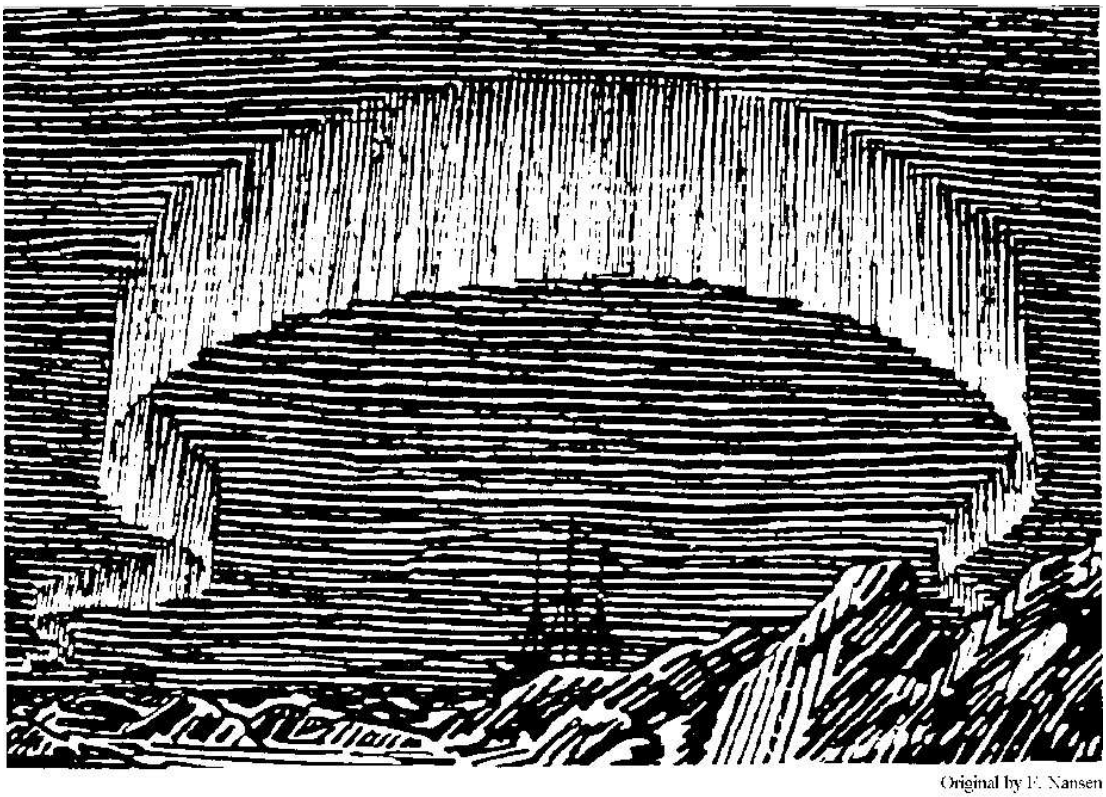
A Dissertation for the Degree of Master of Science

Low-frequency Noise Power Spectrum Density Characterizing of SiGe HBTs

Peng Qi

May 2006



Original by F. Nansen

FACULTY OF SCIENCE

Department of Physics

University of Tromsø, NO-9037 Tromsø, Norway

To my parents

Abstract

The main purposes of this thesis are the Low-Frequency Noise measurement of Silicon-Germanium Heterojunction Bipolar Transistors and its Power Spectrum Density Characterizing.

The new generation $375\ GH_Z$ SiGe HBTs were measured in this work. We show that most of PSDs of the new generation SiGe HBTs have very "bumpy" spectra which is contributed by GR noise sources. We investigated their basic characteristics of LFN such as the dominant noise source, base-current dependence, emitter geometrical scaling dependence and noise variation. They have similar LFN characteristic with the elder generation SiGe HBTs except for the emitter geometrical dependence.

The most important contribution of this work is that we particularly focused on developing a totally automatic mechanism to fit the Low-Frequency Noise Power Spectrum Density of SiGe HBTs so that we can use the magnitude of the fitting curve as the low-frequency noise level at any frequency. A model based predictive and autonomous method was engaged for this purpose. This method offers the possibility that we can automatically predict the noise sources of transistors to get good initial fitting parameters in advance instead of finding each of them by eyes. Experiments with the fitting method shows that:

- Always good fitting for most of the cases;
- Accurately locating each noise source;
- Sometimes meaningless fitting parameters but still good fitting.

Therefore, by using this method, we can find out how each noise source acts on the spectrum, which noise source dominates the spectrum, etc. And some careful interpretations will be presented based on this fitting procedure. Further, this method still leaves large space to be extended, so it is a good basis for future work on fitting.

Acknowledgements

I would like to express my most honest appreciation to all the people who ever gave me help during the last one year. Particularly, I would like to thank my supervisor Jarle. A. Johansen for his continual instruction and encouragement from cover to cover. I have greatly benefited from his profound knowledge and charming personality. I am also grateful the initial suggestions and guidance of my internal supervisor Alfred Hanssen.

I would like to thank my class fellow Anthony Doulgeris for his kind help and concern from the beginning of my study in Tromsø and wise advices for this work. And I am really appreciate the spiritual support from my good Chinese fellow Ciren Yang-Zong, Li Chun, Zhang Qian and my soul mate Dong Yue.

Finally, I would like to thank my father Qi XiangHao and my mother Mao QiYu. Without their continuous support, great concern and thriving expectation, I can not have this work done.

Contents

Chapter 1

Introduction

Since the transistor was born in December 1947, invented by John Bardeen, Walter Brattain and William Shockley in Bell Labs, efforts on searching new semiconductor materials and studying the properties of them have never stopped.

Noise, as one of the essential properties, exists in all kinds of semiconductor materials and acts an important role in the performance of semiconductors, not only for an individual diode or transistor, but also for the whole circuit, and represents an unexpected random interference. This phenomenon limits the minimum signal level that any devices can usefully work on, since there will always be a small but significant amount of noise arising. In another hand, noise also carry some useful information which can help scientists and engineers to characterize the properties of the devices and further take the advantage of it, such as probing the defect density, the material purity and reliability, the condition of circuits, etc. Therefore, studying noise will redound to identify the characteristics of materials and improve the performance of circuits.

The characteristics of Low-Frequency Noise(LFN) which dominates the low-frequency power spectrum in semiconductor transistors will be systematically discussed in this thesis in terms of the analysis of their Power Spectra Density(PSD). Special efforts will be focused on an automatic PSD fitting procedure. Silicon-Germanium Heterojunction Bipolar Transistors(SiGe HBTs) are the candidates to be tested in the experiment since SiGe HBTs are widely used for system-on-chips(SOC) and system-on-package with high frequency and high speed applications in modern industry due to their high-level integration, high speed, low cost, good matching and low noise.

In the theoretical part, we will introduce some element semiconductor concepts to help us understand the origin of the various kinds of noise sources which contribute the LFN PSD of semiconductor devices, diodes and bipolar junction transistors. The SiGe

HBT technology will be preselected. And we will briefly review semiconductor noise sources and state a model for LFN in SiGe HBTs.

Some modern statistical tools such as Multitaper(MT) will be presented to estimate Power Spectrum Density(PSD) from time series in this work as a comparison with the result of the Dynamic Signal Analyzer HP 3561A.

In the experimental part, we will first present the DC and Low-frequency noise measurement system we used in laboratory.

The LFN measurement system employs a computer control mechanism so that we can accurately and automatically set various measurement parameters we expect, such as different base-current. Time series measurements can also be achieved by this system.

The PSDs we measured by using this system are then systematically analyzed to investigate various characteristics of the LFN of SiGe HBTs, such as the dominant noise source, base-current dependence, emitter geometrical scaling dependence and noise variation. The Multitaper PSD estimation will also be done as a comparison of the PSD which is estimated by the dynamic signal analyzer.

Great efforts will be concentrated to an automatic fitting procedure. This procedure efficiently fits many kinds of LFN PSD by a model based method. By this method, we can dramatically identify each noise component of the LFN and see which one is the dominant noise source.

The core of this fitting procedure is non-linear fitting method, and a noise model will be introduced as the fitting function. The most important point in this procedure is that we innovatively choose the initial fitting parameters for this non-linear fitting function by using the classification, that is, we automatically get the initial fitting parameters from a noise model database instead of estimating them by eyes every time. The noise model database actually contains a sequence of synthetic PSDs which can be established by either theoretical noise level or the actual noise level of the test PSDs themselves. By comparing the test PSD with the PSD models in the database, we can choose the parameters of the closest model PSD as the initial fitting parameters. The final parameters will schematically show all of the noise components of LFN. This method was initial developed in order to find the PSD magnitude at any frequency when we analyze the PSDs. Latterly, it was proven as not only a efficient fitting method, but also a good point to study the characteristics of LFN.

Part I

Theoretical Part

Chapter 2

Elementary Semiconductor Concepts

A semiconductor is a material with an electrical conductivity intermediate between that of an insulator and a conductor. The electrical properties of semiconductors are drastically influenced by the material, purity and structure. These characteristics have been well investigated and employed in many different applications. Commonly used semiconductor materials are silicon(Si), germanium(Ge), and some compound semiconductors such as gallium-arsenide (GaAs) and cadmium-telluride (CaTe). Silicon, due a large part to the advanced state of its fabrication technology, is the most important semiconductor, and completely dominate the current commercial market. In this chapter, some basic properties of semiconductors will be introduced in order to help us rather understand the noise sources later on.

2.1 The Semiconductor Lattices

The spatial arrangement of atoms inside a material acts an important role in determining the properties of the material. In a broad sense, the structure of semiconductor lattices can be classified as amorphous, polycrystalline and crystalline. An amorphous Si thin-film transistor is used as the switching element in liquid crystal displays; polycrystalline Si gates are engaged in Metal-Oxide-Semiconductor Field-Effect Transistors(MOSFET's). However, the crystalline semiconductor has been employed to fabricate devices in the vast majority of modern industry. The main goal here is illustrating the crystal structure of the principal semiconductor(crystalline) and how it works as a conductor. Most of the discussions and examples in this section will be based on Si, which is applicable to other semiconductors with different materials, such as Ge and the compounds.

2.1.1 The Unit Cell

The unit cell is the most basic element of a given crystal. A two-dimensional lattice shown in figure 2.1(a). People can choose either of the two alternative unit cells shown in figure 2.1(b) for reproducing the crystal. Theoretically, there could be countless unit cells for a given crystal with different size or angle, but it is favorable to employ a relatively lager unit cell with orthogonal sides instead of a primitive cell(the possibly smallest unit cell) with nonorthogonal sides. Especially in three-dimensional case, a non-cubic unit cell will make the problem more difficult.

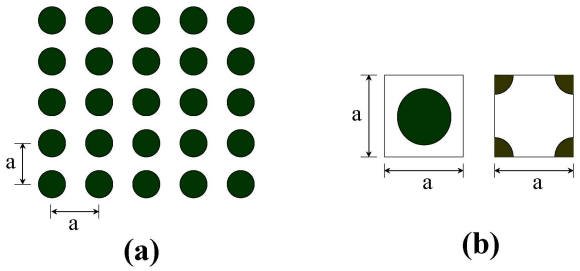


Figure 2.1: The unit cell method of describing atomic arrangements inside the crystals. (a) A two-dimensional schematic lattice. (b) Two alternative unit cells.

2.1.2 The Simple 3D Unit Cells

Semiconductor crystals take three-dimensional forms in nature. Like the two-dimensional case, there are many alternative 3D unit cells. The simplest 3D unit cell is shown in figure 2.2(a) which is called *Simple cubic unit cell* [Pierret, 1988]. Each of the eight atoms in this simple cubic contribute its $1/8$ fractional corner to the unit cell, and its other portions are shared with adjacent unit cell. Thereby, there is altogether one atom in a Simple cubic unit cell. Figure 2.2(b) and (c) show somewhat more complicate unit cells, *bcc* and *fcc*, which contains two and four atoms respectively. The lattice of Gray $S_n(\alpha - S_n)$ and elemental semiconductors like Si and Ge is well known as the *diamond lattice* unit cell shown in figure 2.2(d)[Kasap, 1997] since it shows the same structure with diamond. The diamond cubic unit cell has eight atoms.

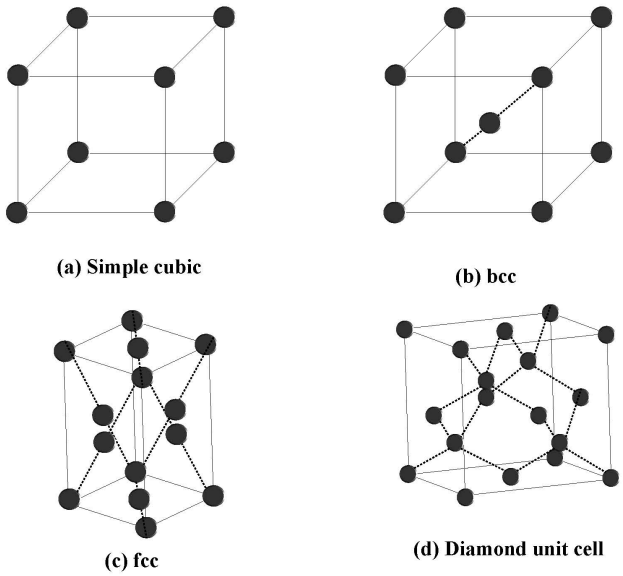


Figure 2.2: Examples of three-dimensional unit cells. (a) Simple cubic unit cell. (b) Body centered cubic unit cell. (c) Face centered cubic unit cell. (d) Diamond lattice unit cell.

2.2 The Carriers

Charge carriers are the entities inside any type of conductor that are responsible for the transport of the charge within the material. For normal conductors such as metal, the electron is the most commonly encountered carrier. In semiconductor, however, there is another kind of carrier called a *hole*, which is effectively an empty electronic state in the valence band and acts as if it is a positively charged "particle". A hole also possess the equal status with electron as the entity of the carrier.

Figure 2.3(a) presents the Bonding Model of the semiconductor, where each circle represents one core of semiconductor such as Si and Ge, and each line represents a shared valence electron. An electron becomes a carrier when it absorbs sufficient energy by which it can break away from the valence bond and wander about inside the lattice. Consequently, a hole carrier will be simultaneously generated with the same energy but opposite charge as the electron. The figure 2.3(b) shows the corresponding energy band to illustrate the energy demand when a electron or a hole becomes a carrier, and it will be further discussed in section 2.3.

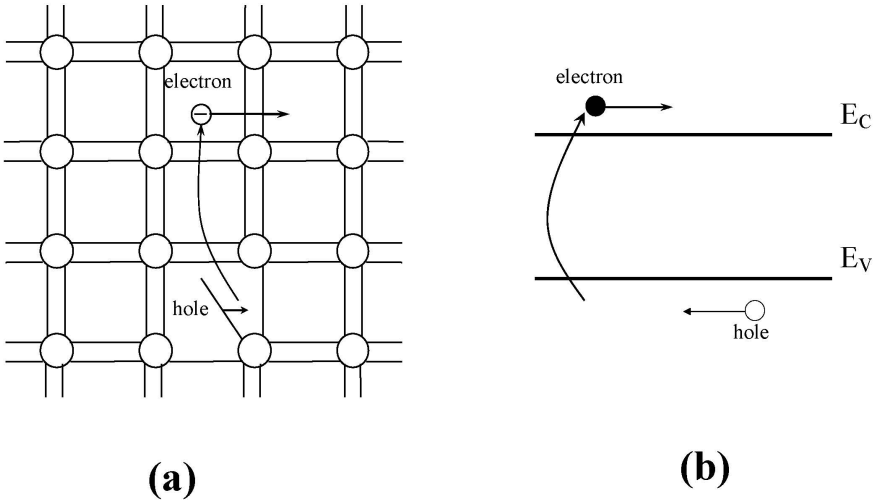


Figure 2.3: Illustration of carriers by means of bonding model. (a) Breaking a valence bond and freeing of an electron, generating a hole simultaneously. (b) The energy band diagram.

2.3 The Energy Band Model

A semiconductor behaves as an insulator at very low temperature, and has an appreciable electrical conductivity at room temperature although much lower conductivity than a conductor. One model advisably interprets how the semiconductors act these behaviours in terms of *energy band*. The conduction band electrons in crystalline semiconductor are not tied to any particular atom(semiconductor core). Instead, they are shared by all the atoms and the status of them changes as a function of time. Therefore, when we talk about the allowed electronic states of the semiconductor, the states are no longer the atomic states, but are associated with the crystal as a whole. For a perfect crystal under equilibrium conditions, a plot of the allowed electron energies versus distance along any preselected crystalline direction(always called the x-direction) is shown in figure 2.4, where the upper band of these allowed states is called *conduction band*(CB); the lower band is *valence band*(VB) and the intervening gap is known as the *forbidden gap* or *band gap*. The introduced E_C is the lowest possible conduction band energy, E_V is the highest possible valence band energy, and the intervening gap $E_G = E_C - E_V$ is the

band gap energy. The vacuum level plotted by the dashed line is the energy level where both the potential energy(PE) and the kinetic energy(KE) of electrons are zero, which means the electrons on this level is just free from the solid. The *Fermi Energy Level*(E_F) is defined as that energy value below which all states are full and above which all states are empty at absolute zero of temperature. It is actually the energy level where the probability of the occupied energy states is 0.5. Its value is up to the semiconductor material themselves, specially decided by the doping condition. The E_{Fi} plotted in this figure is the Fermi Level for intrinsic semiconductor which will be introduced later on. This cited plot, a plot of allowed electron energy presents as a function of position, is the **basic energy band model** [Pierret, 1988].

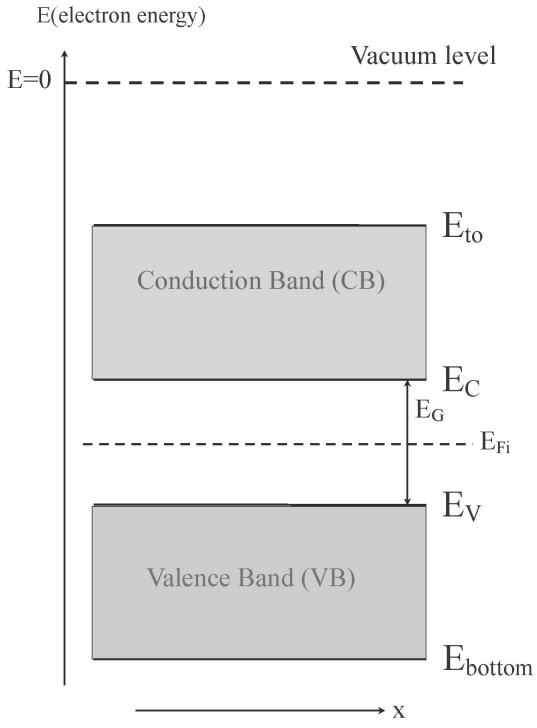


Figure 2.4: The energy band diagram of semiconductor.

The energy of the electron must be at least a little bigger than the E_C so that the electron could break away from the valence band to the conduction band to become a charge carrier. The electrons staying inside the valence band are constrained by the valence force, and can not be a carrier. At low temperature, few electrons can overcome

the valence force to become a free *carrier* since they do not have sufficient energy. On the other hand, higher temperature gives the electrons more thermal energy and can more likely break loose from the valence bond and consequently increases the conductivity. At absolute zero, the uppermost filled electron energy band of the semiconductor will be the fully filled VB, that is, no electron will be the *conduction band* and the semiconductor then becomes an isolator. Further more, base on the foregoing theory, it is not surprising that the electrical resistance of semiconductor changes by voltage, and does not follow the Ohm's law.

Recalling the energy band plot in figure 2.3(b) in section 2.2, the energy of the electrons become greater than the lowest conduction band energy E_C after breaking from the valence bond in valence band for the sake of being a carrier, and the energy of the hole should be correspondingly lower than the highest valence band energy E_V, that is, to be a carrier, the electron and hole should be in the conduction band and the valence band respectively. If the absorbed energy is exactly equal to the gap energy E_G, the electron and hole will stay at their respective critical band energy and have no ability to wander about.

2.4 Intrinsic Semiconductors

An intrinsic semiconductor is a pure semiconductor without any significant impurity species inside. The presence and type of charge carriers is therefore determined by the material itself instead of the impurities, and the amount of two kinds of carriers, electrons and holes, is roughly equal, that is, the number of electrons in the conduction band is equal to the number of holes in the valence band. Moreover, intrinsic semiconductors conductivity can be due to crystal defects or due to thermal excitation.

2.5 Extrinsic Semiconductors

Just as its name implies, an extrinsic semiconductor is an otherwise pure semiconductor which has been doped with impurities such that the concentration of carriers of one polarity exceeds the other type. The electronic properties of extrinsic semiconductors will therefore be severely influenced by the characteristics of the impurities. Two types of doping have been defined, n-type doping and p-type doping, due to the different type of carriers they introduce. The type and extent of the doping is decided by the requirement of the applications. For one particular purpose, the extrinsic semiconductor

could be doped with the corresponding doping type and impurity concentration level to achieve the requirement. For instance,a degenerate semiconductor, which is doped to such high levels that the dopant atoms are an appreciable fraction of the semiconductor atoms, acts more like a conductor than a semiconductor.

2.5.1 n-Type Doping

An n-type semiconductor is obtained by carrying out a process of doping more free negative charge carriers, as a result the concentration of the electron will be much higher than the hole, that is, the electrons are the majority carriers and the holes are the minority carriers.

