_{DE}^+ = N_{AB}^- e^{-V_0/V_T}$, we can write the final formula for the base current density:

$$J_B = \frac{kT_a\mu_p n_{io}^2}{N_{DE}^+ L_{pe}} e^{V_{BE}/V_T} = N_{AB}^- \frac{kT_a\mu_p}{L_{pe}} e^{(V_{BE}-V_0)/V_T} \tag{A.9}$$

A.1.2 Collector Current Density, J_C

To begin in deriving the collector current density, we determine the built in electric field using by equating the hole drift and diffusion currents [52]:

$$q\mu_p(x)\,\mathcal{E}(x) = D_p\frac{dp(x)}{dx}. \tag{A.10}$$

So, the built in electric field is:

$$\mathcal{E}(x) = \frac{V_T}{p(x)}\frac{dp(x)}{dx}. \tag{A.11}$$

Now, the total electron current can be written as:

$$J_n(x) = q\mu_n n(x)\,\mathcal{E}(x) + qD_n\frac{dn(x)}{dx}. \tag{A.12}$$

Substituting equation (A.11) into equation (A.12) and using Einstein's relation ($D = \mu V_T$), we can write the following equation for the electron current density in the semiconductor:

$$J_n(x) = qD_n\left[\frac{n(x)}{p(x)}\frac{dp(x)}{dx} + \frac{dn(x)}{dx}\right] = \frac{qD_n}{p(x)}\frac{d\{n(x)\,p(x)\}}{dx}. \tag{A.13}$$

Next, integration can be carried out from position x to the edge of the base yielding the following expression[2]:

$$\frac{J_n(x)}{qD_n}\int_x^{W_B} p(x')\,dx' = n(x)\,p(x). \tag{A.14}$$

Thus, the minority carrier concentration at the edge of the base side of the base-emitter space charge region is given as:

$$n(0) = \frac{J_n(0)}{qD_n p(0)}\int_0^{W_B} p(x)\,dx. \tag{A.15}$$

[2]The boundary condition $n(W_B) = 0$ has been applied.

Now, we know from standard diffusion theory[3] that the electron concentration under active bias is related to its equilibrium value by an exponential term that is a function of the applied bias:

$$n(0) = n_{p0} e^{V_{BE}/V_T}. \tag{A.16}$$

Thus, we can solve for the collector current density by equating (A.15) with (A.16):

$$J_C = J_n(0) = \frac{qD_n p(0) n_{p0} e^{V_{BE}/V_T}}{\int_0^{W_B} p(x)\, dx} = kT \frac{\mu_n p(0) n_{p0} e^{V_{BE}/V_T}}{\int_0^{W_B} p(x)\, dx}. \tag{A.17}$$

Equation (A.17) is known as the Moll-Ross relation for bipolar transistors [155, 156], and relates the total majority carrier charge in the base to the collector current density. Finally, if we assume that a constant acceptor doping of N_{AB}^- is used in the base and that the majority carrier concentration at the base side of the space charge region is insensitive to applied bias, then $p(x) = p(0) = N_{AB}^-$, and equation (A.17) simplifies drastically:

$$J_C = kT \frac{\mu_n n_{io}^2}{W_B N_{AB}^-} e^{V_{BE}/V_T}. \tag{A.18}$$

A.2 Capacitances

A.2.1 Base-Emitter Diffusion Capacitance, C_{bed}

Charge neutrality requires that the difference in number of majority carriers and minority carriers in a doped semiconductor be equal to the total number of ionized impurities [55]. Thus, for a change of in base minority charge of ΔQ_{nb}, a base current is required to supply a change in the majority concentration of $\Delta Q_{pb} = \Delta Q_{nb}$, and we can define a small-signal base-emitter diffusion capacitance [132]:

$$C_{bed} = \frac{\Delta Q_{pb}}{\Delta V_{BE}} = \frac{\Delta Q_{nb}}{\Delta V_{BE}} = \frac{\Delta Q_{nb}}{\Delta I_C} \frac{\Delta I_C}{\Delta V_{BE}} = g_m \frac{\Delta Q_{nb}}{\Delta I_C}. \tag{A.19}$$

Recall, that the minority current in the base is given as

$$J_C = \frac{qN_{db}D_{nb}}{W_B} \left(e^{V_{BE}/V_T} - 1 \right) \approx \frac{qN_{db}D_{nb}}{W_B} e^{V/V_T}. \tag{A.20}$$

[3] see [52] or [55]

Now, the total minority carrier charge in the base is given as:

$$Q_{nb} = \int_0^{W_B} qAn_p(x)\, dx = \frac{qAn_{p0}W_B}{2} e^{V_{BE}/V_T}. \tag{A.21}$$

Taking the ratio of minority carrier charge to collector current density, we can write an expression for the minority carrier base transit time [132]:

$$\tau_f \equiv \frac{Q_{nb}}{I_C} = \frac{W_B^2}{2\mu_{nb}V_T}. \tag{A.22}$$

Thus, from A.19, the diffusion capacitance can be written as:

$$C_{bed} = g_m\tau_f = \frac{I_C W_B^2}{2\mu_{nb}V_T^2}. \tag{A.23}$$

A.2.2 Depletion Capacitances

A depletion (junction) capacitance occurs due to charge dipole in a pn junction and can be determined as [52]:

$$C_d = \left|\frac{\partial Q_d}{\partial V_d}\right|, \tag{A.24}$$

where Q_d is the charge in each half of the depletion region and V_d is the potential across the depletion region. The total charge on each side of the depletion region can be written as

$$|Q_d| = qAx_{n0}N_d = qAx_{p0}N_a, \tag{A.25}$$

where x_{p0} and x_{n0} are the positions of the edge of the depletion region in the p and n doped regions. In order to simplify (A.25), we can exploit the relationship between the dopant concentrations and applied voltage as given by Streetman [52]:

$$x_{n0} = W\frac{N_a}{N_a + N_d}, \tag{A.26}$$

where

$$W = \sqrt{\frac{2\epsilon(V_0 - V)}{q}\frac{N_a + N_d}{N_aN_d}}. \tag{A.27}$$

Thus, equation (A.25) can be written as

$$|Q_d| = qAW\frac{N_aN_d}{N_a + N_d} = A\sqrt{2q\epsilon(V_0 - V)\frac{N_aN_d}{N_a + N_d}}. \tag{A.28}$$