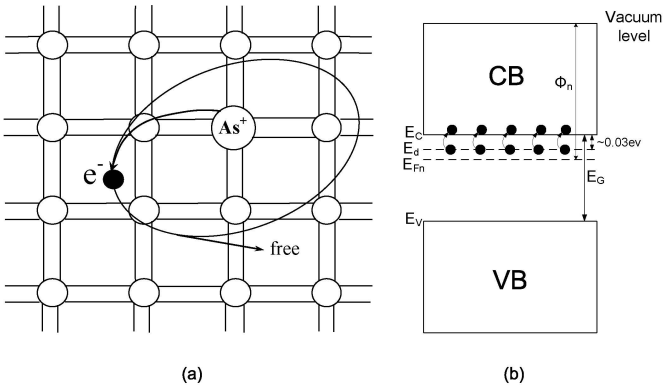


Figure 2.5: (a) Showing an As atom is doped into the pure silicon semiconductor in terms of the bonding model. (b) the corresponding energy band model.

The doping material donates the weakly-bound outer electrons to the semiconductor atoms. This type of doping agent is also known as donor material since it gives away some of its electrons. For example, doping with pentavalent atoms such as phosphorus(P) or arsenic(As) into the silicon semiconductor, as shown in figure 2.5(a)[Kasap, 1997], the As atom share four of its five outer electron with silicon atom just like Si, however the fifth is left orbiting the As site with the energy E_d which is close to the E_C and can therefore easily enter into the conduction band to be a carrier by a very small energy($E_C - E_d$) compared to the E_G. As the schematics of energy bands shows in figure 2.5(b), the fifth electron initially stay at the E_d energy level which is very close to the lowest conduction band level(E_C) so that the room temperature is sufficient to excite

all of them into the CB to be a carrier. The doping-dependent Fermi energy(or level) in n-type (E_{Fn})[1] is lifted towards the CB, which lead to higher electron concentration in the CB under the same temperature. Together, the n-type semiconductor has more carriers available under the same temperature than that of the intrinsic semiconductor. On the other hand, if the temperature is lower, the concentration of the electrons in the conduction band will be significantly decrease since the thermal energy is insufficient small and unable to excite all of the fifth electrons into the CB.

2.5.2 p-Type Doping

A p-type semiconductor is obtained by carrying out a process of doping, that is adding a certain type of atoms, such as the trivalent atoms, to the semiconductor in order to increase the number of free positive charge carriers(hole). Consequently, the concentration of the hole will be much higher than the electron. Therefore we get a opposite situation comparing to the n-type doping, that is, the holes are the majority carriers while the electrons are the minority carriers.

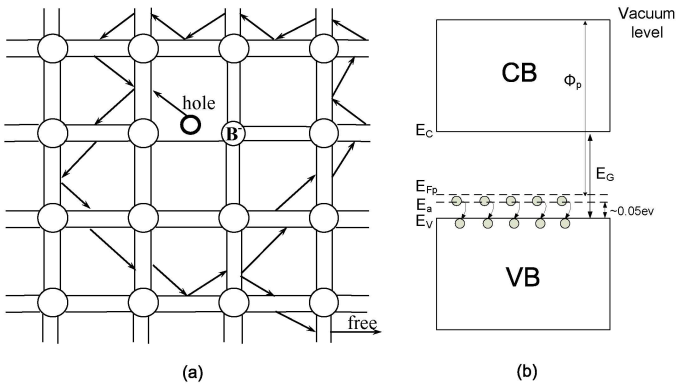


Figure 2.6: (a) Showing an Boron atom is doped into the pure silicon semiconductor in terms of bonding model. (b) The corresponding energy band.

Figure 2.6(a)[Kasap, 1997] shows that a doping material atom(Boron in this case) is added to the pure Si lattice, since Boron only has three valence electrons, when it

[1]For n-type semiconductor, the Fermi level will be higher than that of intrinsic semiconductor and close to the conduction band since more extra electrons have been introduced and more energy states will then be filled in at the same energy level.

substitute a Si atom and share the three electrons with four neighboring Si atoms, one of the covalent bonds has a missing electron, which then becomes a *hole*. The dopant atom(Boron) can accept an electron from a neighboring atoms' covalent bond(Such dopants then are called acceptors) by tunneling to complete the fourth bond, which displace the hole away from the B^-. Since the hole is still attracted by the negative charge left on the Boron(B^- ion), it thereby takes an orbit around the B^- ion. The binding energy between the hole and the B^- ion(E_a) is very small. Even though slightly greater than the binding energy between the electron and the As^+ in the n-type doping, it is still can be overcome by the thermal energy in room temperature. Therefore the *holes* can freely enter into the VB to be carriers under this condition. Figure 2.6(b) illustrates this procedure in terms of the energy band, where the *hole* energy level(E_a) is close to the valence band and the Fermi Lever is lowered towards the VB, this condition is just opposite to that of n-type semiconductor.

2.6 The Diode and Band Bending

The semiconductor diode is essentially a pn junction as shown in figure 2.7. The pn junction are actually made from the same intrinsic semiconductor crystal such as Si in practice by the fabrication process which creates the regions of different doping.

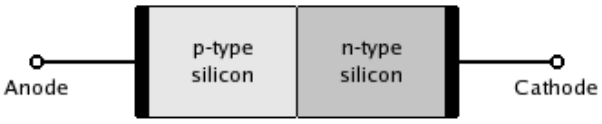


Figure 2.7: The pn junction.

2.6.1 Depletion Region

In figure 2.8(a), when the two types of regions stick together, the free electrons in n-type and the free holes in p-type will diffuse across the junction to the p-type region and n-type region respectively, and the diffusion between the two regions forms the diffusion current I_D [Adel S. Sedra, 1998]. The electrons that diffuse across the junction into the p region soon recombine with some of the majority holes in p-type region. A similar situation happens in the n-type region, with oppositely charged particles. Since the recombination happens near the junction, it gets depleted of free carriers and uncovers

bound charges, positive charges[2] in n-type side and negative charges[3] in p-type side. The bound charges consequently cause an intrinsic electric field. The region where the intrinsic electric field exists is called depletion region. This electric field also accelerate the minority carriers, electrons in p-type and holes in n-type which are thermally generated and diffuse to the edge of the depletion region, and form the minority drift current I_S within the depletion region, the minority carriers are not shown in this figure. The minority drift current is independent of the value of the depletion-layer voltage V_0 as shown in figure 2.8(b), but strongly dependent on temperature due to the thermally generated minority carriers.

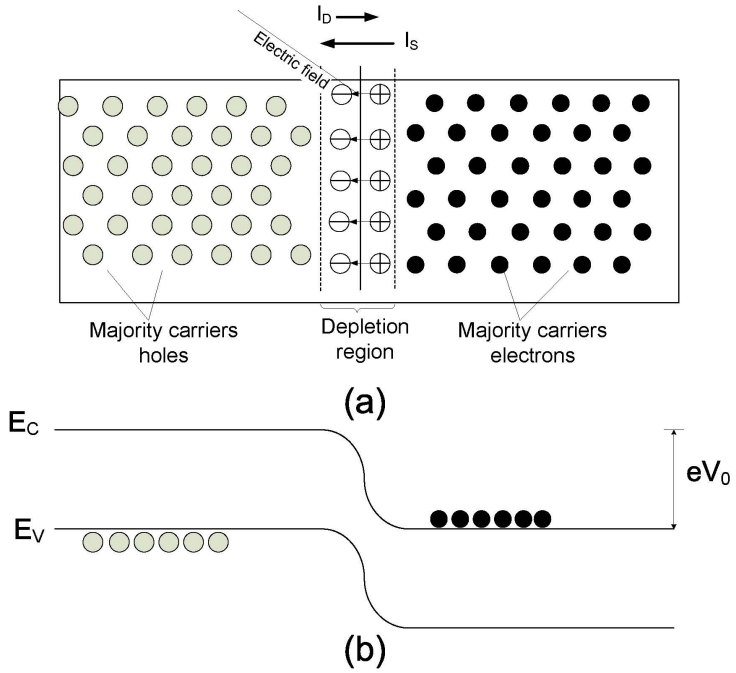


Figure 2.8: (a) The depletion region of pn junction. (b) The energy band.

From the analysis above, the depletion region is growing by the diffusion current I_D and forced by the minority carrier drift current I_S which is driven by the electric field of the depletion region itself. As long as these two part of current components is equal, i.e.

$$I_D = I_S \tag{2.1}$$

the pn junction then reach the equilibrium condition.

[2]The positive charge is the remaining part of pentavalent dopant atom that has lost its 5'th electron, hence it has a net positive charge.

[3]The trivalent dopant atom that has lost a hole, hence it has a net negtive charge.

2.6.2 Band Bending

The E_C and E_V in the energy band model shown in figure 2.4 have been consistently drawn as the energies independent of the position coordinate x. This happens only if the material is electric field free, such as the intrinsic semiconductor. Otherwise, the band energies become a function of position when a electric field exists inside the material such as the pn junction.

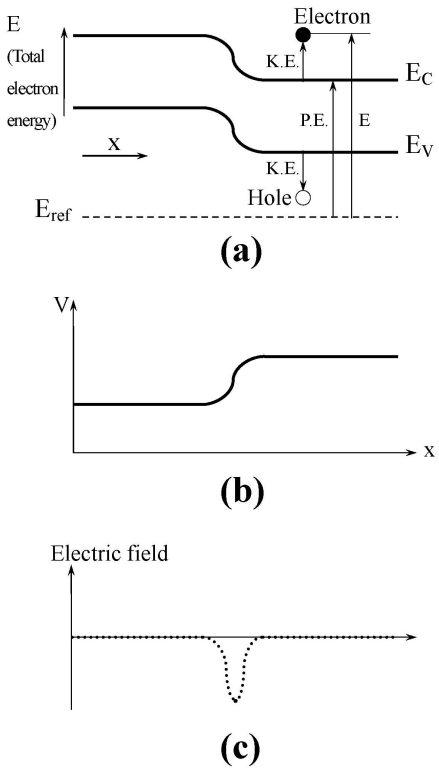


Figure 2.9: The energy band bending due to the electric field. (a) The band bending diagram with Kinetic energy($K.E.$), Potential energy($P.E.$) and total electron energy(E). (b) electrostatic potential. (c) the electric field corresponding the band bending in the part (a) the band bending diagram.

The band bending is exhibited in figure 2.8(b) and figure 2.9(a) by a simple energy band diagram of pn junction. It takes place in the depletion region where the electric field exists and the extent eV_0 is apparently up to the depletion-layer voltage V_0. For silicon, V_0 is between $0.6v$ and $0.8v$.

According to the foregoing energy band model theory, the electron should absorb an

energy in excess of E_G so that it can get the kinetic energy to move around as well as the hole which is simultaneously generated. Then the kinetic energy of electron as shown in figure 2.9(a) is defined by:

$$K.E. = E - E_C \tag{2.2}$$

The corresponding hole kinetic energy is given:

$$K.E. = E_V - E_h \tag{2.3}$$

where the E and E_h are the total energy of the electron and hole respectively compare to the reference energy E_{ref} which is position-independent and could be chosen to be any convenient value[Pierret, 1988]. The potential energy(P.E.), which is defined by $P.E. = E_C - E_{ref}$, is the key to explain the band bending which relates to the electric field. In elementary physics, assuming the existing electric field is the only force associated with the changing of the potential energy, then the potential energy of the electron with the charge $-q$ is given:

$$P.E. = -qV \tag{2.4}$$

Due to the inside electric field($\mathscr{E}$) located inside the depletion region shown in the figure 2.9(c), the electrostatic potential(V) at the corresponding position in figure 2.9(b) increase in terms of the definition $\mathscr{E} = -\frac{dV}{dx}$. The same result if we visually discuss the depletion region in figure 2.8(a) where the intrinsic electric field points to the p-type region and the electrostatic potential will grow along the opposite direction. Therefore, the correlative potential energy consequently decrease by the relationship equation (2.4) which consequently turns out the band bending.

2.7 npn Bipolar Junction Transistor

By definition, the bipolar junction transistor(BJT) is a kind of semiconductor device containing three adjoining, alternately doped regions in which the middle region is very narrow compared with the minority carrier diffusion length for that region. The three regions are,respectively, p type, n type and p type in a pnp transistor, and n type, p type and n type in a npn transistor. And each region is directly connected to a terminal labeled: *emitter(E), base(B)* and *collector(C)*. Figure 2.10(a), (b) show the schematic structure of the two types of BJT. By variation of the donor and acceptor concentration resulting from the fabrication process within the same crystal, different doping for these regions are realized. Figure 2.10(c) is a simplified cross section of the npn transistor

showing this kind of fabrication of the device.

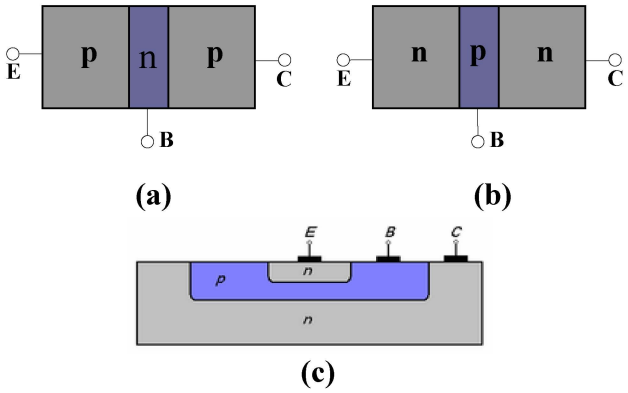


Figure 2.10: (a),(b) the sketch map of pnp and npn corresponding to the terminal name E(emitter), B(base) and C(collector); (c) The simplified cross section of npn bipolar transistor.

In this section, we will mainly discuss the characteristics of npn BJT since all the devices we used in the experiments are npn type.

As shown in figure 2.10(b) and (c), the npn BJT consists of one p-type(the acceptor is normally Boron) doped *base* between two n-type(the donor is normally Arsenic) doped regions *emitter* and *collector*. Furthermore, as will become obvious later, the emitter is generally much higher doped than that of base and the collector is even lower doped than the base.

2.7.1 The Active Operation Mode

The active operation mode is the mode that BJTs are generally functional in the circuits. Figure 2.11 shows the operation setup for npn bipolar transistors in active mode, where the BE junction is forward biased and the CB junction is reversely biased. To make the transistor conduct appreciable current from C to E, V_{BE} must be at least equal to or slightly greater than the cut-in voltage which is usually between $0.6v$ and $0.7v$ for silicon based BJTs. Under the active operation mode, we will subsequently discuss the working mechanism of the BJT.

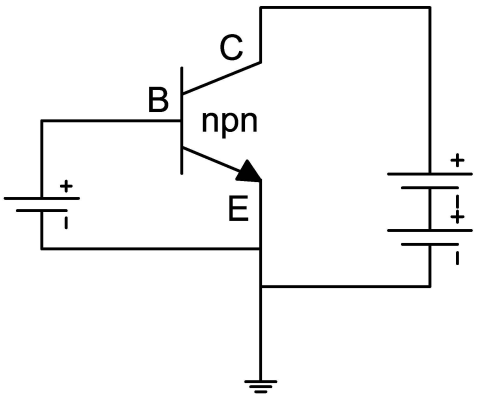


Figure 2.11: The operation setup for active mode.

2.7.2 The Current Components

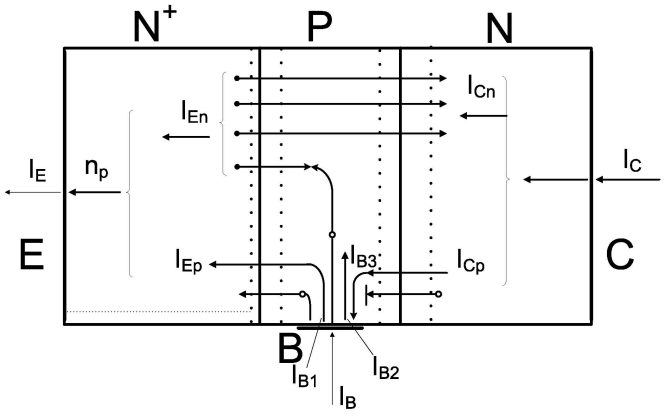


Figure 2.12: The carrier flux and current distribution within the BJT in the active mode.

The figure 2.12 shows the various carriers flux and current components of the npn BJT under the active mode except the recombination-generation currents in the depletion regions. Together we have the current value for each region:

$$I_E = I_{En} + I_{Ep} \tag{2.5}$$

$$I_C = I_{Cn} + I_{Cp} \tag{2.6}$$

$$I_B = I_{B1} + I_{B2} - I_{B3} \tag{2.7}$$

Both the electrons injected from emitter to base(I_{En}) and the holes injected from base to emitter(I_{Ep}) contribute the emitter current which is opposite to the direction of the electron flux and going out of the device. The I_{Ep}, the current from holes being injected across the forward biased BE junction from base into emitter, is very small compare to the I_{En} because of the lower doping and thin base. The I_{B1} in the base corresponds to the I_{Ep}, which is thereby also small. Since the very thin base width compare to the diffusion length of electron, most of the I_{En} will be diffuse through the base into the collector and becomes the I_{Cn}. But still few electrons are recombined with the majority carriers holes in base and form the I_{B2} which exists for complementing the disappeared holes in base and is very small in terms of the few quantity of the recombined electrons. The current I_{B3} is a part of the collector current due to the thermally generated holes in collector, and it is very small too. To sum up above, the BJT have similar magnitude I_E and I_C and both of which are stay at the much higher level than I_B.

2.7.3 The Gain of BJT

The transistor's current *gain*(β) is defined by:

$$\beta = \frac{I_C}{I_B} \tag{2.8}$$

The BJTs are normally expected that it should have high gain, that is, the $I_C \gg I_B$ i.e. $I_E \gg I_B$. There are two ways to achieve this goal, one of them try to increase the difference between the I_{En} and I_{Ep} by heavily doping the emitter and slightly doping the base and the I_{B1} is therefore reduced; another way most decreases the width of the base in order to decrease the number of electrons which is recombined by the holes in base and reduce the I_{B2}.

2.7.4 The Minority Carrier Diffusion in The Base

The minority carriers diffusion in the base is the key to accurately explain the relationship between the transistor's gain and the fabrication of the BJT. Note that there is not any other force to push the electrons through the quasi-neutralize base to the collector but the diffusion, therefore this diffusion significantly influences all the current component and further affects the performance of the transistor such as gain.

Figure 2.13 shows the minority carrier concentration distribution of the BJT in active mode, where P_{n0} and n_{p0} are the thermal equilibrium value of the minority carrier con-

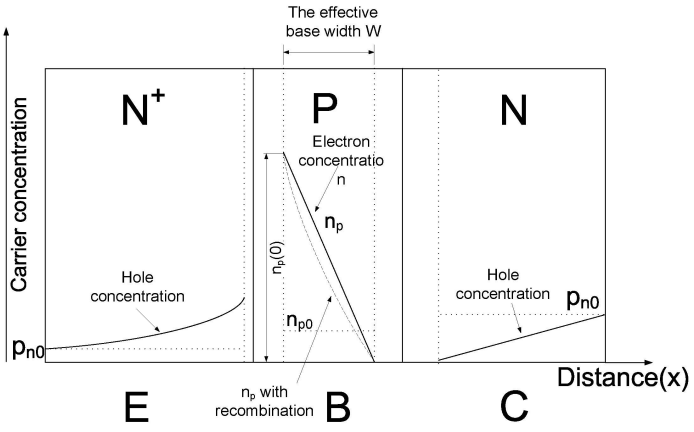


Figure 2.13: The minority carrier current in BJT.

centration in n-type region and p-type region respectively. The electrons injected from emitter to base becomes minority carrier in p-type base, and then across the base into the collector by means of diffusion. The concentration of these electrons is a straight line in the steady state if no recombination takes place. The electrons have the highest concentration($n_p(0)$) at the emitter side and lowest concentration(zero) at the collector side. For any forward biased pn junction, the concentration $n_p(0)$ will be proportional to e^{v_{BE}/V_T}:

$$n_p(0) = n_{p0} e^{v_{BE}/V_T} \tag{2.9}$$

where v_{BE} is the forward BE junction bias voltage and V_T is the thermal voltage. This decline minority carrier concentration profile in the base causes the injected electrons from emitter diffusing through the base toward the collector. And this electron diffusion current I_n in base is directly proportional to the slope of the straight line concentration profile[Adel S. Sedra, 1998], given by:

$$\begin{aligned} I_n &= A_E q D_n \frac{dn_p(x)}{dx} \\ &= A_E q D_n \left(-\frac{n_p(0)}{W} \right) \end{aligned} \tag{2.10}$$

where A_E is the cross section area of BE junction, q is the magnitude of the electron charge, D_n is the electron diffusivity in base and the W is the effective base width. We see that the I_n is reversely proportional to the base width(W), that is, the thinner base the

higher I_n. Since the BC junction is reversely biased, the diffusing electrons that reach the edge of the collector will be swept across the CB junction and totally contribute the collector current i_C, i.e. $I_n = i_C$. Taking the positive direction of i_C to be into the collector terminal, the i_C can be expressed:

$$i_C = I_S e^{v_{BE}/V_T} \tag{2.11}$$

the I_S is saturation current which is given by:

$$I_S = \frac{A_E q D_n n_i^2}{N_A W} \tag{2.12}$$

where the n_i is the intrinsic carrier density and N_A is the doping concentration of the base.

Two parts of base current I_{B1} and I_{B2} as shown in figure 2.11 have been concerned for calculating the i_B, since the I_{B3} which comes of the thermally generated holes is relatively very small. Then the i_B is:

$$i_B = I_S \left(\frac{D_p}{D_n} \frac{N_A}{N_D} \frac{W}{L_p} + \frac{1}{2} \frac{W^2}{D_n \tau_b} \right) e^{v_{BE}/V_T} \tag{2.13}$$

where the D_p is the hole diffusivity in the emitter, L_p is the hole diffusion length in the emitter, N_D is the doping concentration of the emitter and τ_b is the minority-carrier lifetime in the base, that is, the average time for a minority electron to recombine with a majority hole in the base.

Together, we have the current gain of BJT:

$$\beta = \frac{i_C}{i_B} = 1 / \left(\frac{D_p}{D_n} \frac{N_A}{N_D} \frac{W}{L_p} + \frac{1}{2} \frac{W^2}{D_n \tau_b} \right) \tag{2.14}$$

by which, we see that decreasing either the radio of N_A/N_D or the width W can increase the current gain. So a high gain BJT need the base doping concentration N_A is much smaller than the doping level of emitter N_D and the width of base should simultaneously be as thin as possible.

2.7.5 The Energy Band Bending

Figure 2.14[W.Neudeck and Pierret, 1989] shows the npn BJT energy band both for active region operation(the broken line) and for thermal equilibrium(the solid line). The

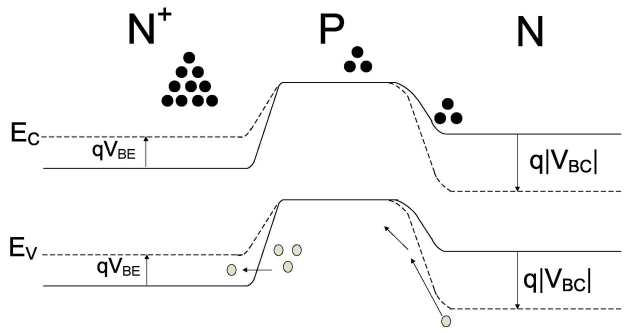


Figure 2.14: The energy band bending of npn BJT in active mode.

forward biased BE junction lowers the barrier for electrons entering into the p-type base from emitter and electrons therefore are allowed to be injected across the BE junction into the base and further diffuse through the very narrow base and slide down into the collector by the potential hill. The magnitude difference of the energy barrier, which has been lowered by the forward bias, is expressed by the potential energy qV_{BE} as shown in figure.

The stronger reversely biased BC junction in the active mode has greater energy difference($q|V_{BC}|$) with the thermal equilibrium compare to qV_{BE} which make the diffusion electrons from base more easily slide down into the collector by this deeper potential hill, at the meantime, it also increases the barrier for electrons in collector which may want to travel into the base.

Chapter 3

SiGe Heterojunction Bipolar Transistors

In modern industry, various applications require highly integrated and low cost ICs operating at very high frequencies, such as RF and microwave circuits. Therefore, it is necessary that the devices employed in these circuits not only should be suited for making highly integrated, low-cost ICs, but also can offer sufficiently high speed for the application at hand. In addition, more and more specific requirements emerge, which also need the highly integrated and low-cost semiconductor material to do the work. With all these expectations, the Silicon-Germanium Heterojunction Bipolar Transistors(SiGe HBTs) arise. Figure 3.1 shows the cross-section of a state-of-the-art $210GH_Z$ SiGe HBT. We can see that only the base of this transistor employs the SiGe alloy, the reason of this will be soon apparent.

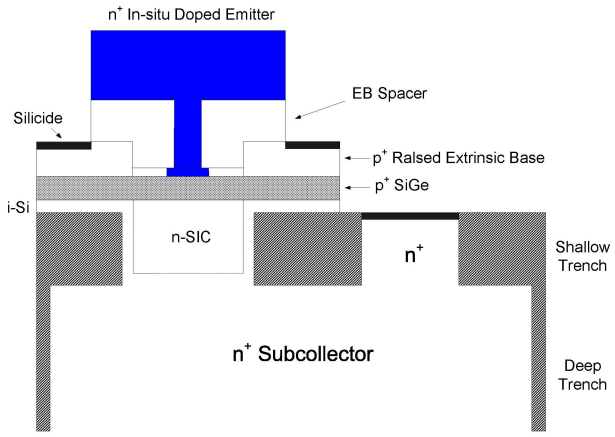


Figure 3.1: The schematic cross-section of the third generation SiGe HBTs.

3.1 The constraint of Si and III-V compounds

Silicon, as introduced in chapter 2, is the most wildly used semiconductor material and it totally dominates the current commercial market due to its advance state of fabrication, highly integration ability and abundant resource. However, its comparatively small carrier mobility for both electrons and holes excludes itself from the high frequency applications since the speed of a device ultimately depends on how fast the charges can be transported across the device under sustainable operating voltages. In fact, the maximum velocity the carriers in Si can obtain under a high electric field is limited to around $1 \times 10^7 cm/sec$ under normal conditions, and is regarded as a *'slow'* semiconductor [Cressler and Niu, 2002]. On the other hand, various III-V compounds(e.g.,GaAs and InP) enjoy much higher carrier mobilities and saturation velocity. Furthermore, the III-V compound devices can also handle many of specific applications by the well known *bandgap engineering* which dramatically offer many specific performances by altering their composition at atomic level [F.Capasso, 1987]. Nevertheless, the III-V devices lose the competition which is associated with making highly integrated and low-cost ICs.

3.2 The Advantages of SiGe HBTs

The emergence of the Silicon-Germanium alloy($Si_{1-x}Ge_x$) in recent decades, which is fabricated by introducing Ge into Si, gives us a completely new way to fabricate a semiconductor material with both the benefits of Si and III-V compounds, that is, the low-cost as Si and a comparable high speed as the III-V compounds. This idea is actually an old one by Kroemer [1957]. Due to the limitation of material growth technology, however, the first SiGe HBT was not born until 1987 [Iyer and et al., 1987]. Excitingly, with the emergence of the first generation SiGe HBTs, the rapid development of this technology was noticeable. Currently, the SiGe alloy has been commonly used in the IC manufacturing industry, where one important engagement of it is producing Silicon-Germanium Heterojunction Bipolar Transistors(SiGe HBTs). Some of the key points of the SiGe alloy and its heterojunction bipolar transistors include:

- The SiGe can be manufactured by the equipment which is used for the conventional silicon wafer. Therefore this process achieve costs that are similar with the silicon manufacturing compared to other far more expensive technologies such as III-V compounds;

- The SiGe allows state-of-the-art CMOS logic to be highly integrated with ultra high performance heterojunction bipolar transistors;
- The SiGe HBTs have significantly higher forward gain and lower reverse gain which leads to better low current and high frequency performance than that of traditional bipolar transistors;
- The bandgap engineering, which is normally available to the compound semiconductors, can be employed to tune the band gap for any particular purposes in terms of being a heterojunction technology.

3.3 The Advantage of SiGe HBTs vs Si BJTs

One important frequency response figure of merit of transistors is the unity-gain cut-off frequency(f_T), given by:

$$\frac{1}{2\pi f_T} = \tau_b + \tau_e + \tau_c + \frac{1}{g_m}(C_{re} + C_{rc}) \tag{3.1}$$

where the τ_b, τ_e and τ_c are the transit time of the base, emitter and collector respectively; g_m is the transconductance; C_{re} and C_{rc} are EB and CB junction depletion capacitances.

As the previous discussions in section 2.7, the bias conditions of EB and BC junction in the active mode make both the transit time of electrons in emitter and collect are significantly small compare to the transit time in base which is purely depend on the diffusion whose velocity is directly decided by the carrier mobility of the semiconductor material. That is, the diffusion time (τ_b in equation (3.1)) is affected by the carrier mobility. Thereby the comparatively small carrier mobility of conventional Si BJT makes large τ_b and subsequently decreases the speed of the transistor. Nevertheless, the SiGe HBTs, by means of bandgap engineering, can offer the opportunity to significantly speed up the velocity of the charge transport in the base and accordingly can produce far higher speed transistors.

The SiGe alloy is fabricated by introducing Ge into Si. Since Ge has a larger lattice constant and a smaller bandgap than Si($0.664\ eV$ vs $1.12\ eV$), the SiGe will have an adjustable bandgap between Ge and Si determined by the specific content of Ge and Si. Therefore a grading of the Ge content in the base of SiGe HBTs, along the x-direction as shown in the figure 3.2(b), will produce a corresponding grading of the bandgap, as shown in figure 3.2(a). The different grading of bandgap sequently establishes a built-in

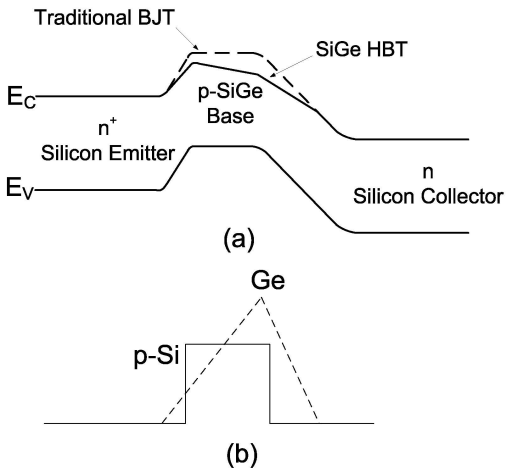


Figure 3.2: The energy band structure of SiGe HBT [Cressler and Niu, 2002].

electric field along the same direction in the base, which will force the electrons across the base far more rapidly than that of pure diffusion. Hence the SiGe HBTs can achieve a much higher speed of operation. Many different x values in $Si_{1-x}Ge_x$ have been used to investigate new characteristics of SiGe technology for achieving some particular requirements [Chen et al., 2001].

3.4 Conclusion

The SiGe HBTs have been more and more employed in virtually all analog and high-frequency applications recently, such as the cellular phone, wireless network and other communication systems in which the transistors with high speed, high-level integration, low-cost, low-noise and good matching are necessary. In addtion, by seamlessly integrating with the conventional Si CMOS, the SiGe HBT BiCMOS technology, which is perhaps the single most important advantage of SiGe HBTs, has over competed III-V HBT technologies by highly integrated ability and low-cost.

In this work, the fourth generation SiGe HBTs with $350\ GH_Z$ maximum unity-gain cutoff frequency(f_T), which was recently issued by IBM , will be the candidate of the measurements to help us investigate the performance of SiGe HBTs, on which the Low-frequency noise characteristics will be systematically investigated by means of the study of its Power Spectrum Density.

Chapter 4

Low-Frequency Noise(LFN) Sources of Semiconductor

The reason why the LFN is so important are given by the following two aspects: First, the magnitude of LFN could be significantly high close to dc, therefore it is of concern for low-noise analog circuits which need to operate at low frequency such as the amplifiers used in a zero intermediate frequency(IF); Second, the LFN can be upconverted to RF frequencies and produce transistor phase noise through the nonlinear $i - v$ relationship of the transistor. Due to the low noise capability, the SiGe HBTs has a significant advantage in many applications, such as the mobile receiver and typically its Low-Noise Amplifier(LNA) in which the amount of added noise must be sufficiently low.

The physical LFN origins will be introduced in this chapter. Afterwards, we will have a model expression for the LFN in terms of the noise sources which contribute all or part of the LFN along the frequency scale up to $10kH_Z$.

4.1 Thermal Noise

The thermal noise(or Nyquist noise) of a conductor is generated by the equilibrium fluctuations of the electric current regardless of any external power supply, due to the random thermal motion of the charge carriers. In general, the power spectral density(PSD) of the voltage across the conductor(R) is given by:

$$S_v(f) = \frac{2Rhf}{e^{\frac{hf}{k_B T}} - 1} \tag{4.1}$$

where f is the frequency, $h = 6.63 \times 10^{-34} Js$ is Planck's constant, $k_B = 1.38 \times 10^{-23} JK^{-1}$ is Boltzmann's constant and T is the absolute temperature of the conductor. In low frequency range, i.e.

$$f \ll \frac{k_B T}{h}$$

we have:

$$S_v = 4kTR \tag{4.2}$$

where the PSD of thermal noise has been presented as no frequency dependence, i.e. white noise. Thermal noise always there as long as the temperature does not become absolute zero and it is generally the white noise floor observed at high frequencies for the resistors.

4.2 Shot Noise

Shot noise comes up due to the fluctuations associate with dc current flow across a potential barrier. Solely the fact that the current is carried by discrete particles causes shot noise. This phenomenon was first observed in radio tubes by Schottky [1918].

The shot noise occurs in transistors when the current results from the discrete random emission of charged particles(electrons and holes) goes through the PN-junctions where the particles need overcome a potential barrier, therefore turn into a completely stochastic manner. The emission of these particles is assumed a Poisson stream. Base on the conventional macroscopic views, the Power Spectrum Density of the base and collector shot noises are:

$$S_I = 2qI \tag{4.3}$$

where I is the dc base or collector current, $q = 1.6 \times 10^{-19}$ is the electron charge.

In bipolar transistors, there is a very popular collector-base junction origin of the $2qI_C$ shot noise theory, that is, any dc current flow across any pn junction has shot noise[van der Ziel, 1955]. However, since the transition of the carriers go through the CB junction, which is normally reverse-biased for low-noise amplification, is a drift process, there is not any intrinsic shot noise when a dc current passing through such a junction alone. By this means, the collector current shot noise shows up only when the current being injected into the CB junction from the emitter has already had shot noise, that means the collector current shot noise results from the flow of emitter majority electrons over the potential barrier of EB junction, and has $S_{I_C} = 2qI_C$ [Niu, 2005]. Since the theory that shot noise is EB junction origin has been established, the transport noise

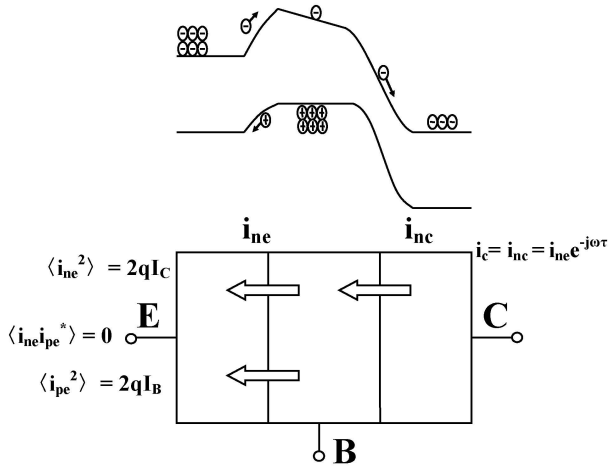


Figure 4.1: Emitter-base junction origin of collector current shot noise in a bipolar transistor.

model can be found as the illustration of figure 4.1. Both the electrons injection into the base from emitter and the holes injection into the emitter from base independently contribute the emitter current shot noise. And the PSD of them are:

$$S_{i_{ne}} = \frac{\langle i_{ne}^2 \rangle}{\Delta f} = 2qI_C \tag{4.4}$$

$$S_{i_{pe}} = \frac{\langle i_{pe}^2 \rangle}{\Delta f} = 2qI_B \tag{4.5}$$

which gives the same spectral expressions as the traditional view on shot-noise. In addition, the $\langle i_{ne} i_{pe}^* \rangle = 0$ due to the independent processes of electron and hole injections. The collector current shot noise S_{I_C} is the transported version of the electron injection into the base by EB junction. So we have:

$$i_c = i_{nc} = i_{ne} e^{-j\omega\tau} \tag{4.6}$$

showing that the collector noise is a delayed version of the emitter-base junction noise. The $i_c = i_{nc}$ noise is a phase delayed version of the i_{ne} noise by a factor depending on frequency(ω). For low frequencies with long wavelengths, $\omega \ll \frac{1}{\tau}$, the phase difference between the noise in EB and BC junction is negligible. At high frequencies, however, when the wavelength becomes comparable to device size, this phase difference is significant.

4.3 Generation-Recombination Noise

The generation-recombination noise(GR noise) is due to the trapping-detrapping processes of carriers among energy states, mostly between an energy band and a discrete energy level(trap) in the bandgap, which produce excess carriers through 'generation' and reduce the number of carriers by 'recombination'. This process turns out the fluctuations in the number of the carriers, which are given by:

$$\frac{d\Delta N}{dt} = -\frac{\Delta N}{\tau} \tag{4.7}$$

where ΔN is the carrier fluctuation and τ is the release time during the trapping-detrapping process. For a two terminal sample with resistance R and voltage V, the PSD are:

$$\frac{S_R}{R^2} = \frac{S_V}{V^2} = \frac{S_N}{\langle N\rangle^2} = \frac{\langle \Delta N^2\rangle}{\langle N\rangle^2}\frac{4\tau}{1+(2\pi f\tau)^2} \tag{4.8}$$

where the S_R, S_V and S_N are PSD of resistance, voltage and number of carriers respectively. $\langle N\rangle$ is the average number of free carriers. This kind of PSD expression gives a

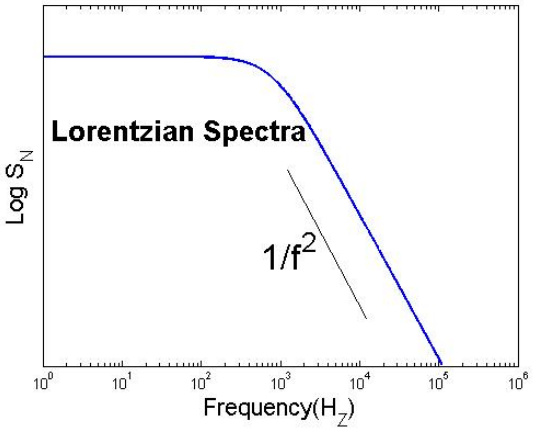


Figure 4.2: The Lorentzian.

Lorentzian noise spectrum as shown in figure 4.2 which is almost constant in the low-frequency range and rolls down as $1/f^2$ at high frequency [Jones, 1994].

4.4 $1/f$ Noise

The $1/f$ noise, also called flicker noise, is a signal with a frequency spectrum such that the power spectral density is proportional to the reciprocal of the frequency. People realized that the $1/f$ noise is a fundamental noise which is intrinsic to the semiconductor devices after it had been found in many semiconductor materials and devices [van der Ziel, 1979].

The common agreement about the origin of the $1/f$ noise is that it comes from the fluctuation of the conductivity(σ) which depends on both the mobility(μ) and number(N) of carriers. Their relationship is given by:

$$\sigma = q(\mu_n n + \mu_p p) \tag{4.9}$$

where μ_n and μ_p are the mobility of electrons and holes, and the n and p are the density of electron and hole respectively.

Hooge [1969] gave an empirical relation for $1/f$ noise base on homogenous samples of semiconductors and metals:

$$\frac{S_I}{I^2} = \frac{S_V}{V^2} = \frac{\alpha_H}{fN} \tag{4.10}$$

where α_H is well known as Hooge constant and initially given about 2×10^{-3}, and N is the carrier number. This relation also exclude the surface effect as the main source of the $1/f$ noise in homogenous samples since it is reversely proportional to the number of mobile carriers. However strong surface $1/f$ noises have been observed in n-type semiconductor [Vandamme, 1989] and BJT [Ziel, 1989] and present different α_H value.

The debate about whether the mobility fluctuation or number fluctuation is the fundamental $1/f$ noise mechanism has been lasted a long time and is still pendent.

4.4.1 Mobility Fluctuation Flicker Noise

This mobility fluctuation theory consider the origin of $1/f$ noise to be carrier scattering by lattice vibrations. The Hooge relation equation (4.10) has been wide-spread employed and connected to this theory, Hooge and Vandamme [1978] found that:

$$\alpha_{meas} = \alpha_{latt} \left(\frac{\mu_{meas}}{\mu_{latt}} \right)^2 \tag{4.11}$$

where α_{meas} is the measured Hooge constant and α_{latt} is Hooge constant when only lattice vibration exists in the test samples. The mobility subscript have the same meaning. By this relation, the Hooge constant in equation (4.10) can be derived from the mobility fluctuation. And it also proved that the lattice scattering is the only reason which cause $1/f$ noise. Later on, Hooge [1994] found that the α_H vary between 10^{-7} and 10^{-2} which indicates that the value of α_H is very sensitive to material quality and relative noise level of material and devices.

4.4.2 Number Fluctuation Flicker Noise

The number fluctuation is the fluctuation of the number of carriers in the conductor, which can be caused by the generation-recombination processes in the oxide-semiconductor surface such as the polysilicon to crystal silicon interfacial oxide and the oxide spacers around the emitter perimeter. If these independent GR-traps have a particular statistical distribution of characteristic time constant $g(\tau) \propto 1/\tau$ on a wide time scale, then the $1/f$ noise can be given by the superposition of these GR-traps [McWhorter, 1955] as shown in figure 4.3 which is similar with the model which was given by Surdin [1939] and Kingston and McWhorter [1956]:

$$S_N(f) = \int_0^\infty 4\langle \Delta N^2 \rangle \frac{\tau}{1+(2\pi f\tau)^2} g(\tau) d\tau \tag{4.12}$$

Many works proved that the oxide-semiconductor surface is not the exclusive resource

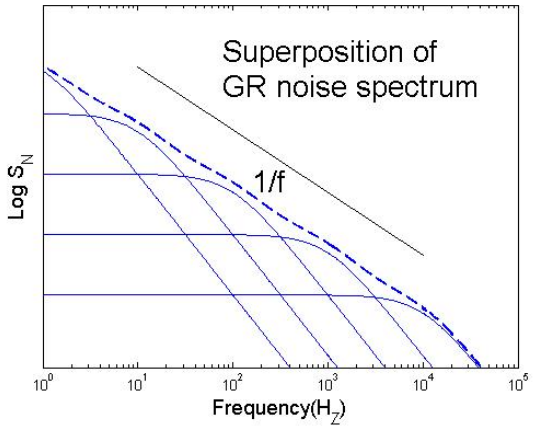


Figure 4.3: The superposition of Lorentzian spectrum.

where the $1/\tau$ distribution can be achieved, for example, D'yakonova et al. [1991] pro-

posed a model where an exponential tail of defect states near the CB causes this kind of distribution as well.