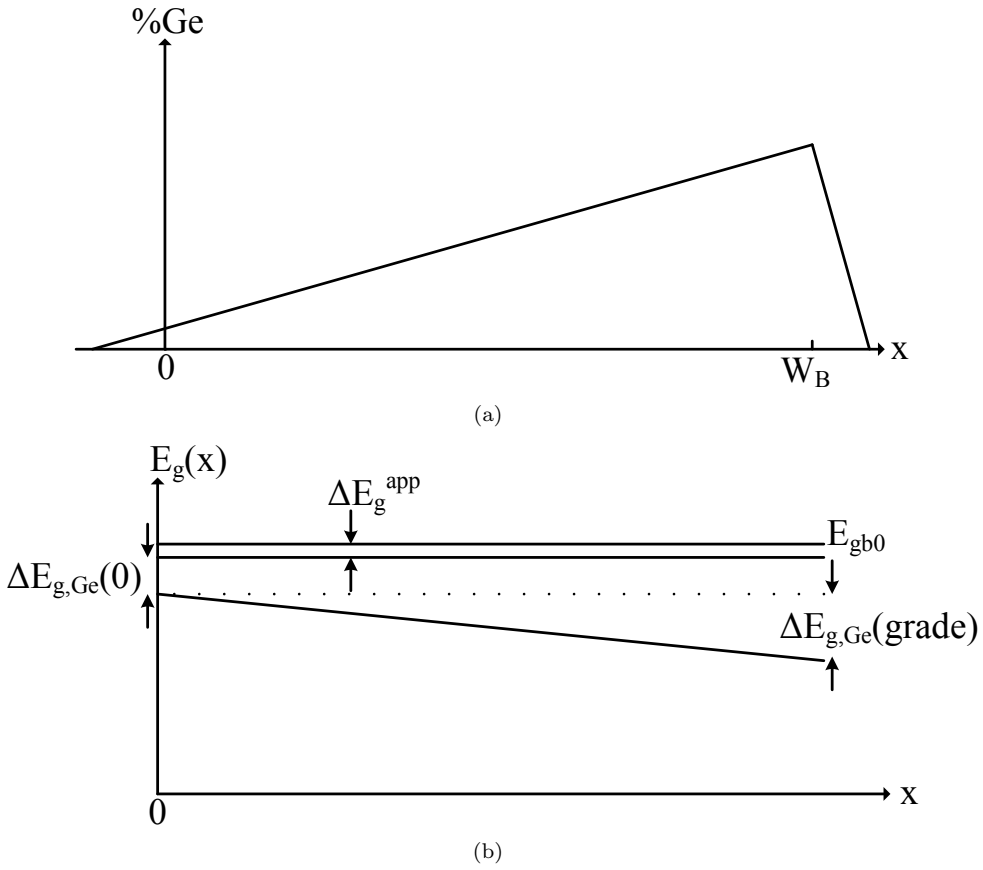


Figure A.1: (a) Drawing of Ge profile in the base. For the purpose of analysis, the depletion regions are not included in the base. (b) Bandgap structure in the base region demonstrating reduction from the Si bandgap due to Ge content and high doping levels. E_{gb0} is the bandgap for intrinsic silicon.

Finally, the an expression for the depletion capacitance can be written as

$$C_d = \epsilon A \sqrt{\frac{q}{2\epsilon\left(V_0 - V\right)} \frac{N_a N_d}{N_a + N_d}} = \frac{\epsilon A}{W}. \tag{A.29}$$

A.3 SiGe

A.3.1 Terminal Currents

As discussed in Chapter 2.1.2 the introduction of Ge to the base in a SiGe transistor results in a bandgap narrowing effect which has a strong effect on the collector current density, making the collector current a strong function of the Ge profile. In this section, the collector current will be derived for a linearly grated Ge profile, which is common in state-of-the-art SiGe HBTs. The

conventions used for labels are consistent with those used by Cressler in [30]. A representative Ge profile is drawn in Fig. A.1(a), with a sketch of the corresponding energy band diagram appearing in Fig. A.1(b). There are two effects that are accounted for in the energy band diagram:

1) A position independent reduction in the bandgap of $\Delta E_{g,app}$ resulting from heavy doping in the base.

2) A position dependent bandgap reduction due of $\Delta E_{g,Ge}(x) = \Delta E_{g,Ge}(0) + (x/W_B)\,\Delta E_{g,Ge}(grade)$ due to the introduction of the Ge grating in the base.

To determine the collector current, we can use the generalized Moll-Ross equation as derived by Kröemer [157]:

$$J_{C,SiGe} = \frac{q\left(e^{V_{BE}/V_T} - 1\right)}{\displaystyle\int_0^{W_B} \frac{p_b(x)\,dx}{D_{nb}(x)\,n_{ib}^2(x)}}, \tag{A.30}$$

where $V_T = kT/q$, p_b is the position dependent hole concentration, D_{nb} is the position dependent electron diffusion constant, and n_{ib} is the position dependent intrinsic carrier concentration. The intrinsic carrier concentration is proportional to the bandgap. Therefore, the square of the position dependent intrinsic carrier concentration in SiGe can be written as [30]:

$$n_{ib,SiGe}(x)^2 = (N_C N_V)_{SiGe}(x)\, e^{-E_{gb0}/kT} e^{\Delta E_{g,app}/kT} e^{\Delta E_{g,Ge}(0)/kT} e^{x/W_B \Delta E_{g,Ge}(grade)/kT}. \tag{A.31}$$

Next, as suggested by Cressler in [30], we define a ratio of effective density of states between SiGe and Si,

$$\gamma(x) = \frac{(N_C N_V)_{SiGe}(x)}{(N_C N_V)_{Si}}, \tag{A.32}$$

allowing us to rewrite equation (A.31) in a more compact form:

$$n_{ib,SiGe} = \gamma(x)\, n_{io,Si}^2 e^{\Delta E_{g,app}/kT} e^{\Delta E_{g,Ge}(0)/kT} e^{x/W_B \Delta E_{g,Ge}(grade)/kT}. \tag{A.33}$$

Finally, by assuming a constant level of ionized impurities in the base, N_{ab}^-, and defining a position averaged diffusion constants and density of states ratio, $\widetilde{D_{nb}}$ and $\widetilde{\gamma}$ the integral in the denominator of (A.30) becomes:

$$\frac{N_{ab}^-}{n_{io,Si}^2 \widetilde{D_{nb}} \widetilde{\gamma} e^{\Delta E_{g,app}/kT} e^{\Delta E_{g,Ge}(0)/kT}} \int_0^{W_B} e^{-x/W_B \Delta E_{g,Ge}(grade)/kT}\, dx. \tag{A.34}$$

Performing the integration and plugging the result into A.30, the collector current density can be written:

$$J_{C,SiGe} = q\frac{n_{io,Si}^2\widetilde{D_{nb}}\widetilde{\gamma}}{W_B N_{AB}^-}\frac{\Delta_{E_{g,Ge}(grade)}/kT}{1-e^{-\Delta E_{g,Ge}(grade)/kT}}e^{\Delta E_{g,app}/kT}e^{\Delta E_{g,Ge}(0)/kT}\left(e^{V_{BE}/V_T}-1\right). \quad \text{(A.35)}$$

For comparison with Si devices, it is desirable to recast equation (A.35) in such a form that $J_{C,Si}$ can be factored out of the expression. This can be done by substituting

$$\widetilde{\eta} \equiv \frac{\widetilde{D_{nb,SiGe}}}{D_{nb,Si}} \quad \text{(A.36)}$$

into equation (A.35). Doing so, the final equation for the collector current reads:

$$J_{C,SiGe} = J_{C,Si}\left[\widetilde{\eta}\widetilde{\gamma}e^{\Delta E_{g,Ge}(0)/kT}\frac{\Delta_{E_{g,Ge}(grade)}/kT}{1-e^{-\Delta E_{g,Ge}(grade)/kT}}\right]. \quad \text{(A.37)}$$

Appendix B

De-Embedding Procedure

B.1 Introduction to De-Embedding Using Y- and Z- Parameters

In general, shunt elements are best removed using a Y- parameters and series elements are best removed using Z-parameters. The two general configurations appear in Fig. B.1 and the intrinsic networks can be found as

$$\mathbf{Y} = \mathbf{Y}' - \begin{bmatrix} Y_1 + Y_2 & -Y_2 \\ -Y_2 & Y_2 + Y_3 \end{bmatrix} \tag{B.1}$$

and

$$\mathbf{Z} = \mathbf{Z}' - \begin{bmatrix} Z_1 + Z_2 & Z_2 \\ Z_2 & Z_2 + Z_3 \end{bmatrix} \tag{B.2}$$

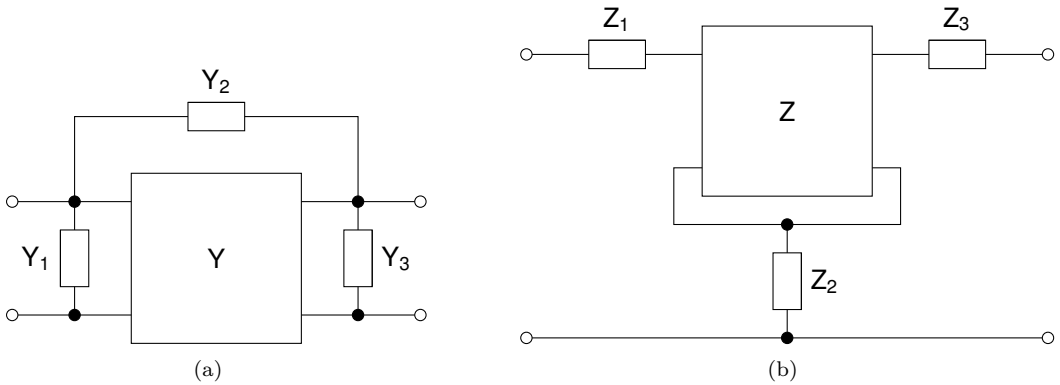


Figure B.1: Equivalent circuits for basic de-embedding

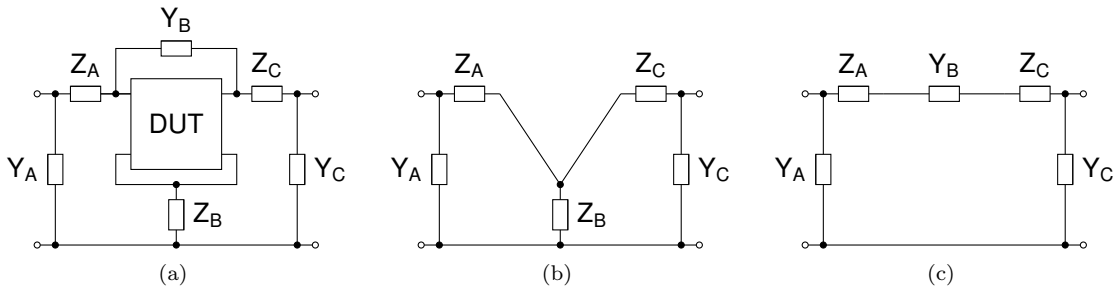


Figure B.2: Equivalent circuit diagrams of transistor, open-circuit, and short-circuit test structures.

B.2 Transistor Test-Structure De-Embedding

In this section, the procedure used to remove the parasitic effects of the bondpads and feedlines is presented. The topic of de-embedding has received considerable attention in the literature and a good summary of the techniques commonly used can be found in Chapter 6 of [158]. In this work, the three-step de-embedding procedure presented by Myslinski, Wiatr, and Schreurs is used [159]. A diagram detailing the parasitic components which need to be removed by the de-embedding appears in Fig. B.2(a). The unknown parasitics associated with the transistor test structure are: 1) Y_A and Y_C, which are the admittances from the input and output bondpads to ground, 2) Z_A and Z_C, which are the series impedances associated with the lines connecting from the bondpads to the transistor input and output terminals, 3) Y_B, which is the admittance between the input and output terminals of the transistor, and 4) Z_B, which is the series resistance in the ground return path.

If all of the parasitics are known, then the following procedure can be used to determine the intrinsic Y parameters of the DUT, $\mathbf{Y_{int}}$, from the measured Y parameters, $\mathbf{Y_m}$:

1) Remove Y_A and Y_C from $\mathbf{Y_m}$ to end up with $\mathbf{Y'_m}$

2) Invert $\mathbf{Y'_m}$ to end up with $\mathbf{Z'_m}$

3) Remove Z_A, Z_B, and Z_C from $\mathbf{Z'_m}$ to end up with $\mathbf{Z''_m}$

4) Invert $\mathbf{Z''_m}$ to end up with $\mathbf{Y''_m}$

5) Remove Y_B from $\mathbf{Y''_m}$ to end up with $\mathbf{Y_{int}}$

In order to determine the unknown parasitics, we use short- and open-circuit test structures, the equivalent circuits of which appear in Figs. B.2(b) and B.2(c). To begin with, we determine Y_A and

Y_C from the Y-parameters of the open-circuit test structure:

$$Y_A = Y_{11}^{OS} + Y_{12}^{OS} \tag{B.3}$$

and

$$Y_C = Y_{22}^{OS} + Y_{12}^{OS}. \tag{B.4}$$

With Y_A and Y_C known, the series elements can be computed from the measured admittance matrix of the short-circuit test structure as

$$Z_A = \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\mathbf{Y^{SS}} - \begin{bmatrix} Y_A & 0 \\ 0 & Y_C \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \tag{B.5}$$

$$Z_B = \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\mathbf{Y^{SS}} - \begin{bmatrix} Y_A & 0 \\ 0 & Y_C \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{B.6}$$

and

$$Z_C = \begin{bmatrix} 0 & 1 \end{bmatrix} \left(\mathbf{Y^{SS}} - \begin{bmatrix} Y_A & 0 \\ 0 & Y_C \end{bmatrix} \right)^{-1} \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{B.7}$$

Finally, with the other elements known, Y_B can be determined as

$$Y_B = \left(-\frac{1}{Y_{12}^{OS}} - Z_A - Z_C \right). \tag{B.8}$$

Appendix C

Extracted Models

C.1 DC Parameters

C.1.1 DC Current Coefficients: Low-Injection Operation

In section 5, it was shown that under low-injection conditions, the base current and collector current densities can be written as

$$J_B = J_{B0} e^{V_{BE}/n_b V_T} + J_{B0,RC} e^{V_{BE}/n_{b,rc} V_T} \tag{C.1}$$

$$J_C = J_{C0} e^{V_{BE}/n_{c0} V_T} + J_{C0,P} e^{V_{BE}/n_{c,p} V_T} \tag{C.2}$$

The coefficients J_{B0}, n_b, $J_{B0,RC}$, and $n_{b,rc}$ appear in Tables C.1-C.4. For the IHP-G4, ST-X1, and ST-X3 devices the base current at 18 K was inconsistent with the model[1] and thus, fitting coefficients are not given. Furthermore, it was not possible to extract the coefficients for the NXP-G3 device at 200 K.