Hooge [2003] investigated that the addition of these individual traps is allowed only when the number of free carriers is much larger than the sum of the carriers in all of the traps. Otherwise, they mix together. The expressions of addition and mix are given:

- Addition: The GR noise spectra is the sum of two or more GR spectra:

$$S = S_A + A_B + \dots$$

- Mixing: the spectrum is one simple Lorentzian with τ is given by:

$$\frac{1}{\tau} = \sum \frac{1}{\tau_i}$$

4.5 LFN Model of SiGe HBTs

All these noise sources above contribute to the SiGe HBTs LFN spectrum independently. In addition, the contribution of thermal noise in the low-frequency range is very small compared to the other. Therefore we can express the LFN:

$$LFN = shotnoise + \sum GRnoise + 1/fnoise$$

where the sum of GR noise is not necessary to follow the $1/\tau$ distribution. In terms of the equation (4.3), equation (4.8) and equation (4.10), this expression can be given:

$$S_{I_B} = 2qI_B + \sum \langle \Delta N^2 \rangle \frac{4\tau}{1 + (2\pi f \tau)^2} + \frac{\alpha_H}{f} \tag{4.13}$$

where S_{I_B} is the spectral density of the base current. Conventionally, the LFN of BJT is denoted by S_{I_B} since the base current is amplified by the transistors themselves and normally constitutes the dominate noise source.

Chapter 5

Power Spectrum Density Estimation

Time series kind signals, such as voltage, current variation, can be collected by laboratory work. These signals include all the information of devices under various working conditions, such as different base current(I_B), biasing and so on. A very important issue in signal process is to estimate the power spectrum density of these time series which are based on a finite set of samples. The PSD can help us to identify the frequencies that carry the signal power or signal energy. Therefore the low-frequency noise PSD shows the signal power along the frequency range from $1\ H_Z$ to $10k\ H_Z$. During recent decades, various digital spectral estimation technologies have been developed, such as Periodogram, Multitaper and so on.

In this section, we will briefly introduce some of these estimation techniques which will be employed in this work later on. Note, using these advance estimation techniques to estimate the PSD is not the main purpose of this thesis, we discussed them here to rather understand the significance of the power spectrum density in this work.

5.1 Periodogram

For a given time series(a realization of stochastic process $X(t)$) based on a finite set of samples $x[n]$, $n = 0, 1, \ldots, N-1$, the periodogram is given [Hanssen, 2003]:

$$\hat{S}_{XX}^{(per)}(f) = \frac{1}{N\Delta t}|\Delta t \sum_{n=0}^{N-1} x[n]exp(-j2\pi fn\Delta t)|^2 \tag{5.1}$$

However, the raw periodogram is not a good spectral estimator since it suffers from spectral bias and variance problems and therefore is treated as an inconsistent esti-

mator. The bias problem arises from a sharp truncation of the sequence, and can be reduced by many different ways, such as first multiplying the finite sequence by a data taper(window), dividing the finite sequence into many segments and averaging, weighted overlapped segment averaging, frequency smoothing and so forth.

5.2 MultiTaper

Multitaper(MT) estimator combines the use of optimal data tapers and average over a set of PSDs which are estimated by these tapers. These optimal tapers $v[n], n = 0, 1, \ldots, N-1$, should follow the maximum "spectral concentration" in which the energy contained in the mainlobe should be maximized relative to the total energy of the taper [Thomson, 1982]. Therefore they are chosen with a discrete Fourier transform $V(f)$, that maximizes the window energy ratio:

$$\lambda = \frac{\int_{-f_B}^{f_B} |V(f)|^2 df}{\int_{-1/2}^{1/2} |V(f)|^2 df} \tag{5.2}$$

where f_B is the expected resolution half-bandwidth of the taper. $\lambda \approx 1$ would therefore be the value that the ideal tapers should have. Slepian [1978] found that the optimal taper $\mathbf{v} = [\mathbf{v[0]}, \mathbf{v[1]}, \ldots, \mathbf{v[N-1]}]^{\mathbf{T}}$ obey the eigenvalue equation:

$$\mathbf{Av} = \lambda \mathbf{v} \tag{5.3}$$

where the A is a matrix with elements $[A]_{nm} = sin[2\pi f_B(n-m)]/[\pi(n-m)], n, m = 0, 1, \ldots, N-1$, corresponding to N pairs $\mathbf{v_k}$ and λ_k, $k = 0, 1, \ldots, N-1$. From these eigenvector and eigenvalue pairs, we obtain a set of orthogonal tapers(eigenvectors $\mathbf{v_k}$) with their corresponding spectral concentration(eigenvalues λ_k). And these orthogonal tapers maximize the ratio(λ) in equation (5.2). The so-called "Discrete Prolate Spheroidal Sequences"(DPSS) was consequently introduced by Slepian [1978] with the resolutions $\mathbf{v_k}$, $\mathbf{v_k^T v_{k'}} = \delta_{\mathbf{k},\mathbf{k'}}$, where $\delta_{k,k'}$ is the Kronecker delta. Once the bandwidth f_B and data length N is decided, one can obtain a sequence of orthogonal taper v_k which could be employed in forming a Multitaper(MT) spectral estimator. The simplest definition of the MT estimator is the average of K tapered "eigenspectra":

$$\hat{S}_{MT}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{S}_{MT}^{(k)}(f) \tag{5.4}$$

where $\hat{S}_{MT}^{(k)}$ is the "eigenspectra" of order k:

$$\hat{S}_{MT}^{(k)} = \left| \sum_{n=0}^{N-1} v_k[n]x[n]exp(-j2\pi fn) \right|^2 ; |f| \leq 1/2 \tag{5.5}$$

where $v_k[n]$ is the kth element of DPSS taper.

The sinusoidal tapers are a simple set of orthogonal tapers that minimize the local bias of the spectral estimator, proposed by Riedel and Sidorenko [1995]:

$$v_k[n] = \left(\frac{2}{N+1} \right)^{1/2} sin \left[\frac{\pi(k+1)(n+1)}{N+1} \right] \tag{5.6}$$

where $k, n = 0, 1, \ldots, N-1$. Note, the sinusoidal tapers have neither eigenvalues connected to, nor the bandwidth f_B.

In practice, both the periodogram and multitaper are evaluated from a finite digital sequence using the Fast Fourier Transform(FFT).

Part II

Experimental Part

Chapter 6

The Measurement Devices and Systems

6.1 Measurement Devices

The devices used in this work are the new generation $375\ GH_Z$ SiGe HBTs issued by IBM recently.

The emitter-area of them are from $0.048\ \mu m^2$ to $1.8\ \mu m^2$. These devices are arranged on six dies, and 20 devices for each. Each terminal of the devices is connected with a square pad so that we can make contact with device by it.

Remarkably, a part of devices are fabricated by several parallel individual SiGe HBTs so that the input current will be evenly separated to every individual transistor, then the output is the sum of every output of them. This structure reduces the influence from each individual transistor, in another word, it eliminates the interferences due to the instability of each transistor in a certain extent. The number of individual transistors in the parallel devices is not same for all of them. There are three types of parallel device consist of 5, 10 and 30 individual transistors respectively. Except for the parallel devices, other devices are fabricated by a single transistor.

6.2 DC Measurements

A conventional way to present the current-voltage behavior of bipolar transistors is the Gummel plot. By definition, the Gummel plot is the combined plot of the collector current, I_C, and the base current, I_B, of a transistor versus the base-emitter voltage, V_{BE}, on a semi-logarithmic scale. This plot is quiet useful in characterizing device figure-of-merit since it reflects on the quality of the emitter-base junction while the base-collector

bias, V_{CB}, is kept at a constant. Moreover, other device parameters, such as the transistor's gain β and leakage currents, can be garnered either quantitatively or qualitatively directly from the Gummel plot.

In this work, by using the Gummel plot of the SiGe HBTs in the measurement, we can easily identify if the transistor is working well or if the probe and the device contact each other tightly so that we can securely continue to do other measurements like LFN measurement following in next chapter.

6.2.1 Gummel Measurement Set-up and Instrument

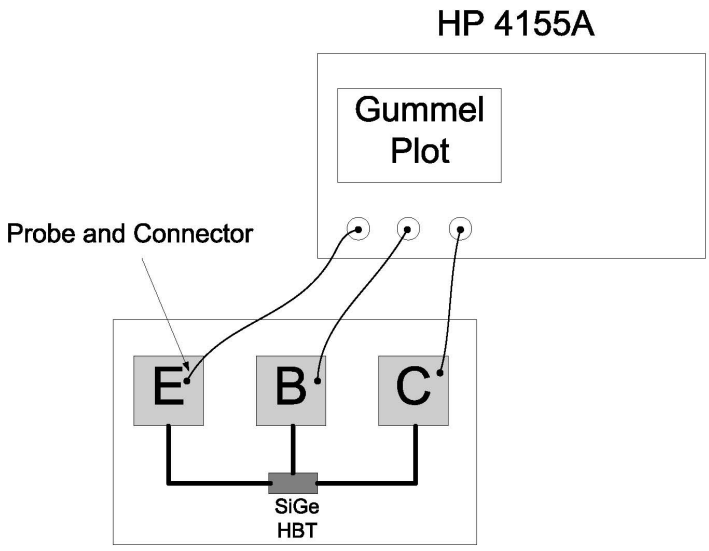


Figure 6.1: The DC measurement set-up for getting Gummel plot.

This measurement set-up is schematically shown in figure 6.1, where the instrument HP 4155A, the semiconductor parameter analyzer, is designed for measuring and analyzing the characteristics of semiconductor device. In this case, three ports of it have been defined as the Base, Collector and Emitter respectively, and are connected with their corresponding terminal of the SiGe HBTs directly by the probes.

The Gummel plot need a stable DC power source to maintain voltage value for each transistor terminal, emitter, base and collector. So the measurement is a DC measurement. HP 4155A can perform rigorous testing, that is, it can force a particular DC voltage or current for the specified duration. Therefore, by configuring the HP 4155A, V_C

and V_B can be set to be constant zero all the way(then $V_{CE} = V_{BE}$), and the scale of $V_{BE} = V_B - V_E$, can be obtained by varying V_E.

6.2.2 The Measured Gummel Plot

The Default configuration measure the values of V_C, V_B, V_E, I_C, I_B and I_E as well as shows the Gummel plot simultaneously on the screen by varying the V_E from $0\ v$ to $-1\ v$ with $-0.002\ v$ each step, that is, V_{BE} varies from $0\ v$ to $1\ v$.

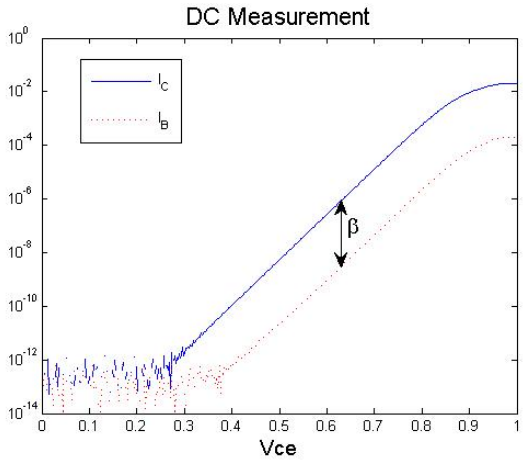


Figure 6.2: DC Measurement.

Figure 6.2 shows one of the Gummel plots we got in many measurements. We see that the device can get gain as long as the V_{BE} is bigger than $0.4\ v$. And the gain does not vary significantly over a wide range, from $0.4\ v$ to $1\ v$ since the I_C and I_B seem parallel each other in this range. Furthermore, both I_C and I_B exponentially increase in the semi-logarithmic scale as the linear rise of V_{BE}. The limitation of the current which the device can handle seems to be reached when V_{BE} is close to $1\ v$ since the curves start to bend down.

This Gummel plot shows that this device is working fine and can be further used in other measurements. On the other hand, the damaged devices or bad connections would give us a totally different Gummel plot, for example, the I_B value of a damaged device could immediately reach a very high current level despite of the I_C once the V_{BE} is larger than zero, which indicates a damaged EB-junction.

Another advantage the Gummel plot offers us is that we can directly calculate the transistor's gain by equation (2.8) from it. Figure 6.3 shows the gain varying along the

V_{BE}, from which we can see that the transistor will get maximum gain when V_{BE} is around $0.7\ v$ which could be the value that people would like to let the device work on.

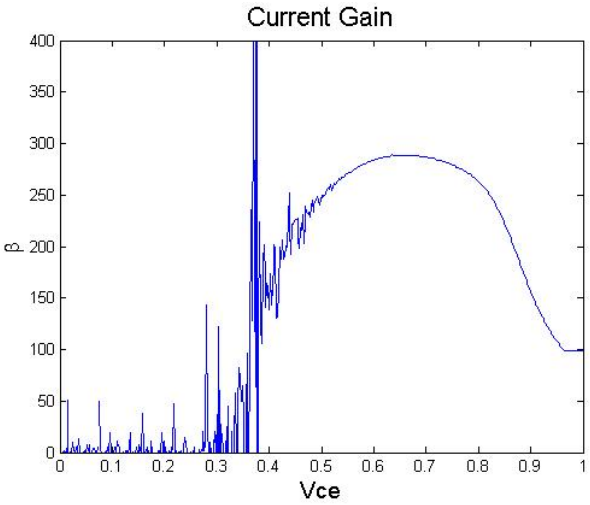


Figure 6.3: Current Gain.

6.3 The LFN Measurement System

A measurement system will be described in this chapter for characterizing the Low-Frequency Noise of SiGe HBTs. Such system is necessary to the measurement since it not only severely reduces the measurement time, but also efficiently increases the measurement precision. A schematic circuit diagram of the measurement system, using the common emitter set-up similar to that of figure 2.11, is shown in figure 6.4.

6.3.1 Dynamic Signal Analyzer

In figure 6.4, the dynamic signal Analyzer HP 3561A is a single channel Fast Fourier Transform(FFT) signal analyzer covering the frequency range from $1\ H_Z$ to $100k\ H_Z$ which is sufficient for the Low-Frequency Noise measurement. The self tests of HP 3561A provide maximum confidence in the operation of the instrument. These tests, in conjunction with the internal calibration signal, make it possible to quickly verify the calibration of the instrument before starting a critical measurement sequence. The remarkable performances of HP 3561A therefore can give us reliable measurement results. In this case, it is also configured to be controlled by the computer program(LabView)

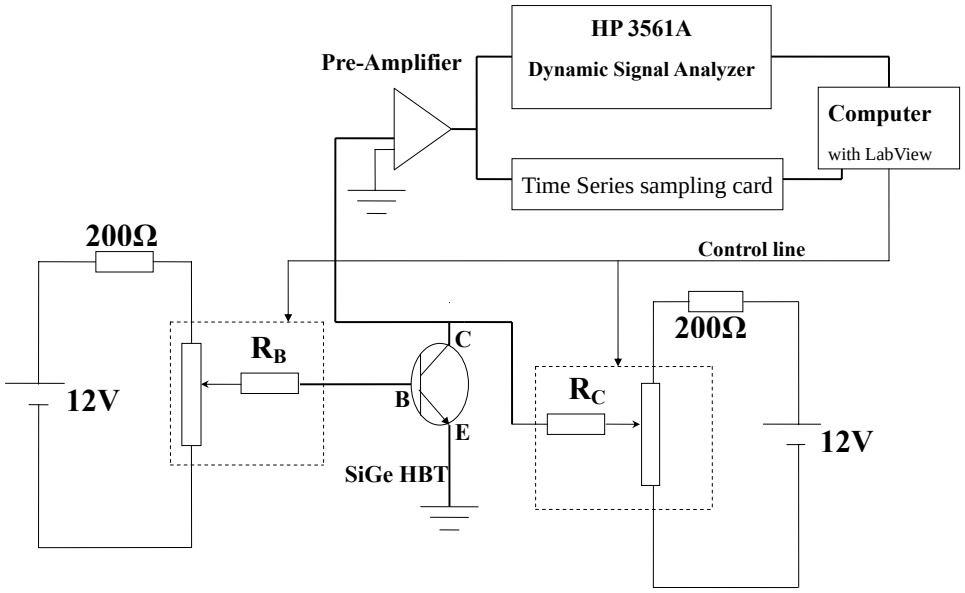


Figure 6.4: Low-Frequency Noise Measurement System.

remotely, so that we can conveniently set all the measurement parameters in LabView and send them to HP 3561A to do the measurement. Once the measurement has been done, all the information will be transferred back to computer, and this information can be used for further analysis.

6.3.2 The Operation Mechanism of Measurement System

From figure 6.4, we see that it is possible to realize many of voltage and current bias requirements for the transistor by tuning the status of the wire-wound resistors. And the tuning procedure is totally automatic under the control of the computer as well.

The bias voltages come from batteries which can significantly lower the interferences from the measurement system itself, because batteries are stable power sources and won't disturb the noise measurements. Wire-wound resistors, which are driven by servo-motors, are used to adjust the I_B and I_C under the command of the computer software, LabView. A particular LabView program has been developed to set I_B value to a specified value and measure other values in the system simultaneously, such as

I_C, I_E, and the voltages of the batteries, etc. Once the measurement system has the I_B set, the collector voltage noise ΔV_C will be amplified by a Low-Noise Pre-amplifier and then sent into the computer controlled Dynamic Signal Analyzer HP 3561A. The dynamic signal analyzer records the time series of a voltage/current fluctuation, and calculates the power spectrum density by the Fast Fourier Transform(FFT). Consequently, the measured PSD will be sent back to the computer and taken over by LabView again for saving the data file as well as for displaying PSD on the screen simultaneously.

For measuring as pure as possible noise spectra of devices, we made the most of various steps to avoid any possible interferences, for instance, shielding the devices in a steel box, using as short as possible cable and shutting down all other irrelevant instruments, even including the computer screen.

To sum up the above, this LFN measurement system can easily characterize the LFN of SiGe HBTs under various conditions, such as different I_B. And we can adjust other circuit parameters(e.g. R_B) in case of necessary to achieve other purposes, for instance, looking for the dominant noise source which will be discussed in section 7.1.

6.3.3 Time Series Sampling

The time series sampling card, shown in the figure 6.4, is a equipment from which we can sample the amplified collector voltages as a time series. The sampling frequency and sampling time scale are also set by a particular LabView program. Generally, we measured 10 seconds time series using $300kH_Z$ as the sampling frequency. This sampling frequency assure that we have not lost any information in the frequency range of our interest.

These sampled time series then can be analyzed by other statistical tools, such as Multitaper and Wavelet, as a comparison of the FFT results from the Dynamic Signal Analyzer HP 3561A.

Chapter 7

LFN Power Spectrum Density Measurements

In this chapter, by applying the LFN measurement system which we mentioned in section 6.3, a series of systematical PSD measurements have been done for investigating the performances of the SiGe HBTs in various aspects, such as finding the dominant noise source, current dependence, emitter-area dependence, noise variation and so on. It is noticeable that most of these measured PSDs show very "bumpy" behaviors instead of the normally expected $1/f$ behavior. That is, these new generation SiGe HBTs normally have strong GR noise spectra along the low-frequency scale.

7.1 Searching The Dominant Noise Source

There are many of noise sources possibly physically located in the regions of the SiGe HBTs. Processes like diffusion, recombination, tunneling, trapping or others could have connection with these noise sources. That is, for each kind of junction, emitter-base junction, emitter-collector or base-collector, these processes may exist to produce the noise. The contribution of all these noise sources should be concerned in the measurement in order to get a clean Power Spectrum Density.

7.1.1 The Noise Model of BJTs

Fortunately, just a few of these noise sources dominate the LFN of a transistor in a particular situation and these noise sources can be put into a transistor model, such as the model in figure 7.2 which shows the most likely noise sources in a small-signal

hybrid π model of one SiGe HBT, i.e. the base current noise, S_{I_B}, collector current noise, S_{I_C}, and parasitic resistance noise, S_{V_rb}, S_{V_rc} and S_{V_re}.

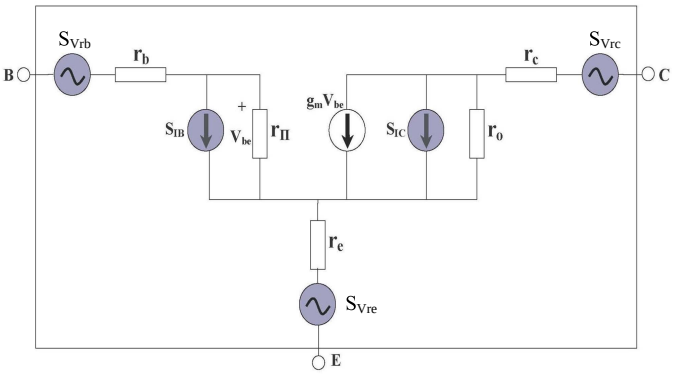


Figure 7.1: π model small-signal circuit with noise source in a SiGe HBT.

For investigating how these noise sources affect the LFN, we need look into the measurement system we actually used. For the measurement system shown in figure 6.4, with the help of figure 7.1, the equivalent circuit can be easily drawn as figure 7.2, where the parasitic resistance noises are not shown in this figure for the sake of brevity, but are concluded in the noise calculation(equation (7.1) and equation (7.2)).

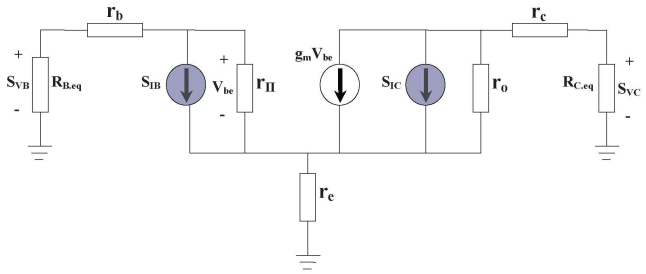


Figure 7.2: Common-emitter equivalent circuit.