The coefficients J_{C0}, n_{c0}, $J_{C0,P}$, and $n_{c,p}$ appear in Tables C.5-C.8. It should be noted that the coefficients associated with the parasitic leakage component of J_C were only required to describe a few of the devices operating at cryogenic temperatures. Specifically, they were required to describe J_C for the IBM-G4 device at 50 K and below as well as the ST-X1, ST-X3, and JAZZ-G3 devices at 77 K and below.

[1] i.e. they had additional structure besides the recombination and diffusion current components.

Table C.1: Extracted base saturation current, J_{B0}, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	1.1e-16	1.6e-25	3.2e-51	1.1e-56	1.7e-59
IHP-G4	7.8e-17	8.7e-24	7.0e-30	4.7e-30	-
ST-G4	4.8e-17	3.9e-27	5.3e-66	3.7e-80	8.2e-85
ST-X2	5.9e-17	1.2e-26	4.5e-48	2.5e-51	6.8e-55
ST-X1	1.3e-16	4.2e-25	4.5e-46	1.6e-47	-
ST-X3	1.7e-17	6.4e-25	2.1e-38	3.1e-42	-
JAZZ-G3	2.1e-16	2.6e-26	2.0e-57	4.5e-55	3.6e-56
NXP-G3	3.1e-14	-	4.0e-52	1.9e-55	2.2e-54

Table C.2: Extracted base ideality factors, n_b, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	1.022	1.076	1.434	2.006	5.324
IHP-G4	1.013	1.152	2.544	3.890	-
ST-G4	1.014	1.000	1.074	1.362	3.585
ST-X2	1.029	1.030	1.541	2.223	5.759
ST-X1	1.081	1.119	1.593	2.392	-
ST-X3	1.008	1.130	1.952	2.743	-
JAZZ-G3	1.063	1.038	1.256	2.060	5.651
NXP-G3	1.142	-	1.339	1.953	5.571

Table C.3: Extracted base recombination saturation currents, J_{br}, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	9.8e-11	1.8e-10	2.2e-17	1.2e-19	1.5e-20
IHP-G4	3.6e-12	1.9e-12	2.7e-11	1.6e-11	-
ST-G4	6.8e-12	4.7e-16	3.9e-23	7.8e-37	2.4e-40
ST-X2	5.3e-12	9.7e-11	1.1e-23	7.8e-20	5.7e-22
ST-X1	1.8e-7	1.7e-14	5.6e-16	3.8e-14	-
ST-X3	6.9e-12	3.0e-12	3.0e-17	1.7e-19	-
JAZZ-G3	7.0e-10	1.7e-11	1.5e-20	7.4e-18	1.5e-17
NXP-G3	1.4e-7	-	3.5e-12	6.7e-12	1.9e-12

Table C.4: Extracted base recombination ideality factors, n_{br}, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	3.385	6.209	5.392	7.105	18.924
IHP-G4	1.752	3.118	14.311	20.788	-
ST-G4	2.737	2.166	3.800	3.265	8.226
ST-X2	3.078	4.770	3.679	7.416	17.923
ST-X1	5.992	3.394	5.716	10.576	-
ST-X3	2.175	3.481	5.271	6.875	-
JAZZ-G3	3.219	3.539	4.156	7.750	22.338
NXP-G3	3.168	-	6.871	11.013	28.921

Table C.5: Extracted collector saturation currents, J_{C0} at low-injection ($A/\mu m^2$).

	300K	200K	77K	50K	18K
IBM-G4	4.4e-14	1.6e-23	6.6e-52	6.2e-78	8.9e-121
IHP-G4	5.5e-14	9.1e-23	8.1e-53	1.4e-69	6.6e-86
ST-G4	1.9e-13	9.6e-22	3.1e-44	7.5e-64	5.3e-92
ST-X2	8.5e-14	2.1e-22	2.5e-58	1.3e-86	6.8e-111
ST-X1	9.8e-14	2.6e-22	2.2e-51	1.2e-65	6.9e-72
ST-X3	1.3e-13	3.4e-22	1.1e-48	5.3e-63	8.9e-72
JAZZ-G3	1.2e-14	8.9e-24	1.3e-54	5.1e-65	7.2e-69
NXP-G3	4.2e-12	4.1e-20	3.5e-49	3.4e-63	1.7e-98

Table C.6: Extracted collector ideality factors, n_c, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	1.025	1.029	1.309	1.349	2.416
IHP-G4	1.015	1.034	1.238	1.458	3.321
ST-G4	1.044	1.070	1.473	1.571	3.013
ST-X2	1.019	1.037	1.116	1.160	2.527
ST-X1	1.025	1.032	1.231	1.496	3.815
ST-X3	1.036	1.039	1.305	1.563	3.841
JAZZ-G3	1.032	1.045	1.261	1.647	4.340
NXP-G3	1.024	1.061	1.236	1.499	2.681

Table C.7: Extracted Collector leakage saturation current, J_{CP0} at low-injection ($A/\mu m^2$).

	300K	200K	77K	50K	18K
IBM-G4	-	-	-	1.3e-18	5.7-24
IHP-G4	-	-	-	-	-
ST-G4	-	-	-	-	-
ST-X2	-	-	-	-	-
ST-X1	-	-	3.7e-18	1.3e-17	8.4e-19
ST-X3	-	-	1.3e-19	1.4e-19	1.4e-20
JAZZ-G3	-	-	3.8-19	1.9e-19	7.0e-20
NXP-G3	-	-	-	-	-

Table C.8: Extracted collector leakage ideality factors, n_{cp}, at low-injection.

	300K	200K	77K	50K	18K
IBM-G4	-	-	-	8.303	15.897
IHP-G4	-	-	-	-	-
ST-G4	-	-	-	-	-
ST-X2	-	-	-	-	-
ST-X1	-	-	5.083	8.298	20.704
ST-X3	-	-	4.614	7.177	18.430
JAZZ-G3	-	-	4.881	7.385	19.935
NXP-G3	-	-	-	-	-

C.2 f_t and f_{max}

The peak values of f_t and f_{max} were extracted using the method presented in Chapter 6 and the results appear below in Tables C.9-C.12.

Table C.9: Extracted peak f_t

Technology	300K	200K	77K	50K	18K
IBM-G4	218	268	331	339	344
IHP-G4	246	297	356	349	348
ST-G4	277	331	415	411	416
ST-X2	239	268	327	340	359
ST-X1	254	287	339	343	355
ST-X3	248	297	359	368	373
JAZZ-G3	122	156	173	179	177

Table C.10: Extracted J_C for peak f_t

Technology	300K	200K	77K	50K	18K
IBM-G4	11.8	12.5	12.1	11.7	12.3
IHP-G4	9.7	10.0	10.1	8.6	5.3
ST-G4	14.6	15.2	13.1	13.2	12.7
ST-X2	10.7	10.7	9.3	9.7	9.2
ST-X1	11.2	12.4	11.4	10.8	11.5
ST-X3	11.5	12.6	12.1	11.7	12.1
JAZZ-G3	5.0	5.6	5.9	5.9	5.6

Table C.11: Extracted peak f_{max}

Technology	300K	200K	77K	50K	18K
IBM-G4	244	277	302	311	336
IHP-G4	192	281	289	315	271
ST-G4	261	310	311	340	359
ST-X2	231	278	296	298	363
ST-X1	211	255	233	227	282
ST-X3	210	248	238	238	283
JAZZ-G3	149	163	170	172	165

C.3 Small-Signal Model Component Values

In this section, the small-signal component values of the bias independent resistances are presented along with equations describing the bias dependence of the capacitors.

Table C.12: Extracted J_C for peak f_{max}

Technology	300K	200K	77K	50K	18K
IBM-G4	11.3	11.9	10.7	10.9	9.7
IHP-G4	10.6	10.4	10.9	9.3	5.3
ST-G4	11.8	13.1	16.4	13.9	13.6
ST-X2	8.2	10.6	11.1	8.3	10.1
ST-X1	9.8	9.5	9.9	8.2	7.0
ST-X3	9.0	9.2	9.6	10.9	7.6
JAZZ-G3	4.5	5.4	4.8	4.8	4.8

C.3.1 Capacitances

The base-collector and collector to substrate capacitances can be described in terms of depletion capacitances. Thus, the equation for C_{CB} and C_{CS} is given as

$$C_X = \frac{C_{X0}}{(1 + V_X/V_{X0})^{mx}}, \tag{C.3}$$

where X is either CB or CS depending upon which junction is being described. The coefficients required to evaluate equation (C.3) for the collector-base junction appear in Tables C.13-C.15 and the coefficients required for the collector-substrate junction appear in Tables C.16-C.18.

Table C.13: Extracted C_{CB0}, $fF/\mu\text{m}^2$

Technology	300K	200K	77K	50K	18K
IBM-G4	13.0	12.5	12.2	12.2	11.8
IHP-G4	14.7	14.1	13.4	13.3	13.4
ST-G4	19.7	18.9	18.5	18.4	18.3
ST-X2	14.1	13.4	13.1	13.0	12.1
ST-X1	12.4	11.9	11.6	11.5	11.5
ST-X3	12.9	12.3	12.0	12.0	11.9
JAZZ-G3	10.3	9.8	9.5	9.5	9.2

Table C.14: Extracted V_{CB0}

Technology	300K	200K	77K	50K	18K
IBM-G4	0.689	0.855	1.062	1.123	1.067
IHP-G4	0.710	0.763	0.746	0.745	0.779
ST-G4	0.709	0.775	0.821	0.830	0.836
ST-X2	0.644	0.713	0.775	0.770	0.787
ST-X1	0.661	0.729	0.772	0.779	0.779
ST-X3	0.669	0.737	0.790	0.799	0.789
JAZZ-G3	0.699	0.882	1.033	1.047	1.084

Table C.15: Extracted m_{cb}

Technology	300K	200K	77K	50K	18K
IBM-G4	0.1602	0.1640	0.1834	0.1908	0.1756
IHP-G4	0.2458	0.2324	0.2073	0.1976	0.1918
ST-G4	0.2252	0.2179	0.2123	0.2121	0.2064
ST-X2	0.2222	0.2121	0.2073	0.1986	0.1734
ST-X1	0.2193	0.2092	0.2014	0.1992	0.1960
ST-X3	0.2252	0.2159	0.2118	0.2107	0.2014
JAZZ-G3	0.2126	0.2295	0.2408	0.2395	0.2306

Table C.16: Extracted C_{CS0}, $fF/\mu\text{m}^2$

Technology	300K	200K	77K	50K	18K
IBM-G4	5.6	8.7	7.7	6.8	2.0
IHP-G4	4.8	4.7	3.6	3.2	1.1
ST-G4	10.3	11.7	8.5	8.6	5.8
ST-X2	5.3	5.0	4.4	3.9	2.1
ST-X1	4.8	4.6	4.1	3.9	3.7
ST-X3	4.7	4.5	4.5	4.1	3.5
JAZZ-G3	7.4	8.6	7.9	7.3	3.2

Table C.17: Extracted V_{CS0}

Technology	300K	200K	77K	50K	18K
IBM-G4	1.133	1.272	0.963	1.073	-
IHP-G4	0.987	0.962	0.440	0.490	-
ST-G4	0.413	0.378	0.527	0.481	-
ST-X2	0.466	0.507	0.667	0.649	-
ST-X1	0.461	0.501	0.671	0.698	0.742
ST-X3	0.485	0.506	0.619	0.608	0.701
JAZZ-G3	0.469	0.582	0.729	0.847	-

Table C.18: Extracted m_{cs}

Technology	300K	200K	77K	50K	18K
IBM-G4	0.049	0.567	0.400	0.379	-
IHP-G4	0.317	0.407	0.186	0.117	-
ST-G4	0.126	0.048	0.077	0.052	-
ST-X2	0.155	0.157	0.127	0.077	-
ST-X1	0.134	0.141	0.122	0.109	0.086
ST-X3	0.137	0.148	0.109	0.091	0.070
JAZZ-G3	0.080	0.167	0.149	0.126	-

Numerical fittings for the base–emitter capacitances were also extracted. The equation that was used for this fitting is

$$C_{BE} = C_{BE0} + \frac{C_{BE1}}{(1 - V_{BE}/V_{BE0})^{m_b e}}. \tag{C.4}$$

The coefficients required for the evaluation of C_{BE} are given in Tables C.19–C.22.

Table C.19: Extracted C_{BE0} [fF/μm^2]

Technology	300K	200K	77K	50K	18K
IBM-G4	27.5	29.2	36.5	35.6	40.3
IHP-G4	17.1	15.6	11.9	16.3	18.9
ST-G4	19.2	19.5	24.0	23.6	25.9
ST-X2	18.9	17.9	19.4	19.4	16.9
ST-X1	25.2	20.7	21.0	27.9	29.7
ST-X3	18.9	23.0	23.6	26.6	28.2
JAZZ-G3	19.2	18.6	18.6	33.2	32.5

Table C.20: Extracted C_{BE1} [fF/μm^2]

Technology	300K	200K	77K	50K	18K
IBM-G4	2.0×10^{-16}	8.8×10^{-17}	1.1×10^{-18}	1.4×10^{-18}	2.2×10^{-19}
IHP-G4	9.4×10^{-17}	5.9×10^{-16}	1.4×10^{-15}	2.7×10^{-16}	1.6×10^{-19}
ST-G4	8.2×10^{-17}	9.9×10^{-17}	8.3×10^{-19}	1.2×10^{-18}	2.5×10^{-19}
ST-X2	6.5×10^{-18}	2.1×10^{-17}	3.5×10^{-19}	4.7×10^{-20}	2.9×10^{-20}
ST-X1	1.1×10^{-17}	2.2×10^{-16}	2.9×10^{-16}	5.6×10^{-18}	6.9×10^{-19}
ST-X3	4.9×10^{-16}	3.1×10^{-17}	1.5×10^{-17}	1.2×10^{-20}	5.6×10^{-21}
JAZZ-G3	1.4×10^{-15}	1.6×10^{-15}	1.5×10^{-15}	2.3×10^{-18}	3.6×10^{-19}

Table C.21: Extracted V_{BE0}

Technology	300K	200K	77K	50K	18K
IBM-G4	1.000	1.048	1.108	1.085	1.097
IHP-G4	0.957	0.973	1.030	1.042	1.095
ST-G4	0.939	9.998	1.076	1.074	1.079
ST-X2	1.056	1.051	1.121	1.136	1.148
ST-X1	1.044	1.014	1.021	1.049	1.045
ST-X3	0.932	1.027	1.045	1.105	1.097
JAZZ-G3	0.905	1.000	1.069	1.122	1.185

Table C.22: Extracted m_{be}

Technology	300K	200K	77K	50K	18K
IBM-G4	2.721	2.618	4.406	2.957	4.071
IHP-G4	2.625	1.239	0.862	1.341	4.533
ST-G4	2.866	2.284	4.000	3.608	4.096
ST-X2	5.216	3.240	4.877	5.877	6.299
ST-X1	4.773	2.669	1.701	3.363	4.105
ST-X3	1.807	3.127	2.956	6.886	6.963
JAZZ-G3	1.124	1.127	1.103	3.877	5.631

C.3.2 Resistances

The bias independent resistances are given below.