Assuming all these noise sources are independent each other, then the spectrum of

ΔV_C can be expressed as[Jin, 2004]:

$$\begin{aligned}
\frac{S_{V_C}}{R_{C.eq}^2} &= \frac{1}{m}\left(m_{I_B}S_{I_B} + m_{I_C}S_{I_C} + n_{rb}S_{V_{rb}} + n_{re}S_{V_{re}}\right) \qquad (7.1)\\
m &= \left[r_O\left(r_b + R_B + r_\pi + (1+\beta)\,r_e\right) + \left(R_C + r_c + r_e\right)\left(r_e + r_b + r_\pi + R_B\right) - r_e^2\right]^2\\
m_{I_B} &= \left[r_O\beta\left(R_B + r_b + r_e\right) + r_\pi r_e\right]^2\\
m_{I_C} &= \left[r_O\left(R_B + r_b + r_e + r_\pi\right)\right]^2\\
n_{rb} &= \left(\beta r_O - r_e\right)^2\\
n_{rc} &= \left(r_b + r_e + r_\pi + R_B\right)^2\\
n_{re} &= \left(\beta r_O + r_\pi + r_b + R_B\right)^2
\end{aligned}$$

A accustomed way to find the dominate noise source is changing the external base resistance $R_{B.eq}$, that is the R_B in figure 6.4. The contribution of $S_{I_B}, S_{I_C}, S_{V_{rb}}, S_{V_{rc}}$ and $S_{V_{re}}$ to the LFN as the changing of $R_{B.eq}$ can be calculated by:

$$\begin{aligned}
\frac{S_{V_C}}{R_{C.eq}^2} &= k_{I_B}S_{I_B} + k_{I_C}S_{I_C} + l_{rb}S_{V_{rb}} + l_{rc}S_{V_{rc}} + l_{re}S_{V_{re}} \qquad (7.2)\\
k_{I_B} &= \left[\beta\frac{R_B + r_b + r_e}{R_B + r_b + r_e(1+\beta) + r_\pi}\right]^2\\
k_{I_C} &= \left[\frac{R_B + r_b + r_e + r_\pi}{R_B + r_b + r_e(1+\beta) + r_\pi}\right]^2\\
l_{rb} &= \left[\frac{\beta}{r_b + R_B + (1+\beta)r_e + r_\pi}\right]^2\\
l_{re} &= \left[\frac{\beta r_O + R_B + r_b}{r_O(r_b + R_B + (1+\beta)r_e + r_\pi)}\right]^2\\
l_{rc} &= \left[\frac{R_B + r_b + r_e + r_\pi}{r_O(r_b + R_B + (1+\beta)r_e + r_\pi)}\right]^2
\end{aligned}$$

which is a simplification of equation (7.1) due to the usual condition $r_O \gg R_{C,eq}, r_c, r_e$.

If some typical number for circuit components($I_B = 10\mu A, \beta = 100, R_C = 1k, r_O = 50k, r_\pi = 2.6k, r_b = r_e = r_c = 10$) and noise spectra measured at 1 H_Z($S_{I_B} = 1\times10^{-20}$ $A^2/H_Z, S_{I_C} = 5\times10^{-18}$ $A^2/H_Z, S_{V_{rb}} = 1.6\times10^{-19}$ $V^2/H_Z, S_{V_{rc}} = 1.6\times10^{-15}$ V^2/H_Z, and $S_{V_{re}} = 1.7\times10^{-15}$ V^2/H_Z) are chosen [L.S.Vempati et al., 1996] [R. Brederlow and Thewes, 2001], then we can plot the output noise spectra S_{I_C}, which is calculated by these typical values, as shown in Figure 7.3.

The situation, which is shown in this plot, where S_{I_C} increases by increasing $R_{B,eq}$,

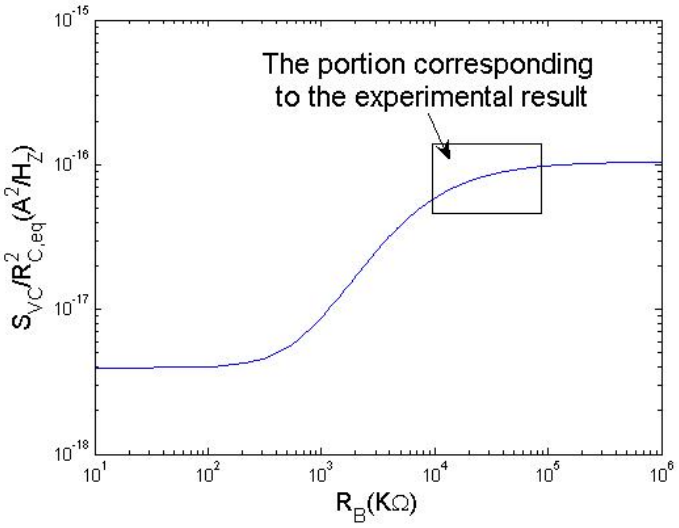


Figure 7.3: One typical noise output versus $R_{B,eq}$.

states that the base current noise spectra S_{I_B} dominates the noise output. According to the knowledge of elementary circuit principle, the current always would like to follow the way where smaller resistance exists, the lager $R_{B,eq}$ could force the current to choose the easier way, that is, somewhat more current flow into the base and increase the part of S_{I_B} which will be amplified to the collector by the transistor and contribute a part of S_{I_C}. Many works such as S. P. O. Bruce and Rydberg [1999] and L.S.Vempati et al. [1996] show that the base current noise is usually the dominant noise source in common-emitter configuration since it is amplified by the transistor. This fact is represented more significantly in SiGe HBTs because of its characteristic of high gain(β). The relationship between S_{I_B} and S_{I_C} can be easily expressed as:

$$S_{I_B} = \frac{S_{I_C}}{\beta^2} \tag{7.3}$$

where $S_{I_C} = \frac{S_{V_C}}{R^2_{C,eq}}$

7.1.2 The Measurement and Result

A set of Power Spectrum Density have been measured at $R_B = 10k\ \Omega$, $20k\ \Omega$, $30k\ \Omega \cdots$ $100k\ \Omega$ by using one particular device with $A_E = 1.8\ \mu m^2$ under $I_B = 0.2\ \mu A$. The LFN measurement system somehow does not function to find any working point when we choose the R_B value which is less than $10k\ \Omega$. In addition, the maximum value of the

Wire-wound resistor is $100k\ \Omega$. So only the PSDs that R_B within the range $10k - 100k$ have been measured.

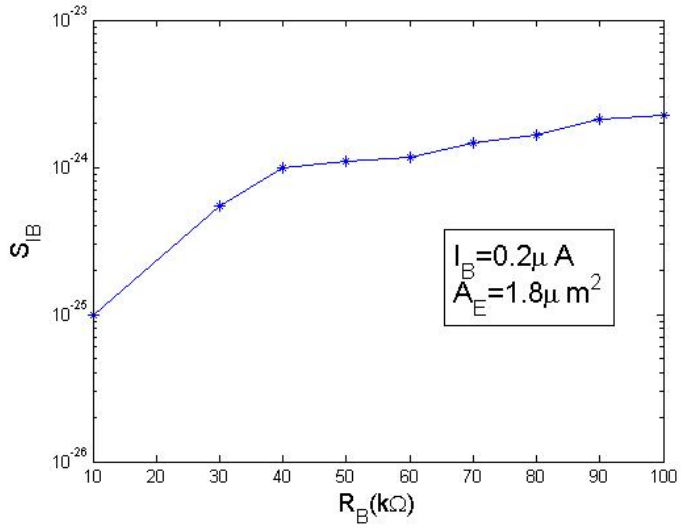


Figure 7.4: The base resistance dependence of the S_{I_B}.

By picking up each PSD value at $f = 10\ H_Z$, we can easily plot these noise spectrum levels versus the corresponding R_B in Figure 7.4. We can see that the singularity occurs around $R_B = 40k\ \Omega$. From $10k\ \Omega$ to $40k\ \Omega$, the noise levels increase quickly. Nevertheless, they are almost stay at the same level after $40k\ \Omega$, even though still increasing slightly. Despite of the relatively narrow measurement scale, this result plot shows the consistent trend compare to the corresponding portion of the figure 7.3, which proves that this result is coherent with the foregoing theoretical conclusion, i.e. the dominant noise source is the base current noise.

7.2 The Base-Current Dependence of LFN

From section 7.1, we found that the Base Current I_B is the dominant noise source which means that the noise level of LFN will vary by changing the value of I_B. It is significant that the impact of I_B should be well known in advance before the transistor can be used in the circuit. P.-F.Lu [1987] found the relationship $S_{I_B} \sim I_B^2$ in self-aligned bipolar transistor, and Plana et al. [1995], L.S.Vempati et al. [1996], A.Gruhle and C.Mahner [1997] and S. P. O. Bruce and Rydberg [1999] also found the similar principle in SiGe heterojunction bipolar transistor.

In this section, the PSD of the SiGe HBTs have been measured at several different I_B, we also chose the devices with different emitter area to see if the I_B has a similar impact for those different candidates.

7.2.1 A Theoretical Model for LFN I_B Dependence

For elucidating the experimental discovery, $S_{I_B} \sim I_B^2$, a theoretical model is established in terms of unit area(A_{E_U}). Assuming that both the base current density J_B and the noise sources are homogenously distributed over the emitter and spatially uncorrelated, we can divide the emitter into n independent unit areas and the J_B is evenly distributed among all these n unit areas. In each unit area, the noise spectrum is given by:

$$\begin{aligned} S_{I_BU} \sim A_{E_U} J_{B_U}^k &= A_{E_U} \left(\frac{I_B}{n \cdot A_{E_U}} \right)^k \\ &= \frac{I_B^k}{n^k \cdot A_{E_U}^{k-1}} \end{aligned} \tag{7.4}$$

where the parameter k in this model is generally equal to 2 according to the observations above and the base current density J_B is the fundamental mechanism in this model. The total noise spectrum is therefore the sum of all these individual spectrum:

$$\begin{aligned} S_{I_B} &= n \cdot S_{I_BU} \\ &= \frac{I_B^k}{n^{k-1} \cdot A_{E_U}^{k-1}} \\ &= I_B^k \cdot A_E^{1-k} \end{aligned} \tag{7.5}$$

A similar theoretical model has been given by H.A.W.Markus and T.G.M.Kleinpenning [1995].

7.2.2 The Measurements and Results

Five I_B values have been taken in this experiment, $0.2\ \mu A$, $0.5\ \mu A$, $0.8\ \mu A$, $1\ \mu A$ and $2\ \mu A$. The SiGe HBTs with $0.3\mu m^2$ and $0.6\mu m^2$ were chosen, and there are several devices for each area in order to get better statistical results.

Figure 7.5 shows the magnitude of base current noise spectrum S_{I_B} versus the corresponding I_B. The value of these S_{I_B} are picked up at $10\ H_Z$. The gradients we got for these two kinds of devices are 2.4750 and 2.0532 respectively which are close to the

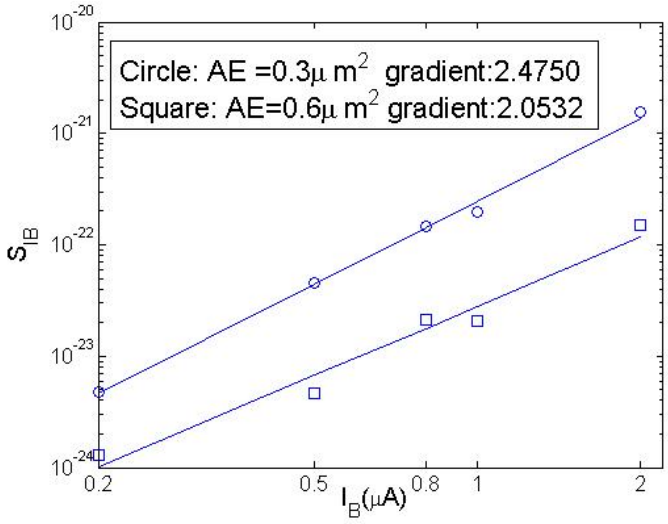


Figure 7.5: $S_{I_B} \propto I_B^2$.

generally observed value $k = 2$. Thereby, we extrapolate that the PSD of SiGe HBTs with different emitter-area have a similar I_B dependence, i.e. $S_{I_B} \sim I_B^2$.

7.3 The Emitter Geometrical Scaling Dependence of LFN

In modern industry, transistors are greatly down scaled geometrically in order to improve their performance and adapt to higher integration level. One design factor associated with geometrical scale is that the Low-Frequency noise shows different frequency dependence and noise level from device to device with distinct geometrical property.

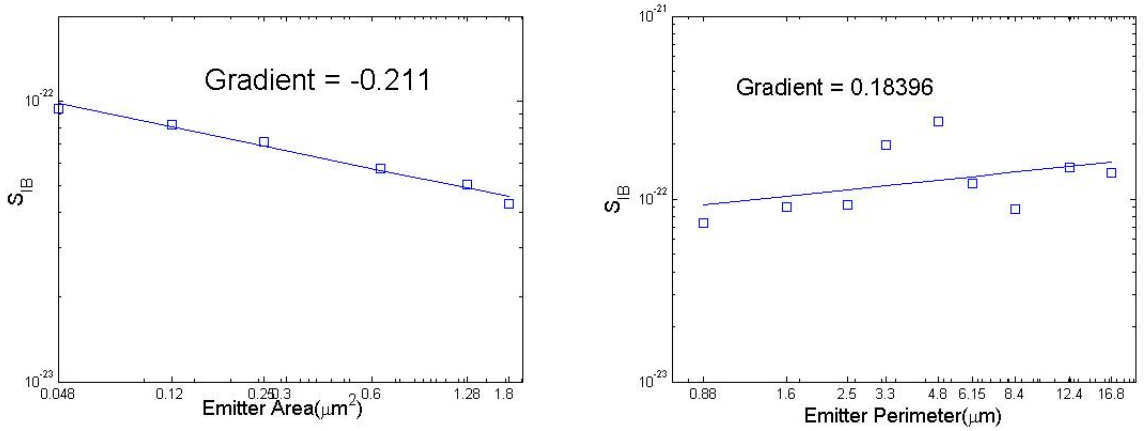
In this section, we will measure many devices with different type of geometrical scaling, the Emitter-Area and Emitter-Perimeter, to study the geometrical scaling dependence of LFN.

7.3.1 The Theoretical Model for Emitter-Area Dependence

In terms of the theoretical model shown in equation (7.5) in section 7.2 and the usual value $k = 2$, under the same I_B value, the S_{I_B} therefore is reversely proportional to A_E, i.e. $S_{I_B} \sim A_E^{-1}$. This trend also has been observed in many works [Jin et al., 2003] [L.S.Vempati et al., 1996].

7.3.2 The Measurements and Results

Many of devices with different Emitter-Area($0.048\ \mu m^2$, $0.12\ \mu m^2$, $0.25\ \mu m^2$, $0.3\ \mu m^2$, $0.6\ \mu m^2$, $1.28\ \mu m^2$ and $1.8\ \mu m^2$) and Emitter-Perimeter($0.88\ \mu m$, $1.6\ \mu m$, $2.5\ \mu m$, $3.3\ \mu m$, $4.8\ \mu m$, $6.15\ \mu m$, $8.4\ \mu m$, $12.4\ \mu m$, $16.8\ \mu m$) have been measured at $I_B = 1\ \mu A$.



(a) Emitter-Area dependence of LFN

(b) Perimeter dependence of LFN

Figure 7.6: The Emitter Geometrical Scaling Dependence of LFN.

Figure 7.6(a) shows the emitter-area dependence of S_{I_B}, $S_{I_B} \propto A_E^{-0.2}$, which is far away from the theoretical model as well as the general observation, $S_{I_B} \propto A_E^{-1}$. This difference happened possibly due to two reasons. First, the general observation is obtained from either very big devices or old fabrication technologies. These new generation "bumpy" SiGe HBTs, however, are not necessary to follow this relation due to their extremely down-scaled sizes and new fabrication technology; Second, we do not have sufficient number of the device for each kind of emitter-area to get good statistically average results, moreover the devices we used sometimes behavior as very noisy or quiet device compared to the noise level of other same size devices. The Emitter-Perimeter dependence presents a contrary trend compare to the Emitter-Area dependence as shown in figure 7.6(b) where the magnitude of the S_{I_B} increases by the increasing of the emitter perimeter, which could be meaningful since the longer perimeter of emitter will introduce more GR traps [Jin et al., 2003] and consequently higher the noise level.

7.4 The Noise Variation

The stability of the device should be concerned. Therefore, investigations about the noise variation on various devices with different emitter-area or under different base-current will be systematically performed in this section. And the noise variation coefficient(δ) will be calculated for scaling the stability. This coefficient is given by [Johansen et al., 2003][Jin et al., 2002][Joseph et al., 2001]:

$$\begin{aligned} \delta &= \frac{1}{S_{I_B.avg}} \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (S_{I_B.i} - S_{I_B.avg})^2} \\ S_{I_B.avg} &= \frac{1}{N} \sum_{i=1}^{N} S_{I_B.i} \end{aligned} \tag{7.6}$$

where the N is the number of samples and i indicates the $i'th$ sample.

7.4.1 The Variations Among Devices With Same A_E

Under the same I_B value, Low-Frequency Noise has been measured for several devices. It shows a larger statistical scatter in the very small devices compared to the larger devices.

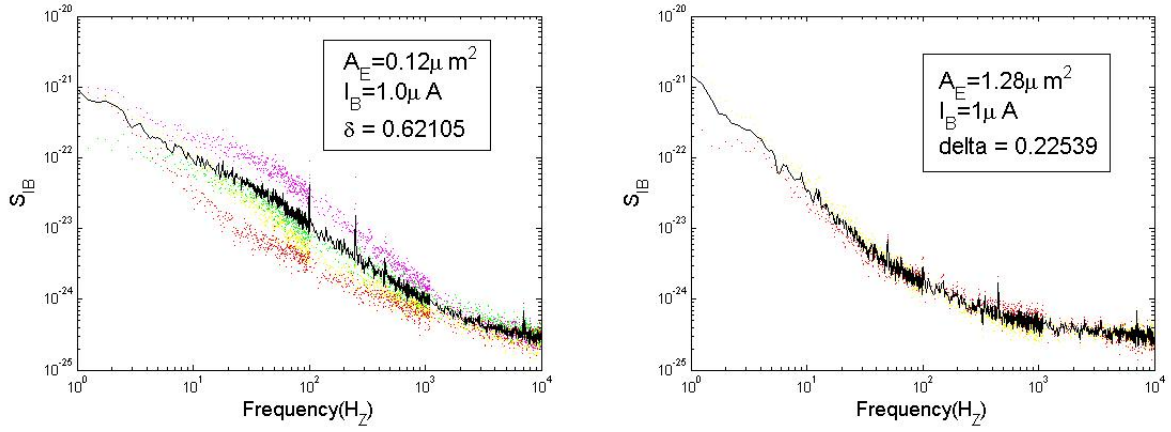
The Measurements and Results

The devices with two kinds of emitter-area, $A_E = 0.12\ \mu m^2$ and $A_E = 1.8\ \mu m^2$, have been chosen as the small and large devices respectively, and are measured at $I_B = 1.0 \mu A$.

Figure 7.7 shows that the variation of the small devices, $\delta = 0.62$, is much bigger than that of large devices, $\delta = 0.23$. Therefore, the large devices with the same geometrical scaling offer good LFN stability. The LFN of small devices, however, could behave very differently from device to device even though they have the same geometrical scaling.

7.4.2 The Noise Variation of One Single Device - Single Variation

After investigating the variation from device to device, we want to know the variation among several PSDs which are measured on one particular device at different time instead of one type of devices with the same A_E. For the brevity and distinguishing from the variation we mentioned in section 7.4.1, we call this variation as *single variation*. The I_B and A_E dependence of single variation will be investigated in this section.



(a) The noise variation on very small devices (b) The noise variation on large devices

Figure 7.7: The variation depended on emitter-area.

I_B Dependence of Single Variation

Two devices with $A_E = 0.12\ \mu m^2$ and $A_E = 1.8\ \mu m^2$ respectively have been chosen. At $I_B = 0.2, 0.5, 1.0$ and $2.0\ \mu A$, eight PSD-samples have been respectively measured for both of them.

The figure 7.8 shows the single variation of these two devices versus the base current I_B. We see that the larger I_B offer smaller single variation for both of them, that is, the LFN of the devices is more stable under higher base current. In addition, comparing between these two kinds A_E of device, the small device shows a strong distinction among the single variations so that we can clearly see the I_B dependence, that is, the variation decreases by increasing I_B. Therefore, the small device operates more steadily at large I_B. It is still hard to tell if the largest I_B must most minimize the variation, but we may use the single variation to find the optimum I_B. On the other hand, the single variations of large device are actually not significantly affected by I_B. Therefore, we believe that the large device would perform more reliably compared to small device under various I_B conditions.