Table C.23: Emitter resistance, extracted using Gummel method, $\Omega \cdot \mu\text{m}^2$

Technology	300K	200K	77K	50K	18K
IBM-G4	1.10	1.38	1.41	1.53	1.71
IHP-G4	1.47	2.40	2.78	2.92	3.31
ST-G4	1.61	1.45	1.10	1.15	1.20
ST-X2	2.41	2.49	3.27	3.38	3.41
ST-X1	2.90	3.21	2.73	2.77	3.38
ST-X3	2.74	2.96	2.83	2.97	3.15
JAZZ-G3	3.78	3.70	1.88	1.95	2.46

Table C.24: Extracted Collector Resistance, $\Omega \cdot \mu\text{m}^2$

Technology	300K	200K	77K	50K	18K
IBM-G4	5.3	4.4	3.0	2.8	2.3
IHP-G4	12.1	10.0	7.8	7.0	6.0
ST-G4	4.8	4.5	3.9	3.7	3.3
ST-X2	1.8	3.7	1.0	0	0
ST-X1	6.6	4.8	4.3	4.0	3.1
ST-X3	6.3	4.6	3.3	2.9	2.8
JAZZ-G3	12.5	11.4	14.8	0	0

Appendix D

Network Parameters for HBT equivalent-circuit models

The equivalent circuit diagram for a SiGe HBT appears in Fig. D.1. In this section, the network parameters associated with each sub-block used for parameter extraction will be presented.

D.1 Intrinsic HF Circuit

The network parameters of the intrinsic circuit are given as

$$\mathbf{Y^I} = \begin{bmatrix} g_{be} + j\omega\left(C_{BE} + C_{CB}\right) & -j\omega C_{CB} \\ \widetilde{g_m} - j\omega C_{CB} & j\omega C_{CB} \end{bmatrix} \tag{D.1}$$

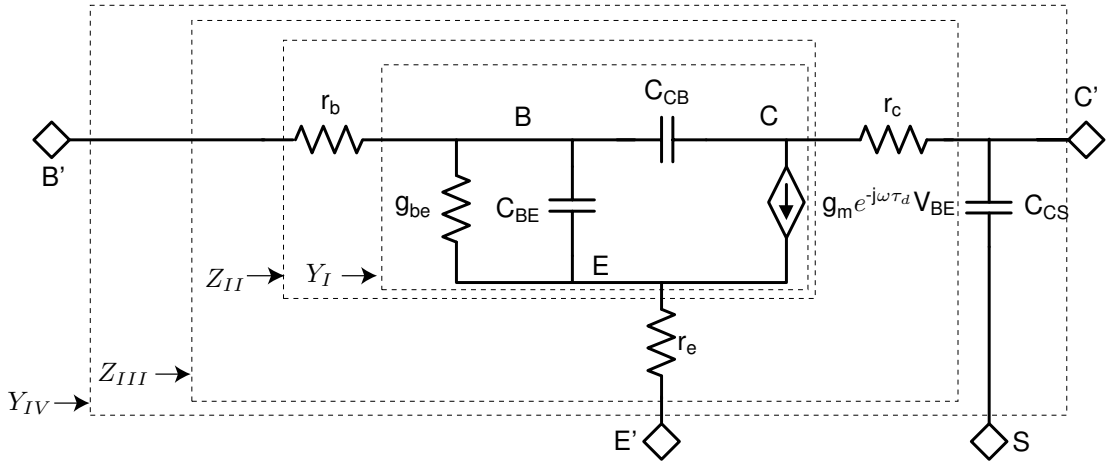


Figure D.1: SiGe HBT small-signal model with sub-network blocks identified

and

$$\mathbf{Z^I} = \frac{1}{\widetilde{g_m} + g_{be} + j\omega C_{CB}} \begin{bmatrix} 1 & 1 \\ 1 - \frac{\widetilde{g_m}}{j\omega C_{CB}} & 1 + \frac{C_{BE}}{C_{CB}} + \frac{g_{be}}{j\omega C_{CB}} \end{bmatrix}, \tag{D.2}$$

where

$$\widetilde{g_m} = g_m e^{-j\omega\tau_d}. \tag{D.3}$$

D.2 HF Circuit Including Base Resistance

The network parameters of the network block labeled "Z_{II}" are given as

$$\mathbf{Z^{II}} = \frac{1}{\widetilde{g_m} + g_{be} + j\omega C_{CB}} \begin{bmatrix} 1 + r_b\left(\widetilde{g_m} + g_{be} + j\omega C_{BE}\right) & 1 \\ 1 - \frac{\widetilde{g_m}}{j\omega C_{CB}} & 1 + \frac{C_{BE}}{C_{CB}} + \frac{g_{be}}{j\omega C_{CB}} \end{bmatrix} \tag{D.4}$$

and

$$\mathbf{Y^{II}} = \frac{\begin{bmatrix} g_{be} + j\omega\left(C_{BE} + C_{CB}\right) & -j\omega C_{CB} \\ \widetilde{g_m} - j\omega C_{CB} & \jmath\omega C_{CB}\left(1 + r_b\left(\widetilde{g_m} + g_{be} + j\omega C_{BE}\right)\right) \end{bmatrix}}{1 + r_b\left(g_{be} + j\omega\left(C_{CB} + C_{BE}\right)\right)}. \tag{D.5}$$

Appendix E

Noise in Bipolar Circuits

E.1 Noise in SiGe HBTs

The spectral densities of the equivalent input-referred noise generators for a SiGe HBT can be written in chain representation as [30]

$$S_{v_n} \approx \frac{\overline{|i_{nc}|^2}}{|Y_{21}|^2} + 4kT_a\left(r_b + r_e\right), \tag{E.1}$$

$$S_{i_n} \approx \overline{|i_{nc}|^2}\frac{|Y_{11}|^2}{|Y_{21}|^2} + \overline{|i_{nb}|^2}, \tag{E.2}$$

and

$$S_{i_n v_n^*} \approx \overline{|i_{nc}|^2}\frac{Y_{11}}{|Y_{21}|^2}. \tag{E.3}$$

For the case in which $f << g_m/C_{mu}$ and r_e is small[1],

$$Y_{21} \approx g_m \tag{E.4}$$

$$Y_{11} \approx \frac{g_m}{\beta_{AC}} + j\omega\left(C_\pi + C_\mu\right) \tag{E.5}$$