A_E Dependence of Single Variation

In last section, when we investigate the I_B dependence of single variation, we found that the large device shows a better stability(smaller single variation) compared to the small device under every base current I_B. Therefore it is clear that the single variation

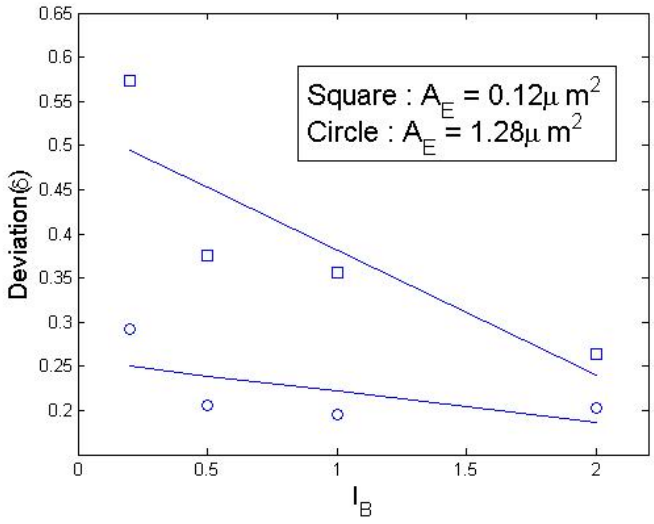


Figure 7.8: The noise deviation of a particular device under different I_B value.

has A_E dependence.

Four devices with $A_E = 0.12, 0.64, 1.28$ and $1.8\ \mu m^2$ have been chosen. Under $I_B = 2\mu A$, which is supposed to be the most stable condition we can get, eight samples for each device have been measured and applied to calculate the noise variation coefficient of these four devices.

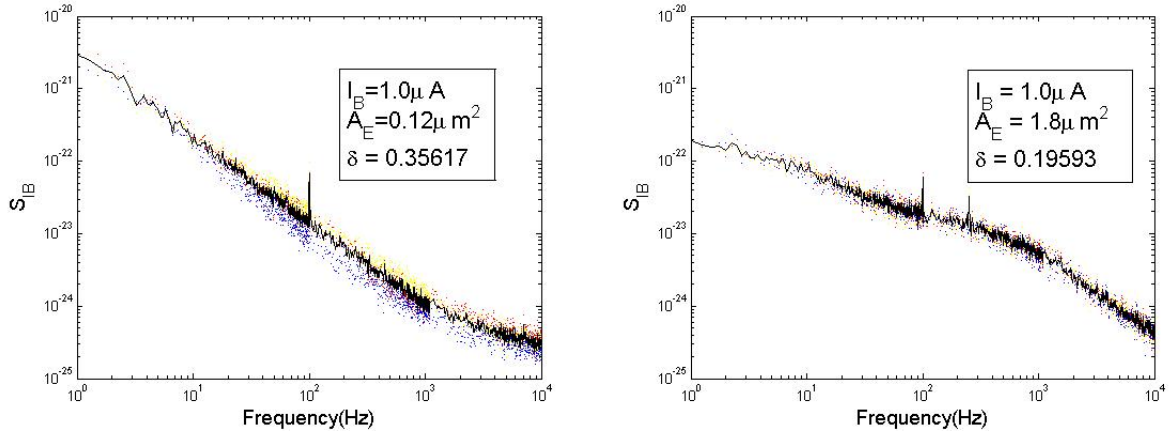
Figure 7.9(a) and figure 7.9(b) show the PSD of three of the samples for small and large device respectively. The solid lines is the average value of them. From equation (7.6), the deviation of both of them can be calculated, the deviation of small device,$\delta = 0.356$, is obviously bigger than the result of large device ,$\delta = 0.196$.

Figure 7.10 shows all these four single variations, we can see the trend that the smaller devices show stronger single variations compared to the bigger devices.

Note, we did not calculate the exact value of both the I_B dependence and the A_E dependence because of the reason that we do not have sufficient number of good device to get the expected statistical results.

The Single Variation Due To The Fabrication of The Transistor

Comparing the single variation between the parallel and single devices, introduced in section 6.1, is another case which has been concerned. Two SiGe HBTs with the same emitter area, $A_E = 1.28\ \mu m^2$, have been chosen. One of them is fabricated by 10 parallel individual transistors, and another one just include one single transistor. Again, eight



(a) The single variation of Small device

(b) The single variation of Large device

Figure 7.9: The single variation of two devices with different A_E.

PSD-samples have been measured for the statistical reason.

Figure 7.11 shows the results of the comparison between the single and parallel devices. Intuitively, the parallel device should have a smaller deviation compared to the single device because the output of the parallel device is averaged by its internal parallel individual devices. In fact, the measurement results are consistent with this intuition. The single variation of the parallel device is 0.2 which has sufficient distinction from the result of the single device, $\delta = 0.3$.

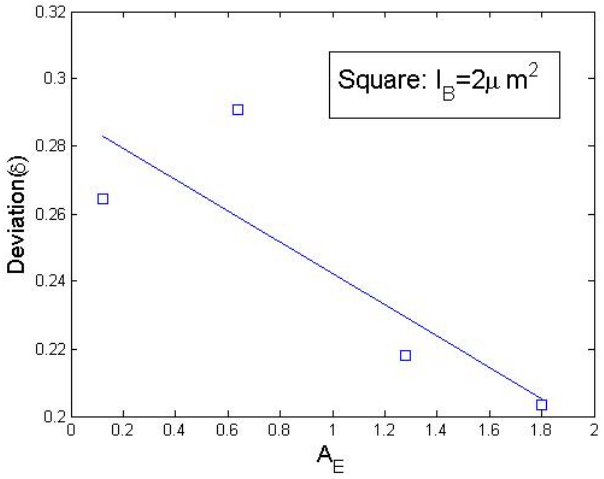
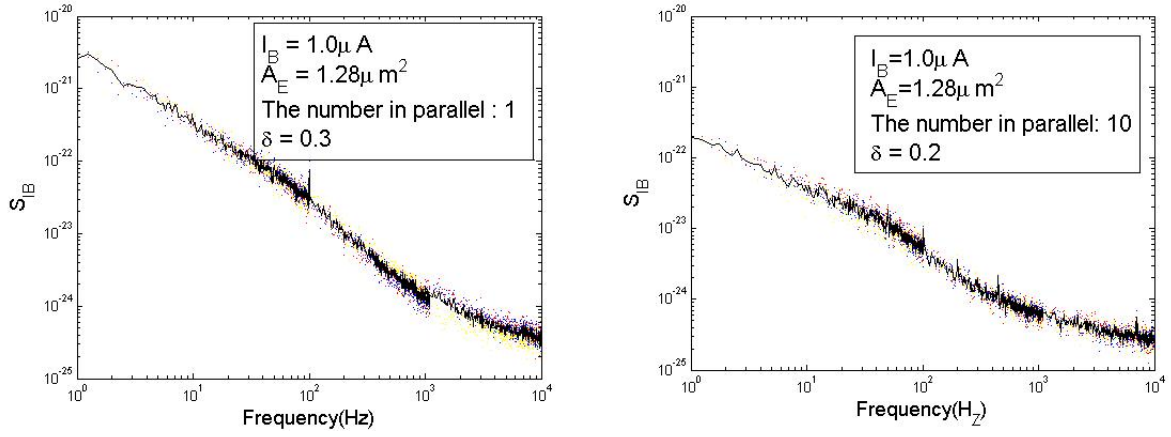


Figure 7.10: Under the same I_B, the noise deviation changing by the A_E.



(a) The single variation of Single device

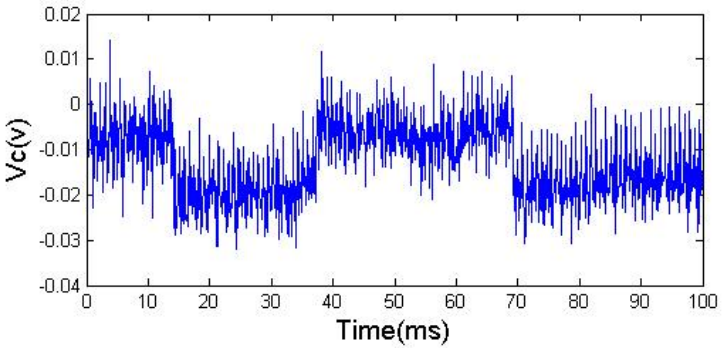
(b) The single variation of Parallel device

Figure 7.11: The single variation due to the transistor fabrications.

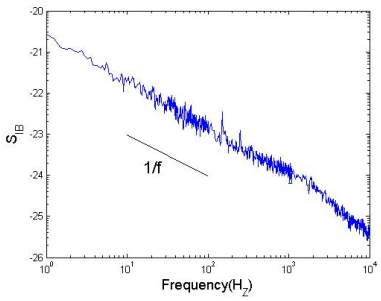
Chapter 8

Time Domain Analysis

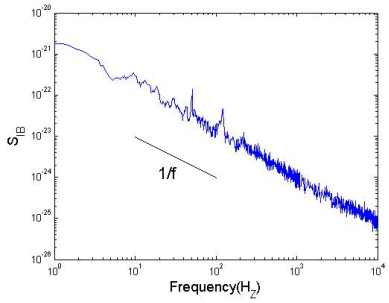
In this chapter, we will use one of the PSD estimators, Multitaper(MT) introduced in chapter 5, to estimate the PSD of a time series which is sampled by the "time series sampling card" shown in figure 6.4. The estimation result shows the consistent PSD with that of estimated by 3561A dynamic signal analyzer.



(a) Timeseries(100 ms)



(b) PSD estimated by 3561A



(c) PSD estimated by Multitaper

Figure 8.1: PSD estimations.

A part of $10\ sec$ time series sampled under $300k\ H_Z$ is shown in Figure 8.1(a). In terms of the MT estimation method introduced in chapter 5, by using 10 sinusoidal tapers(equation (5.6)), we estimated the PSD in each $1\ sec$ respectively and averaged them in order to get a better statistical result. The estimated PSD is then shown in figure 8.1(c).

Comparing the PSD estimated by MT with the one measured by 3561A in figure 8.1(b), both of them give the similar PSD estimation. In laboratory, however, we prefer to use the dynamic signal analyzer to measure the PSD instead of sampling the time series separately because the time series are somehow difficult to be measured without interferences, on the other hand, as we introduced in section 6.3.1, HP 3561A provides maximum confidence in the operation of the instrument by its self tests.

Chapter 9

Fitting the Power Spectrum Density

In this chapter, a method will be introduced to fit the power spectrum density based on the established model in terms of a non-linear fitting method. The background of developing this method is the requirement of investigating the various device properties such as the I_B dependence in section 7.2, the emitter geometrical scaling dependence and so on. As we already known, these investigations normally compare the magnitude of the PSD, which are measured from different conditions, at one particular frequency. Therefore, picking up the correct magnitude value for all the chosen PSD is the foundation of studying the device properties. However, the drastic fluctuation of the original PSDs is definitely disadvantageous for this purpose. So a smooth fitting curve of the original PSD is necessary before we try to get the magnitude. Then the value of the fitting curve can be used as the meaningful PSD magnitude. Moreover, we hope that the fitting curve should be able to reflect the physical characteristics of the device in some sense in order to help us study the noise property of each individual device.

Non-linear fitting method will be employed as the core fitting procedure. To make this method work, a proper fitting function of the non-linear fitting procedure is necessary. For this purpose a theoretical model of low-frequency noise PSD mentioned in chapter 4, which consists of three noise levels, $1/f$ noise, GR noise and shot noise, will be engaged as the fitting function. We will use these three noise levels and all the GR noise centers as the fitting parameter. And the initial condition of the fitting parameter will be chosen by using the classification between the test PSD and the PSD model database for the sake of having a automatical fitting mechanism instead of visually predict all the parameters every time. The PSD model database can be established by using theoretical noise levels or practical noise level which will accordingly be called *theoretical fitting method* and *practical fitting method* respectively. Consequently, the final condition

of the fitting parameter can not only give the fitting curve, but also clearly show the information of each noise component after the fitting procedure.

Furthermore, this fitting method can be applied to identify some parameters such as K_F, C from device to device.

9.1 Non-linear Fitting Procedure

Normally both Polynomial fitting and Non-linear fitting can be employed to fit a data set. Polynomial-fitting has been tested and it works only when the $1/f$ noise dominate the Low Frequency noise. Once there are several strong GR-noise inside the PSD, the result of the Polynomial-fitting can not be trusted any more. In addition, this method can not give any physical information about the device as we expected in its result. On the other hand, the non-linear fitting method can possibly fit any kinds of shape by finding its local minimizer as long as it has a proper fitting function and initial conditions. Moreover, with the help of the fitting function, we could easily get the approximate physical property of the device if only the measured PSD can be fitted well. Because of these advantages of the Non-linear fitting method, we employed it as the core fitting method. In the next step, we should have a good fitting function for it. Based on the theoretical LFN model we discussed in chapter 4, we now develop a functional form of this model to be used as the fitting function.

9.1.1 The Fitting Function - The Theoretical Model of LFN

From the Theoretical Part in chapter 4, we established that the Low-Frequency of semiconductor consists of $1/f$ noise, GR-noise and shot noise, etc.

$1/f$ noise

In circuit simulation programs like SPICE, S_{I_B} is given, modeled by two parameter K_F and A_F:

$$S_{I_B} = \frac{K_F \times I_B^{A_F}}{f} \tag{9.1}$$

In equation (9.1), K_F is an empirical parameter which represent the magnitude of the flicker noise at $1H_Z$ and $1A$. It is a constant number for each device, but has the experimental relationship $K_F \sim A_E^{-1}$ [Sanden et al., 2002]. A_F represent the current dependence of $1/f$ noise which normally is proportional to I_B^2 and thus $A_F = 2$, but this value

will be changed to 2.2 in the theoretical fitting method(section 9.3) in terms of some experimental results.

Generation-Recombination noise

GR noise, from a number of independent sources can be represented by:

$$S_{I_B} = \sum_{i=1}^{N_T} C \frac{J_B^2 \tau_i}{1 + (2\pi f \tau_i)^2} \tag{9.2}$$

Where N_T is the number of traps, C is a constant, $J_B = I_B / A_E$ is the base current density.

Shot noise

The level of shot noise is based on the current:

$$S = 2qI \tag{9.3}$$

Low Frequency noise

Low Frequency noise is the superposition of $1/f$ noise, GR noise and shot noise, i.e. equation (9.1), equation (9.2) and equation (9.3):

$$\begin{aligned} S_{I_B} &= 1/f + \sum GR + 2qI \\ &= \frac{K_F \times I_B^{A_F}}{f} + \sum_{i=1}^{N_T} C \frac{J_B^2 \tau_i}{1 + (2\pi f \tau_i)^2} + 2qI_B \end{aligned} \tag{9.4}$$

The test PSD will be fitted by using this model as the fitting function for the non-linear procedure.

9.1.2 Fitting Parameters

The non-linear fitting procedure uses the parameters in the fitting function to achieve the fitting purpose. The parameters are changed during the procedure to find the best fitting result. In this case, the levels of each noise source type($1/f$, GR and shot noise) are independent and need to be fitted respectively, therefore they all should be a part of the fitting parameters. For $1/f$ and shot noise, only the magnitude needs to be fitted. For GR noise, however, both the magnitude and the time constants of the GR noise

centers should be considered. Eventually, we have four types of fitting parameters in the fitting function, the magnitude of $1/f$ noise, GR noise, shot noise and the time constants of the GR noise centers. In equation (9.4), they are $K_F \times I_B^{A_F}$, $C \times J_B^2$, $2qI_B$ and τ_i respectively. Having the final conditions of these combined fitting parameters, we then can calculate all the parameters in the model, i.e. K_F, C, J_B, I_B, where the I_B is actually not the measured I_B but the fitted I_B, and J_B can be calculated by $J_B = I_B/A_E$.

9.1.3 The Initial Condition of Fitting Function

The non-linear fitting procedure always need a threshold for its fitting parameters when it starts to search the local minimizer. Here, this threshold is called initial condition. Correspondingly, final condition is used to name the end point of this fitting procedure. In this case, the initial condition becomes the initial assumption of the fitting parameters which are given in section 9.1.2. As we will present, a good initial condition(closer to the real noise status of the test PSD) is necessary to quickly find the best fitting result. Since we already have the fitting function and its fitting parameter, the initial condition becomes the next important issue which should be addressed.

Testing the significance of the initial condition

Trials have been done in which we artificially built some PSDs with the preset fitting parameters, i.e. the known noise levels of $1/f$ noise, shot noise and GR noise, and the GR time constants τ_i in equation (9.4). Since we know exactly each component of these synthetic PSDs, the non-linear fitting gave us a perfect fitting curve at all time if only the preset parameters themselves have been used directly as the initial condition of the fitting parameters, and the final condition are consistent to the preset values. For more widespread cases, we can still get the consistent result when the initial conditions are offset within a certain extent. However, the fitting curves somehow show slightly different behavior even they are still acceptable, and the better initial condition(closer to the real value) generally gives the better fitting result. Unacceptable results occurred when we used initial conditions out of a certain range. Therefore, we say that this fitting method is not very sensitive of the initial condition, but the initial condition should be as good as possible to get the best fitting.

Establishing the initial condition model database

For a given PSD, according to the theoretical mode equation (9.4), it is possible to theoretically predict the $1/f$ noise level, the shot noise level and even the GR noise level in terms of the given I_B, A_E and the empirical K_F, A_F, C value. It is impossible, however, to invisibly predict the GR noise centers since the distribution of the GR noise centers vary from device to device due to the different physical conditions of the devices themselves which can not be predicted by any theory. Therefore, we developed a method which on beforehand establishes a model database for finding the best initial condition. Here, the model is actually the synthetic PSDs established by the preset parameters which we exactly know. For one particular test PSD, given the I_B and A_E, all three noise level can be calculated by equation (9.4). Since the noise level can be well predicted this way, it is not necessary to consider other levels. Therefore the three noise level will be identical for all models within one database, the only difference among the models of the database are the different combinations of the GR noise centers. Furthermore, there is another database in which the corresponding preset fitting parameters are stored. Having this kind of model database, we can find the most analogical model for the test PSD within the database by means of classification. It is supposed that the corresponding fitting parameter of the chosen model can be taken as the best estimated initial condition for the non-linear fitting method. Therefore, establishing a proper model database is significant.

Unfortunately, the empirical values of K_F, A_F, C change from device to device, and even the actual value of I_B and A_E can not be trusted to determine the magnitude of the noise level in many cases, which will be shown in the following sections. Thereby, a totally automatical fitting procedure will be introduced in terms of all parameters of equation (9.4) coming from the test PSD itself.

9.2 The Method of Classification

The Nearest Neighbor method is used as the method of classification in order to find the best initial condition. This method can classify the test vectors to the nearest feature by calculating the distances among them. The distance is calculated by:

$$D = \sqrt{\sum_{i=1}^{N} |MODEL - PSD|^2} \tag{9.5}$$

Intuitively, as long as the model database has been properly established, the model with the most similar shape comparing to the test PSD from the database will be chosen by this way. Since we already know the parameters for this particular model, the non-linear fitting will have a good starting point.

9.3 Theoretical Fitting Method

In this section, we will use the non-linear fitting procedure, whose initial conditions are collected from the theoretical models which have been established by the given I_B and A_E, and the empirical value of K_F, A_F and C, to fit the test PSD. We will see that the theoretical method gives good result for some normal devices, but is not the ideal way in the diverse unpredicted situations that often occur which could strongly influence the results, such as much higher or lower noise level than the normal case.

9.3.1 The Construction of The Model Database

From the theory(equation (9.4)) and experimental data in chapter 7 above, the noise level will vary along with the change of the base current I_B and the device area A_E. So it is necessary to have models for every combination of I_B and A_E, as well as the same quantities of GR noise for each of these combinations. In terms of the work of Sanden et al. [2002], K_F does not change much in the emitter area range of the devices we used, therefore, only one proper value has been chosen for all kinds of devices in our case. Base on the devices we used in this thesis and the consequence of many trials for finding the best value for K_F, C, A_F , Table (9.1) lists all the parameter we can theoretically use in this case.

Parameter	value
Base current I_B	$0.2, 0.5, 0.8, 1$ and $2(\mu A)$
Area of device A_E	$0.048, 0.12, 0.25, 0.3, 0.64, 1.28$ and $1.8(\mu m^2)$
K_F	10^{-8}
A_F	2.2
f_c	$10^{-8}, 10^{-5}, 0.01, 0.1, 1, 5, 10, 50, 100, 200$
	$400, 800, 1600, 3200, 6400$
C	7.2×10^{-33}

Table 9.1: Initial condition fitting parameters.

The value of I_B and A_E come from the value we used in experiments. The value of

K_F, A_F and C are experimentally tested to be the best choice for our devices in terms of some empirical numerical value [Sanden et al., 2002]. f_c are the centers of GR noise in frequency scale. The τ, in equation (9.4), will be decided by these f_c with the relationship $\tau = \frac{1}{2\pi f_c}$.

From the data in table (9.1), we can build the model database and the corresponding fitting parameters database. However, for one particular combination of I_B and A_E, the number of models in the model database is determined by the number of possible combinations of parameters, and will lead to millions of combinations if all of the possibilities are considered. To reduce the complexity, we need to reduce the number of GR centers in each model. To achieve a clear distinction between the sum of GR noise and 1/f noise we allow only up to seven different GR noise centers in each model. The database then can be constructed by this set of models, and their corresponding preset fitting parameters are stored in another database.

9.3.2 Experiments and Results

Two devices with $A_E = 0.6\ \mu m^2$ have been chosen to be tested under $I_B = 0.5\ \mu A, I_B = 1.0\ \mu A$ and $I_B = 2.0\ \mu A$ respectively. In terms of these actual value and possible combination of I_B and A_E, three corresponding model databases are established. One of the devices is classified as a normal device and another one is much quieter than the theoretical noise level. They will be referred to as the normal and quiet device respectively during the experiment. We will test our fitting method on these devices below.

Fitting the normal device by using the relevant model database

For the normal device, the test PSD, which is measured under $I_B = 1.0\ \mu A$, using the relevant model database which is built with $I_B = 1.0\ \mu A, A_E = 0.6\ \mu m^2$, is perfectly fitted, as shown in figure 9.1. With this method, the initial condition and the final condition can be recorded during the procedure. Table (9.2) shows the value of these parameters, where $J_B = I_B/A_E$ is the base-current density. The initial and final conditions are close each other which means that we found the right model for this PSD and these parameters can be trusted to explain some physical properties of this device, such as the $1/f$ Level and the number of GR center.

For a more detailed view, figure 9.2(a) shows all the components which compose the fitting curve in figure 9.1. We see that there are two GR noises at $0.5205\ H_Z$ and $1083.6541\ H_Z$ respectively mainly dominate the PSD, and the contributions of $1/f$ noise and other

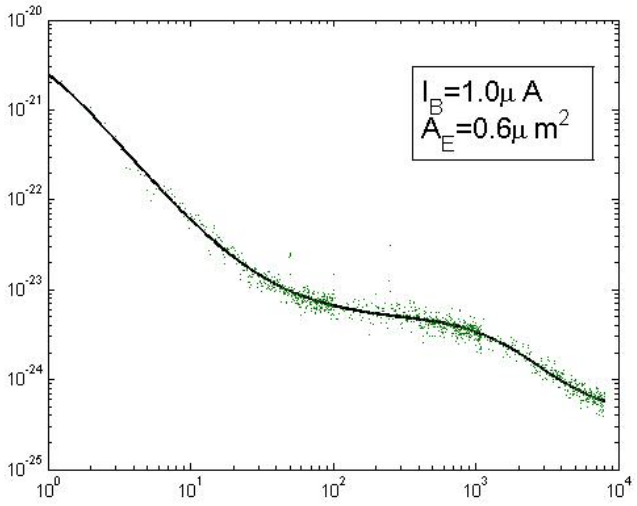


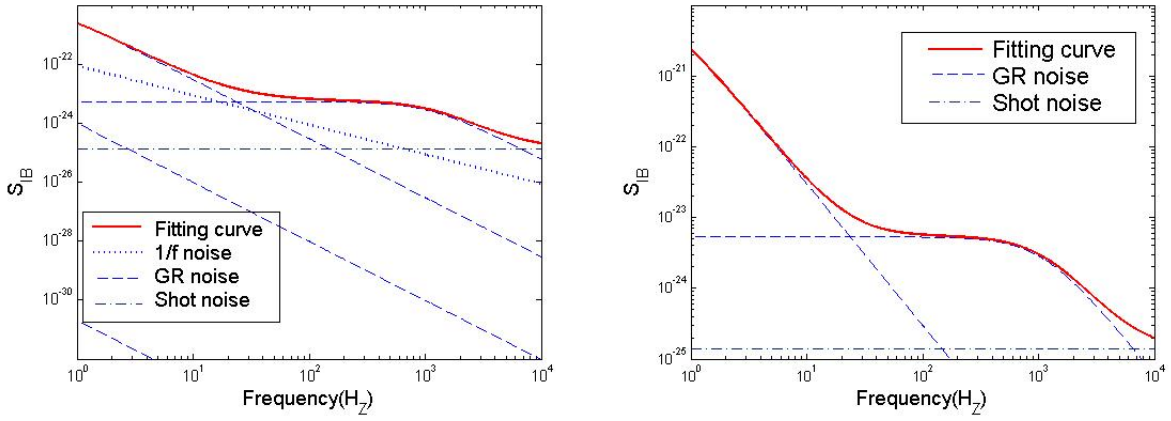
Figure 9.1: Using relevant model database to fit Power Spectrum Density and getting perfect fitting and meaningful parameters.

	Initial	**Final**
K_F	1×10^{-8}	8.83×10^{-9}
J_B	1.67×10^{6}	7.14×10^{5}
C	7.2×10^{-33}	7.1×10^{-32}
GR Noise Center(H_Z)	1×10^{-10}	3.4871×10^{-11}
	0.001	1.6901×10^{-4}
	0.1	0.5205
	800	1083.6541

Table 9.2: Consistent initial and final condition by using the relevant model database.

two GR noises are comparably much smaller. Therefore we can simply remove them and can still get a similar fitting curve as shown in figure 9.2(b). This advantage can efficiently help us to identify the noise origin for each device.

Unfortunately, also using the relevant model, we sometimes got meaningless final condition even if the test PSD can be fitted well. At $I_B = 2.0\ \mu A$ for the same device, this phenomenon shows up in figure 9.3, where a very good fitting result is presented with several meaningless final conditions which are listed in table (9.3). These meaningless parameters present negative final condition include the GR noise level and two of the GR centers. In fact, the non-linear fitting procedure itself does not care about the sign of the value as long as it can get good fitting. It will be shown in section 9.4 that the final parameters are sensitive to the initial parameters during the non-linear fit procedure,



(a) All PSD components (b) The dominant PSD components

Figure 9.2: The LFN PSD components derived by theoretical fitting method.

so a slight difference of the initial condition could lead to meaningless results. On the other hand, it is also possible to get meaningful results by a slight offset of the initial condition which led to the meaningless one.

	Initial	**Final**
K_F	1×10^{-8}	1.84×10^{-8}
J_B	3.33×10^{6}	6.32×10^{6}
C	7.2×10^{-33}	-2.32×10^{-34}
GR Noise Center(H_Z)	0.1	-0.0249
	0.5	0.0728
	1	-0.6826
	5	22.2906
	10	244.8538

Table 9.3: Meaningless Final Condition when $I_B = 2.0\ \mu A$, $A_E = 0.6\ \mu m^2$ by using the relevant model database.

Fitting the normal device by using an irrelevant model database

The so called irrelevant model we mentioned here is the model which has strong deviation in noise level from the theoretical noise level of the test PSD. This deviation lead to the bad initial condition for the non-linear fitting and consequently a bad fitting curve.

For the normal device under $I_B = 2.0\ \mu A$ again, when we change the model to the one which is established by $I_B = 1.0\ \mu A$, $A_E = 0.6\ \mu m^2$, the fitting procedure return

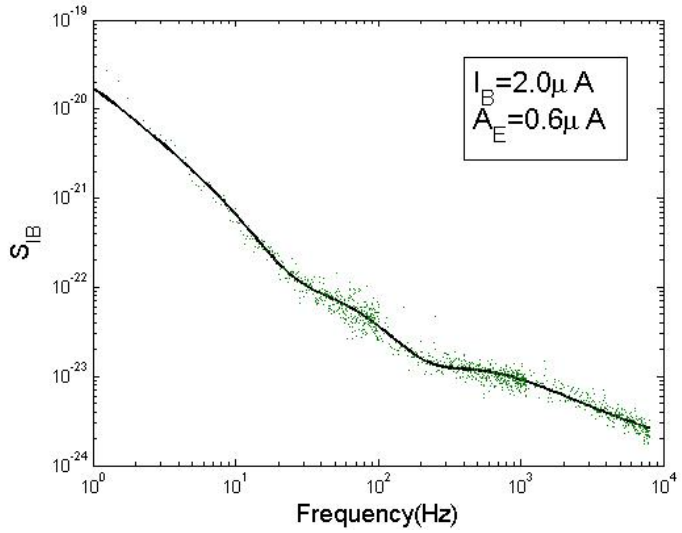


Figure 9.3: Using relevant model database to fit Power Spectrum Density and getting good fitting but meaningless final conditions.

a worse fitting curve, shown in figure 9.4, than the one we got in figure 9.3, since the model gave an initial condition with much lower noise level.

Fitting the quiet device by using a relevant model database

The quiet device shows much lower noise level than the normal one, even they have the same physical fabrication parameter. The fitting result shown in figure 9.5 is obviously unacceptable. It happened since the test PSD is much lower than any models in the model database which has been established by the theoretical parameters a normal device should have. Therefore the Nearest Neighbor Classification justly found the closest model which only include $1/f$ noise and shot noise. However, the test PSD apparently shows strong GR noise around $1000\ H_Z$ which the non-linear fitting procedure can never work it out if there is no GR components in the initial condition. Further more, some improper initial condition lead to totally unfit result which could include negative values. Therefore, the initial condition is a critical issue.

9.3.3 Summary of The Theoretical Fitting Method

To sum up the above arguments, the theoretical fitting method can not be trusted to be the automatical mechanism for fitting PSDs in all cases due to its constraint conditions. The so called constraint conditions of this method are actually its unreliable

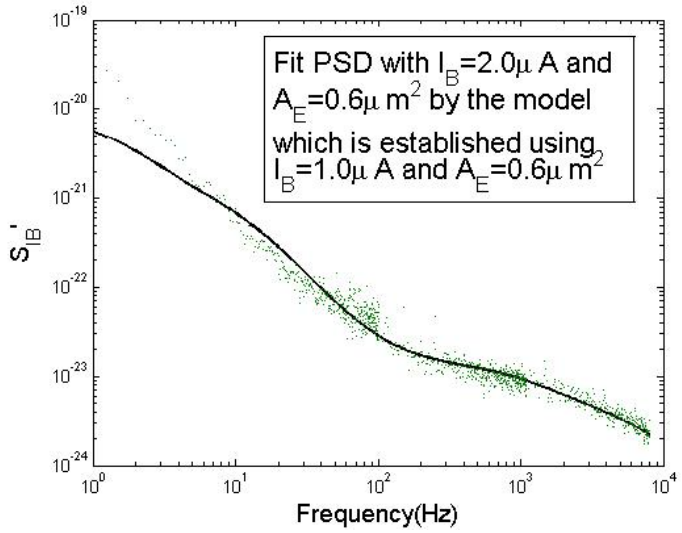


Figure 9.4: Using the improper model database.

assumptions for the initial condition since many interferences make the devices behaviour deviate from the theoretical value. In fact, the model database has given the most possible combinations of the GR noise time constants whatever where the noise level is. So the problem becomes that this method sometimes can not offer the noise level which the test PSD really stay at.

9.4 Practical Fitting Method

In practice, there are many instabilities affecting the Low-Frequency Noise, which cause some devices to present much higher or lower Low-Frequency Noise level than what the theoretical model predicts. This unpredictable noise level is just the reason why the theoretical method is not the ideal way to be the automatical mechanism as we discussed in section 9.3.3. Therefore we develop a practical method which focus on the magnitude of each test PSD reality itself instead of the fixed theoretical parameter value to avoid this disadvantage. However, selecting the true noise level for each particular PSD means that we have to establish the model database every time, which will severely increase the calculation time. Therefore strategies for reducing the calculation time will also be concerned in this section.

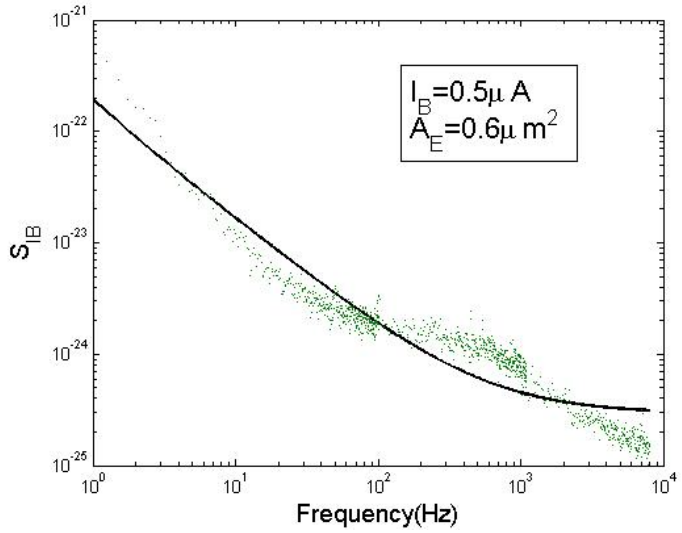


Figure 9.5: Noise Level is too low to be fitted by using the relevant theoretical model database.

9.4.1 Establish The Model Database in Terms of Actual Conditions

The practical fitting method still uses the same parameters(the magnitude of $1/f$ noise, GR noise and shot noise, and the locations of the GR noise centers) to establish the model database as the theoretical method did. The only difference is the way of getting the three noise levels for the model database.

Finding the $1/f$ noise, GR noise level

The magnitude of the PSD is mainly dominated by $1/f$ noise and GR noise at low frequency scale approximately up to $1000\ H_Z$. Therefore we picked up the actual noise level at very low frequency, which is the average value of the first ten points of the test PSD, as the sum of $1/f$ and GR noise level. Then the next question becomes how to assign the proportion between the $1/f$ and the GR noise level since they both contribute within the same frequency range. According to the experimental data of the simulations to many PSD which come from many different type of devices, we eventually found that the 0.7 times the averaged value of the first ten points is the best quantity as the $1/f$ noise level($K_F \times I_B^{A_F}$) for all these tested PSD(but not for every individual PSD), and 20 times $1/f$ noise level has been set to be the value of $C \times J_B^2$ in equation (9.4) which contributes a portion of GR noise level. In fact, the $1/f$ noise level we found to be 0.7, can sometimes be even bigger, and in some cases, the fitting procedure are also

forced to choose a big value in order to get meaningful final condition.

Finding the shot noise level

The shot noise is much more stable than the $1/f$ noise and the GR noise because it only relates to the base current. If we set the shot noise level as a constant, that is, excluding it from the fitting parameters, it could reduce the calculation time in a certain extent. For this reason, we tried to exclude it from the fitting parameter and just set the theoretical value for the PSD model, equation (9.4). Unfortunately, in most of the trials we got the worse fitting even sometimes the fitting curve is still acceptable. For some noisy device, the results become totally unusable. Nevertheless, this does not mean that the shot noise is unstable. In fact, the GR noise and cut-off frequency of the measurement set-up will make the shot noise measurement difficult. So it is necessary to have the shot noise as a fitting parameter since the measured PSD possibly would not show the true shot noise.

As we know, shot noise dominates the higher frequency range of LFN. So it is reasonable to use the mean value of the last ten points as the shot noise level of the initial condition in practice.

The Combinations of GR noise centers

The combinations of the GR noise centers are preset in the same way as the theoretical method since there is no way to predict them in practice either. However we just include up to five GR noise centers for each model within the model database. This is sufficient to yield good results, and it reduces the calculation time to a reasonable amount.

Together, we have all the three noise levels and the combination of the GR noise centers. Then the initial condition of the model is supposed to be able to automatically adapt to any test PSD in this way and overcome the disadvantage of the theoretical method. The model database can consequently be established by these parameters we practically get.

9.4.2 Mechanism for Getting Best fitting

Much work have been done above in order to get a good fitting curve which should have minimum variation from the test PSD as well as a meaningful final condition. Though we think that the best initial condition of the fitting parameter can be found in terms of the most proper model database we got, the non-linear fitting procedure sometimes

returns either bigger variation or meaningless parameters, or both of them. In order to avoid these outcomes, a mechanism have been developed.

In fact, the program has been modified to adjust the initial $1/f$ noise level and the GR noise level simultaneously and reversely with a same preset ratio every time if the fitted curve gives meaningless parameter or big variation[1] from the test PSD and do the same fitting procedure again with the new initial condition. A maximum number of iteration has been set if the procedure can not find a satisfactory fitting curve which should have both meaningful final parameters and small variation, and the best result among these iterations will be picked up as the final parameters even they still could be meaningless or have big variation. Despite of the condition of the final parameters, they can be used anyway as a smooth fitting of the test PSD to pick up its value at any frequency as long as the fitting curve is visually acceptable. This procedure is schematically shown in the appendix in section 11.1.

9.4.3 Experiments and Results

This method has been tested on all the PSD measured at $I_B = 1\ \mu A$, which we got from many different devices. The trial results show that this method can properly fit more than 95% PSD with meaningful parameters and sufficiently small variation, the others are also fitted well but with meaningless parameters. Furthermore, if we properly set the $1/f$ noise level in terms of the actual mean value of the first ten points of PSD, they finally can be perfectly fitted and with meaningful final conditions as well.

We randomly chose a PSD with $A_E = 1.28\ \mu m^2$ and measured at $I_B = 1\ \mu A$ as the test PSD to show a particular example. Using the proportion number 0.7 we have given in section 9.4.1, the test PSD is perfectly fitted. Figure 9.6 and table (9.4) show the fitting result and the corresponding parameters respectively. We see that the PSD shown in figure 9.6 is a little bit steeper than $1/f$. As a good interpretation to this phenomenon, the final parameters in the table (9.4) show that all the GR noise centers, frequency f_c, centralize at very low frequency range(all below $7\ H_Z$ in the table), moreover we could also extrapolate that there might be one very strong GR noise located there. The δ in the figure is the variation between the fitting curve and the test PSD as defined by equation (7.6)[2] and the variation at this extent is totally acceptable.

[1] The value of variation has been experimentally found that it normally should be smaller than 0.13 for an acceptable fitting curve.

[2] We can also use the conventional way, Euclidean distance, to calculate the variation between two curves, i.e. $\delta = \sqrt{\sum (A-B)^2}$, where A and B are vectors.

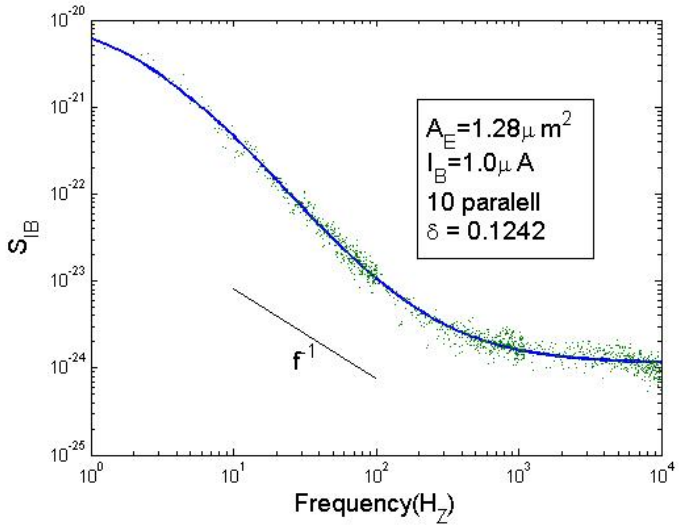


Figure 9.6: Fitting result by practical method.

	Initial	**Final**
K_F	6.61×10^{-10}	8.32×10^{-10}
J_B	7.81×10^{5}	2.73×10^{6}
C	9.62×10^{-31}	4.52×10^{-33}
GR Noise Center(H_Z)	1×10^{-8}	3.63×10^{-8}
	1×10^{-5}	1.6243
	0.01	0.0173
	0.1	1.6249
	5	6.2922

Table 9.4: Practical fitting parameters.

The K_F, C and J_B in table (9.4) are calculated by the $1/f$ noise level, GR noise level and shot noise level respectively in terms of equation (9.4). We assume the A_F in the $1/f$ noise level $K_F \times I_B^{A_F}$ is equal to 2 in terms of the conclusion in section 7.2, and this value will also be used in all the experiments.

9.4.4 Summary of The Practical Fitting Method

To sum up, the practical fitting procedure has proven a strong ability to automatically fit any PSD. However the advantage has its price. Compared to the theoretical fitting, this method severely increases the calculation time since it has to build the model database every time, which normally takes a longer time than the fitting procedure for each

particular test PSD. The theoretical fitting method always has the model database built in advance for all the PSD which come from the same I_B and A_E. Even though, there are still some ways to reduce this disadvantage, for instance, choosing a more efficient distribution and number of GR noise centers to get a robust and compact model database.

9.5 Discussions Based On the Practical Fitting Method

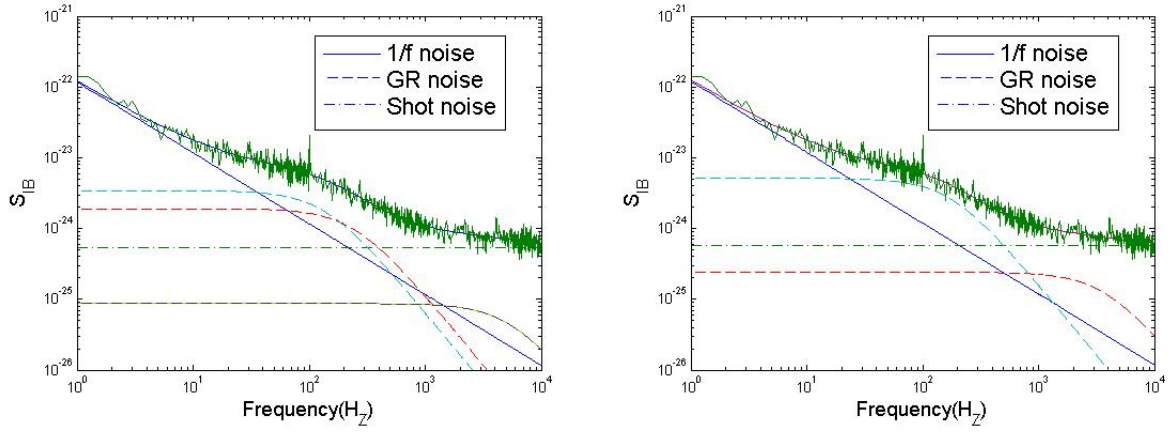
One important reason that the practical fitting method is said to be a robust fitting method is that this method can have the PSD fitted with different initial conditions to a certain extent. Furthermore, these different initial conditions lead to different final conditions by all of which the PSDs can be perfectly fitted. That is, in spite of different parameters among these final conditions, they can achieve a similarly good fitting curve anyway. Then how about the empirical value K_F and C? Do they change significantly among these final conditions for one particular PSD? Furthermore, do they have I_B or A_E dependence?

9.5.1 The Variations of Fitting Parameters For One Particular PSD

Here, we focus on the behavior of the fitting parameters, K_F, J_B and C, on one particular PSD. By changing the proportion of the $1/f$ noise level compared to the actual mean value of the first ten points of PSD, we can easily get different initial fitting parameters and further the different final conditions. The proportional numbers are actually chosen from 0.5 to 3. Within this broad range, all the fitting result present identical good fitting curves even they do not contain the same parameters.