With the spectral densities known explicitly, the noise parameters T_{min}, Y_{OPT}, and R_n can be computed as [144, 160],

$$R_n = \frac{S_{v_n}}{4kT_0}, \tag{E.6}$$

$$G_{OPT} = \frac{1}{S_{v_n}}\sqrt{S_{v_n}S_{i_n} - \Im\left\{S_{i_n v_n^*}\right\}^2}, \tag{E.7}$$

[1] The assumption being made here in terms of r_e being small is that g_m is not reduced much due to r_e. Not making this assumption is the equivalent of replacing g_m with $G_m = \frac{g_m}{1+g_m r_e}$.

$$B_{OPT} = -\frac{\Im\left\{S_{i_n v_n^*}\right\}}{S_{v_n}}, \tag{E.8}$$

and

$$T_{MIN} = \frac{1}{2k}\left[\sqrt{S_{i_n}S_{v_n} - \Im\left\{S_{i_n v_n^*}\right\}^2} + \Re\left\{S_{i_n v_n^*}\right\}\right]. \tag{E.9}$$

Thus, the noise paramters of an HBT can be written as

$$R_n = \frac{T_a}{T_0}\left[\frac{n}{2g_m} + r_b + r_e\right], \tag{E.10}$$

$$G_{OPT} = \frac{g_m}{1 + 2g_m\left(r_b + r_e\right)/n_c}\sqrt{\left(\frac{1}{\beta_{DC}} + \frac{1}{\beta_{AC}^2}\right)\left(1 + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\right) + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2}, \tag{E.11}$$

$$B_{OPT} = -\frac{f}{f_t}\frac{g_m}{1 + 2g_m\left(r_b + r_e\right)/n_c}, \tag{E.12}$$

and

$$T_{MIN} = T_a n_c\left[\frac{1}{\beta_{AC}} + \sqrt{\left(\frac{1}{\beta_{DC}} + \frac{1}{\beta_{AC}^2}\right)\left(1 + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\right) + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2}\right] \tag{E.13}$$

For SiGe $HBTs$, $\beta_{AC}^2 \gg \beta_{DC}$ and equation E.13 reduces to:

$$T_{MIN} = T_a n_c\sqrt{\frac{1}{\beta_{DC}}\left(1 + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\right) + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2} \tag{E.14}$$

In general, when dealing with transistors, it is desirable to use invariants as noise parameters. Thus, the parameter $N = R_n G_{OPT}$ will be used instead of R_n. For a BJT with high β_{AC}, N can be written as

$$N = \frac{1}{2}\frac{T_a n_c}{T_0}\sqrt{\frac{1}{\beta_{DC}}\left(1 + 2\frac{g_b\left(r_b + r_e\right)}{n_c}\right) + 2\frac{g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2} \approx \frac{T_{min}}{2T_0} \tag{E.15}$$

Also, from equations (E.10)–(E.12), the noise conductance and optimum source as

$$g_n = \frac{T_a}{T_0}\left[\frac{1}{\beta_{DC}} + \frac{1}{\beta_{AC}^2} + \left(\frac{f}{f_t}\right)^2\right], \tag{E.16}$$

$$R_{OPT} = \frac{\beta_{DC}}{g_m\left(1 + \beta_{DC}\left(f/f_t\right)^2\right)}\sqrt{\frac{1}{\beta_{DC}}\left(1 + \frac{2g_m\left(r_b + r_e\right)}{n_c}\right) + \frac{2g_m\left(r_b + r_e\right)}{n_c}\left(\frac{f}{f_t}\right)^2}, \tag{E.17}$$

and

$$X_{OPT} = \frac{-B_{OPT}}{G_{OPT}^2 + B_{OPT}^2} = \frac{-\omega\left(C_{BE} + C_{CB}\right)}{\omega^2\left(C_{BE} + C_{CB}\right)^2 + g_m^2/\beta_{DC}}. \tag{E.18}$$

E.2 Amplifiers Employing Resistive Feedback

The analysis of a resistive feedback amplifier is equivalent to the analysis of the circuit shown in Fig. E.1(a), where the Y-parameter block is assumed noiseless and noise sources i_1, i_2, and i_f are uncorrelated. To work this problem, the effect of the feedback resistor will be approximated by ignoring its contribution to equivalent noise current at the output as shown in Fig. E.1(b). This approximation is valid so long as the closed-loop circuit has high current gain. The short-circuit output current noise can then be written as

$$\overline{|i_{on}|^2} = \overline{|i_1'|^2}\frac{|Z_S|^2\,|Y_{21}|^2}{|1 + Y_{11}Z_S|} + \overline{|i_2|^2} + 2\Re\left\{\overline{i_1 i_2^*}\frac{Y_{21}Z_S}{1 + Y_{11}Z_S}\right\}, \tag{E.19}$$

where $\overline{|i_1'|^2} = \overline{|i_1|^2} + \overline{|i_f|^2}$. Similarly, the short-circuit output current from a generator with impedance Z_S can be written as

$$\overline{|i_{og}|^2} = \overline{|i_g|^2}\frac{|Y_{21}|^2\,|Z_S|^2}{|1 + Y_{11}Z_S|^2}. \tag{E.20}$$

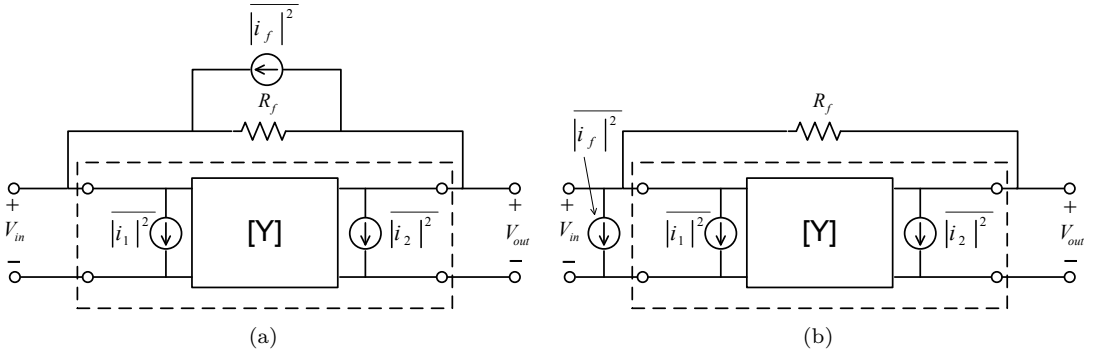


Figure E.1: (a). Generic two-port network with shunt resistive feedback applied. The two-port network is represented in terms of Y-parameters and the noise is represented by an equivalent input and output current source. The feedback network is located outside the dotted line. (b). Simplified equivalent circuit in which the current noise due to the feedback resistor has been moved to the input. This simplification involves ignoring a fully correlated current noise source at the output and is valid so long as the close loop circuit has high gain.

Defining an effective temperature of the source resistance, T_e, and using the thermal noise relation $\overline{|i_g|^2} = 4kT_e\Re\{Z_S\}/|Z_S|^2$ and equating $\overline{|i_{on}|^2}$ with $\overline{|i_{og}|^2}$, we can write the following relationship:

$$T_e = \frac{1}{4kR_S}\left[\overline{|i_1'|^2}|Z_S|^2 + \overline{|i_2|^2}\frac{|1+Y_{11}Z_S|^2}{|Y_{21}|^2} + 2\Re\left\{\frac{\overline{i_1 i_2^*}Z_S\left(1+Y_{11}^*Z_S\right)}{Y_{21}^*}\right\}\right]. \tag{E.21}$$

Now, making the approximation[2]. that $|Y_{11}Z_S| << 1$, which is valid for circuits in which high β_{DC} transistors and large feedback resistors are used, equation (E.21) can be simplified considerably:

$$T_e \approx \frac{1}{4kR_S}\left[\overline{|i_1'|^2}|Z_S|^2 + \frac{\overline{|i_2|^2}}{|Y_{21}|^2} + 2\Re\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}Z_S\right\}\right]. \tag{E.22}$$

It is our desire to minimize equation (E.22) with respect to the complex source impedance $Z_S = R_S + jX_S$. Taking derivatives of equation (E.22) with respect to R_S and X_S, we arrive at the following set of equations:

$$\frac{\partial T_e}{\partial X_S} = \frac{1}{4k}\left[\frac{2X_{OPT}}{R_{OPT}}\overline{|i_1'|^2} - \frac{2}{R_{OPT}}\Im\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}\right] = 0 \tag{E.23}$$

and

$$\frac{\partial T_e}{\partial R_S} = \overline{|i_1'|^2} - \frac{1}{R_{OPT}^2}\left[\overline{|i_1'|^2}X_{OPT}^2 + \frac{\overline{|i_2|^2}}{|Y_{21}|^2} - 2X_{OPT}\Im\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}\right] = 0. \tag{E.24}$$

Solving equations (E.23) and (E.24) simultaneously yields the desired real and imaginary components of the optimum source impedance:

$$R_{OPT} = \sqrt{\frac{\overline{|i_2|^2}}{\overline{|i_1'|^2}|Y_{21}|^2} - \Im\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}^2} \tag{E.25}$$

and

$$X_{OPT} = \frac{1}{\overline{|i_1'|^2}}\Im\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}. \tag{E.26}$$

Thus, plugging equations (E.25) and (E.26) into equation (E.22), we arrive at the expression for the minimum noise temperature in terms of the input and output referred noise powers:

$$T_{MIN} = \frac{1}{2k}\left[\sqrt{\frac{\overline{|i_1'|^2}\,\overline{|i_2|^2}}{|Y_{21}|^2} - \Im\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}^2} + \Re\left\{\frac{\overline{i_1 i_2^*}}{Y_{21}^*}\right\}\right] \tag{E.27}$$

[2]This approximation is valid so long as $f/f_t << \sqrt{1/\left(g_m^2|Z_S|^2\right) - \left(g_f/g_m\right)^2 - 2g_f/\left(g_m\beta_{AC}\right) - 1/\beta_{AC}^2} \approx 1/\left(g_m|Z_S|\right)$.

Finally, the last noise parameter, g_n can be obtained from the following equation:

$$T_e = T_{MIN} + \frac{g_n T_0}{R_S}\left[\left(R_S - R_{OPT}\right)^2 + \left(X_S - X_{OPT}\right)^2\right]. \tag{E.28}$$

Running through this exercise, the noise conductance is found to be:

$$g_n = \frac{\overline{\left|i_1'\right|^2} + \overline{\left|i_2\right|^2}}{4kT_0}. \tag{E.29}$$

At this point, a generic equation for each of the noise parameters in terms of the equivalent input and output noise current sources has been found. The only simplification used is to assume that the input admittance looking into the network when the output is terminated in a short circuit is much lower than the source admittance. This assumption is quite reasonable for the case of a resistive feedback amplifier[3], which we will discuss now. If such an amplifier uses a bipolar device, the input noise current power can be written as $\overline{\left|i_1'\right|^2} = 4kT_a g_f + 2qI_C\beta_{DC}$ and the output noise can be approximated[4] as $\overline{\left|i_2\right|^2} \approx 2qI_C$, and the transimpedance can be approximated[5] as $Y_{21} \approx g_m - g_f$. Thus, for a resistive bipolar feedback amplifier, the noise parameters can be approximated as:

$$T_{MIN} = T_a \frac{n_c}{g_m\sqrt{\beta_{DC}}\left(1 - g_f/g_m\right)^2}\sqrt{1 + 2\beta_{DC}\frac{g_f}{g_{m,ideal}}} \approx T_a \frac{n_c}{g_m\sqrt{\beta_{DC}}}\sqrt{1 + 2\beta_{DC}\frac{g_f}{g_{m,ideal}}}, \tag{E.30}$$

$$R_{OPT} = \frac{\sqrt{\beta_{DC}}}{g_m\left(1 - g_f/g_m\right)}\sqrt{\frac{1}{1 + 2\beta_{DC}g_f/\left(qI_C/kT_a\right)}} \approx \frac{\sqrt{\beta_{DC}}}{g_m}\sqrt{\frac{1}{1 + 2\beta_{DC}g_f/g_{m,ideal}}}, \tag{E.31}$$

$$X_{OPT} \approx 0, \tag{E.32}$$

and

$$g_n = \frac{T_a}{T_0}\frac{n_c}{2g_m\beta_{DC}}\left(1 + 2\beta_{DC}\frac{g_f}{g_{m,ideal}}\right), \tag{E.33}$$

where $g_{m,ideal} = n_c g_m = I_C/\left(kT_a/q\right)$ is the ideal transconductance.

Now, for a shunt feedback amplifier, it is well known that the input impedance and gain are easily related to the open loop voltage gain [152]:

$$A_{V,CL} \approx -\frac{A_{V,OL}}{1 + A_{V,OL}R_S/R_F} \tag{E.34}$$

[3]It is assumed that the amplifier is operating at frequencies well below f_t.

[4]We have neglected the output noise from the feedback resistor. This is valid provided the closed loop current gain is high.

[5]This assumes that $f << g_m/\left(2\pi C\mu\right)$, which is valid to frequencies nearing roughly $f_t/5$.

and

$$R_{IN,CL} \approx \frac{R_F}{1 + A_{V,OL}}, \tag{E.35}$$

where 'OL' stands for 'open-loop', 'CL' stands for 'closed-loop', and $A_{V,OL} \equiv -g_m R_L || R_F$. Thus, from a design perspective, it would be nice to be able to compute the closed loop noise parameters directly from the open-loop noise parameters as a function of the feedback resistance. To accomplish this task, we begin by defining the parameter θ:

$$\theta \equiv \sqrt{1 + 2\beta_{DC} \frac{g_f}{g_{m,ideal}}}. \tag{E.36}$$

Now, the noise parameters of the closed-loop amplifier can be expressed in terms of the open-loop noise parameters and the variable θ:

$$T_{MIN,CL} = \theta\, T_{MIN,OL}, \tag{E.37}$$

$$R_{OPT,CL} = \frac{1}{\theta} R_{OPT,OL}, \tag{E.38}$$

$$X_{OPT,CL} = X_{OPT,OL} = 0, \tag{E.39}$$

and

$$g_{n,CL} = \theta^2 g_{n,OL}. \tag{E.40}$$