Figure 9.7 shows two fitting curves[3] with five and two GR noise centers respectively. We see that the $1/f$ noise dominates the lowest frequency range($1\ H_Z$ - $10\ H_Z$) and GR noise take over the rest. In addition, for all the fitting curves with other parameters, the $1/f$ noise level always keeps its position as shown in this figure. Moreover, for one particular PSD, the J_B should be similar to each other. Therefore, in terms of equation (9.1), K_F should be stable. As we have acclaimed in section 9.1.1, C is supposed to be a constant for one particular device in the noise model. Since it directly contribute the GR noise level, its value then is related to the conditions of GR noise, the number and the positions of GR noise centers. As the fact, the positions of GR noise traps physically ex-

[3]Using the proportional number 1 and 3 respectively to get the initial conditions



(a) Fitting with five GR noise centers

(b) Fitting with two GR noise centers

Figure 9.7: Many different final fitting parameters shows the similar good fitting curve for one particular PSD which is measured under $I_B = 1\ \mu A$, $A_E = 0.64\ \mu m^2$.

ist, that is, all of these fitting parameters should be found to present these traps anyway as long as the good fitting curve is achieved. So it turns out that the main matter which affects C is the number of GR noise centers. In an another word, the point is how many GR noise centers will be used in this fitting procedure to fit one GR signature which is identified by eyes.

Table (9.5) and table (9.6) show the initial fitting parameters and final fitting parameters respectively for the cases with different initial proportional number, where the initial parameters are found by classification from the model databases. And each of the model databases are established by the corresponding $1/f$ noise proportional number.

We can see that all the final parameters K_F, J_B and C are statistically identical to each other over the range that the proportional number are from 0.5 to 0.9, 1 and 1.5, 2 and 2.5 respectively. However, they change significantly among these ranges especially when the number of GR noise centers is different. Together, we can tell that the fitting method can fit the PSD with identical final fitting parameters if only the initial condition has been chosen in a similar manner. These experimental results are not exactly consistent to what we suggested above due to the fact that the shot noise is normally difficult to measure due to the influence from cut-off frequency or GR noise. The parameter J_B is apparently affected by the different final GR noise condition so that it further influences K_F by equation (9.1).

Interestingly, the case with proportional number 3(in table (9.6)) shows two GR noise

Proportion	$K_F(10^{-10})$	$J_B(10^5)$	$C(10^{-33})$	$f_c(H_Z)$				
0.5	2.15	5.26	1.55	1	5	50	100	200
0.6	1.11	5.26	3.64	1	10	50	100	200
0.7	1.44	5.26	3.91	1	50	100	200	400
0.8	8.89	5.26	1.16	1	50	100	200	400
0.9	11.1	5.26	1.2	1	100	200	400	800
1	1.22	5.26	8.5	10	200	400	800	1600
1.5	4.23	5.26	6.5	100	400	800	1600	3200
2	1.77	5.26	21.9	200	800	3200	6400	
2.5	1.99	5.26	29.8	800	1600	3200	6400	
3	0.61	5.26	107	3200	6400			

Table 9.5: The various initial fitting parameters.

Proportion	$K_F(10^{-11})$	$J_B(10^6)$	$C(10^{-33})$	$f_c(H_Z)$				
0.5	3.27	1.48	0.93	145	145	335	3510	3510
0.6	2.8	1.46	1.04	91	254	296	4049	4059
0.7	3.27	1.48	0.93	146	146	331	3496	3496
0.8	3.25	1.48	0.91	131	174	285	3286	3288
0.9	3.27	1.48	0.93	145	145	333	3503	3503
1	4.15	1.31	1.73	137	250	5451	5451	5451
1.5	4.15	1.31	1.73	137	249	5448	5448	5448
2	6.22	1.08	4.93	181	5990	9056	9056	
2.5	6.24	1.08	4.94	181	6127	8943	8943	
3	3.61	1.42	2.84	178	3813			

Table 9.6: The various final fitting parameters.

centers which just locate at the position we can visually predict as shown in figure 9.7(b). Then it is considered as the most efficient parameters for the PSD model, equation (9.4).

To sum up, based on the practical fitting method, the fitting parameters K_F, J_B and C are statistically stable for one particular PSD to a certain initial condition extent.

9.5.2 Identify the GR noise component

By using the practical fitting procedure, we want to show if the PSD measured at different base currents for the same device will show a similar GR noise component. In another word, if fluctuations of the base current will introduce fluctuations of the GR centers.

Figure 9.8 shows two PSDs which are measured under different base current from

the same device which is fabricated by 10 parallel transistors. Table (9.7) lists the final conditions of this two fitting curve. Apparently, these two PSD have the similar GR noise centers which means that the increased base current does not activate new GR traps in this device.

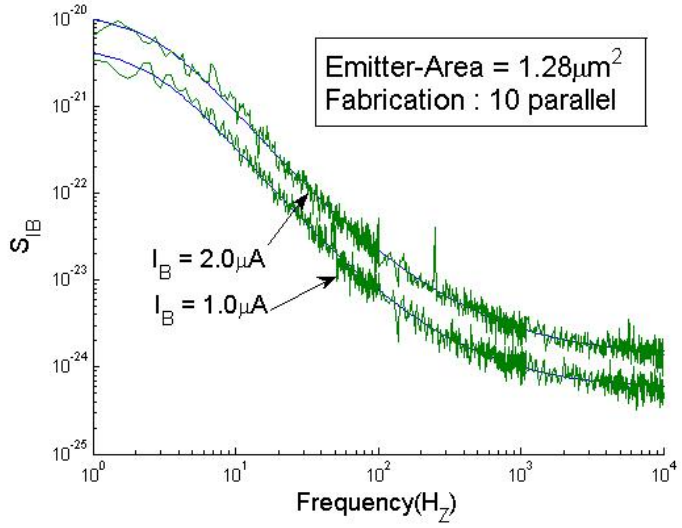


Figure 9.8: For one parallel device, its PSD shows similar GR noise component under different base current.

$I_B(\mu A)$	$K_F(10^{-10})$	$J_B(10^6)$	$C(10^{-32})$	$f_c(H_Z)$
1	1.32	1.38	3.71	2.76
2	0.67	3.43	1.46	2.87

Table 9.7: The two PSD with different I_B conditions present the same GR noise centers in a parallel device.

The situation becomes different when we do the same experiment on a single device with the same emitter area, which is proven to be relatively unstable compared to the parallel device(section 7.4.2). In table (9.8), it is obvious that the higher base current activates more GR noise traps.

9.5.3 The I_B Dependence Of Noise Model parameters

By definition, both K_F and C are constant for one particular device. So only J_B should have I_B dependence, $J_B \propto I_B$. In this section, we will see if the practical fitting method can find this relationship.

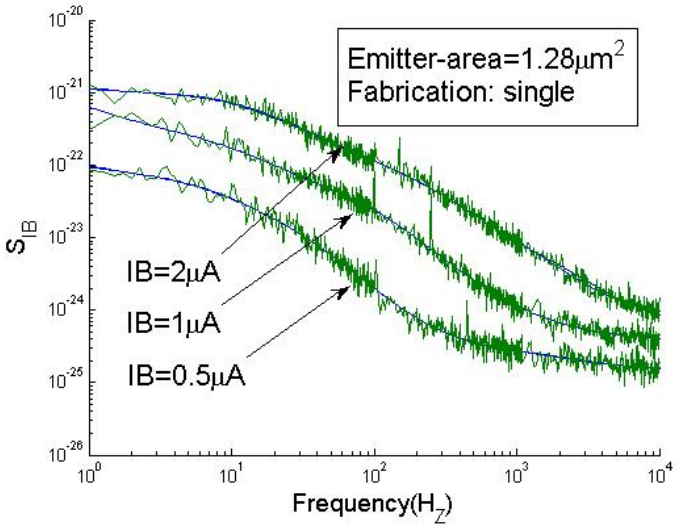


Figure 9.9: For one single device, its PSDs show different GR noise component under different base current.

$I_B(\mu A)$	$K_F(10^{-10})$	$J_B(10^5)$	$C(10^{-32})$	$f_c(H_Z)$				
0.5	3.36	3.9	1.26	11.5	23.8			
1	3.41	8.84	1.43	11.3	69.5	117		
2	1.18	18.8	1.38	13.2	91	93.5	314	1191

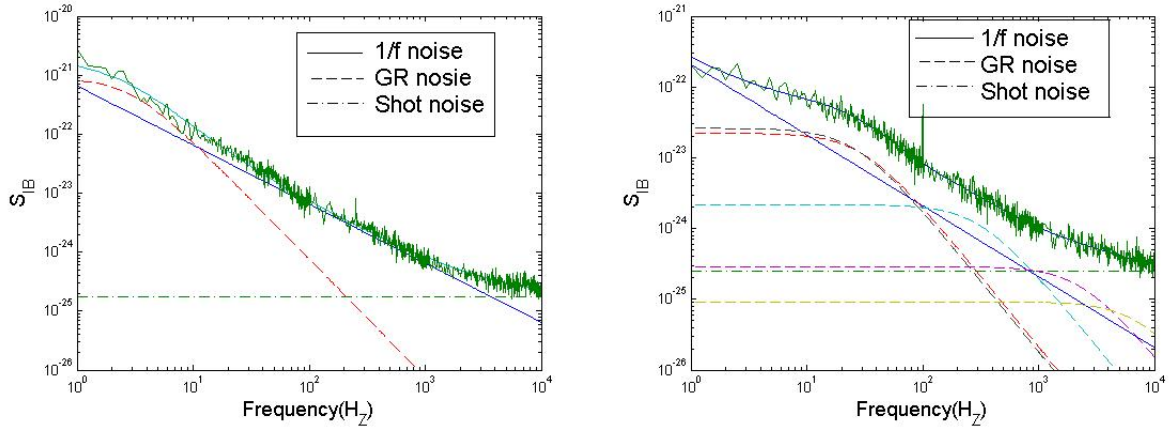
Table 9.8: The three PSD with different I_B conditions present variant GR noise centers in a single device.

From foregoing discussion, the fitting parameters K_F, J_B and C are related to the $1/f$ noise, shot noise and GR noise respectively. In section 7.2, we know that the PSD of $1/f$ noise and GR noise has I_B dependence, $S_{I_B} \propto I_B^2$, which are consistent with the expression of noise model equation (9.4) where all the three noise sources are related to I_B. Separately, the $1/f$ noise is proportional to $I_B^{A_F}$; the GR noise is proportional to J_B^2 and the shot noise is directly related to I_B.

The practical fitting method, however, have K_F and C as the fitting parameter, that is, they will be fitted anyway. In spite of this, we prefer that they do not have any I_B dependence, and all the noise source level variations are contributed by the fitted I_B as the theory indicates above.

Here, we discovered two kinds of device, the LFN of one of them is mainly dominated by $1/f$ noise and GR noise dominate the LFN of another one, as shown in figure 9.10.

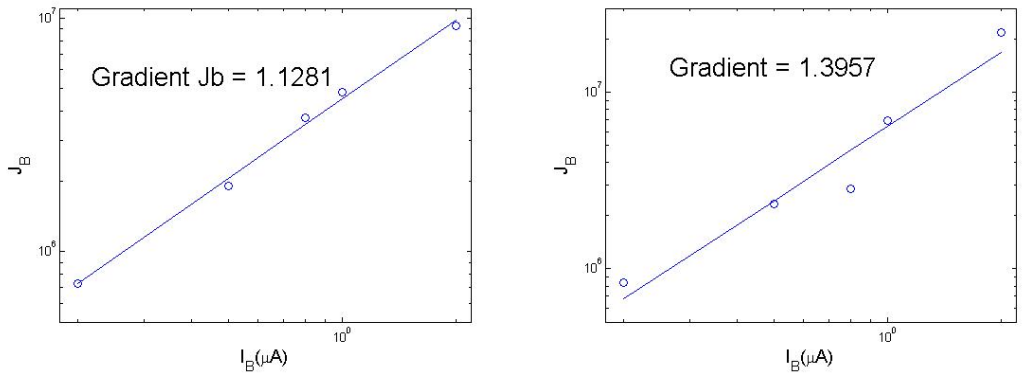
By fitting their PSDs, which are measured at $I_B = 0.2, 0.5, 0.8, 1$ and $2\ \mu A$, we got the



(a) The $1/f$ noise dominate the PSD (b) The GR noise dominate the PSD

Figure 9.10: Two types of device with $1/f$ noise dominant and GR noise dominant respectively.

plots with K_F, J_B and C versus I_B respectively.

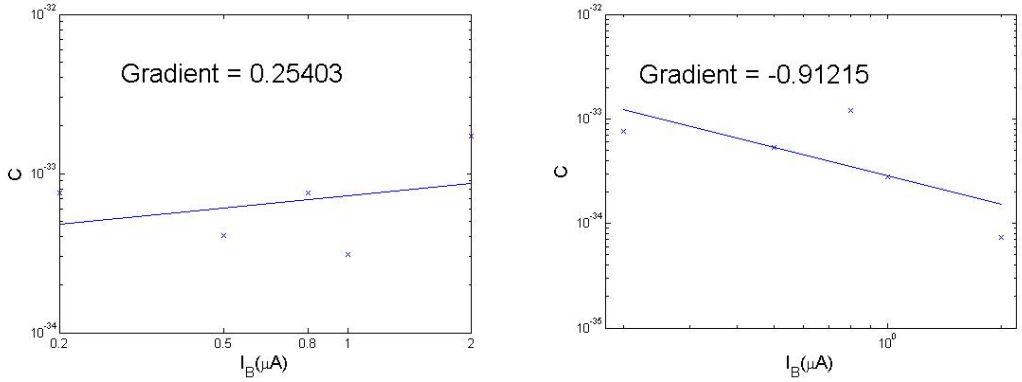


(a) J_B under $1/f$ noise dominant (b) J_B under GR noise dominant

Figure 9.11: J_B variation with $1/f$ noise dominant and GR noise dominant respectively.

In figure 9.11, we see that J_B always approximately has the expected I_B dependence, $J_B \propto I_B$, despite which noise source dominates. Again, because of the large possibility that the shot noise could be affected by the cut-off frequency and GR noise, the fitted J_B therefore shows slightly offset from the expected relationship in both cases, especially when GR noise dominates the PSD, the influence is even more obvious.

As shown in section 9.5.1, C and K_F vary due to the situation of GR noise. In fig-



(a) C under $1/f$ noise dominant (b) C under GR noise dominant

Figure 9.12: C variation with $1/f$ noise dominant and GR noise dominant respectively.

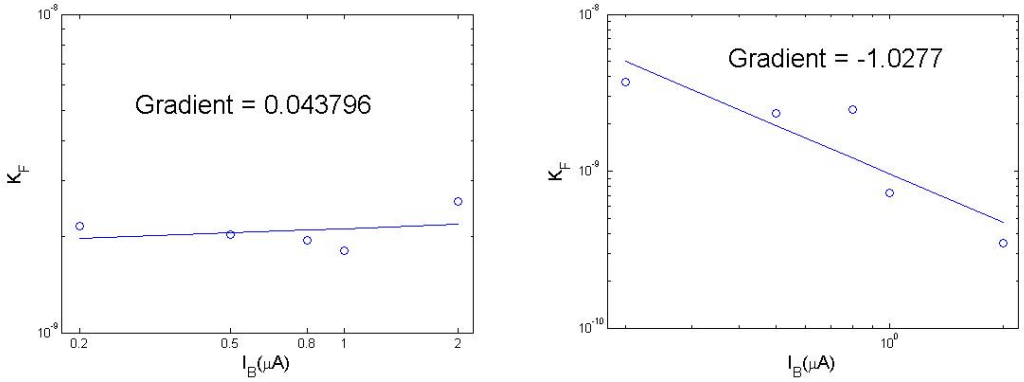
ure 9.12 and figure 9.13, we see that C and K_F are stable as we expect in the case in which the $1/f$ noise dominates the spectra. When GR noise dominates, however, C and K_F drastically show the I_B dependence, which occurred due to the fitted J_B. From figure 9.11(b), we know that the fitted shot noise has been severely affected by the GR noises since J_B is proportional to $I_B^{1.4}$. Once J_B is affected, C and K_F will be consequently affected by the relationships shown in equation (9.2) and equation (9.1) respectively. And the gradients of them exactly show these relations. In terms of these discussions, we can tell that the practical method could be engaged to identify if the GR noise sources influence the measurement of the shot noise.

9.5.4 The Emitter Geometrical Dependence of Fitting Parameters

Following the same idea in section 9.5.3, we will investigate the emitter geometrical(emitter area and perimeter) dependence of these fitting parameters. Under the same base current $I_B = 1\,\mu A$, We measured many devices with different emitter areas($0.048, 0.12, 0.25, 0.3, 1.28$ and $1.8\,\mu m^2$) and different emitter perimeters($0.88, 1.6, 2.5, 3.3, 4.8, 6.15, 8.4, 12.4$ and 16.8 μm) respectively. By fitting them, we got series of fitting parameters discussed below.

The Parameter K_F

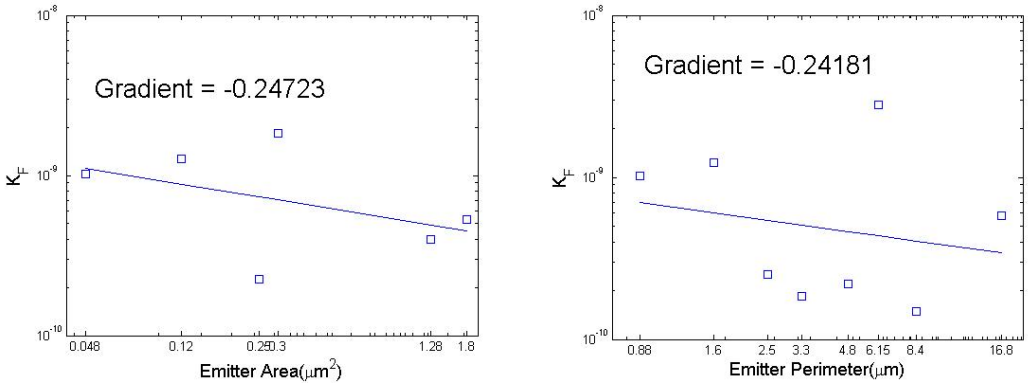
K_F is normally found to be inversely proportional to A_E [Sanden and Deen, 2002]. Figure 9.14(a) shows K_F which we calculated from the practical fitting method we developed in section 9.4. The gradient shows that the K_F is proportional to $A_E^{-0.2}$ instead of



(a) K_F under $1/f$ noise dominant (b) K_F under GR noise dominant

Figure 9.13: K_F variation with $1/f$ noise dominant and GR noise dominant respectively.

the commonly accepted, $K_F \propto A_E^{-1}$. This happened due a big part to a similar reason when we discussed the A_E dependence of LFN in section 7.3. Subsequently, this result is almost consistent to the result of Emitter-Area dependence which we got previously in figure 7.6(a). The K_F versus Emitter-Perimeter, shown in figure 9.14(b), shows the similar trend as the perimeter does. Since the emitter-perimeter is also a essential property of device like emitter-area, it is not surprised that K_F also has emitter-perimeter dependence.

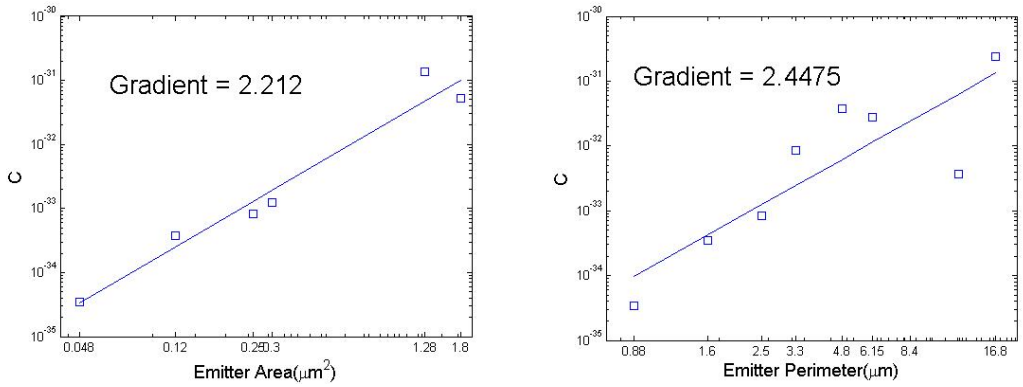


(a) K_F versus A_E (b) K_F versus Emitter-Perimeter

Figure 9.14: The emitter geometrical dependence of K_F.

The Parameter C

Knowing that C is related to the GR noise level of the devices and both the bigger emitter area and longer emitter perimeter could introduce more GR traps, therefore C should have both these two kinds of dependence. In our fitting results, C increase as the increasing of both the area and perimeter as shown in figure 9.15. From these result, we can argue that the larger device has bigger C.



(a) C versus A_E (b) C versus Emitter-Perimeter

Figure 9.15: The emitter geometrical dependence of C.

The Parameter J_B

Under the same base current(I_B), J_B should be proportional to the emitter area(A_E) in terms of the relationship $J_B = \frac{I_B}{A_E}$. Figure 9.16 shows this relationship.

9.6 Summary of The Fitting Method

To sum up, the theoretical fitting method can only be used for normal devices, therefore it is not useful in practice due to the various interferences. However, the practical fitting method has been proven to be an efficient and robust method to not only automatically fit any power spectrum density, but also dramatically show all of the noise sources of the transistors. We simply summarize some advantages and disadvantages of the practical fitting method below:

- Advantages

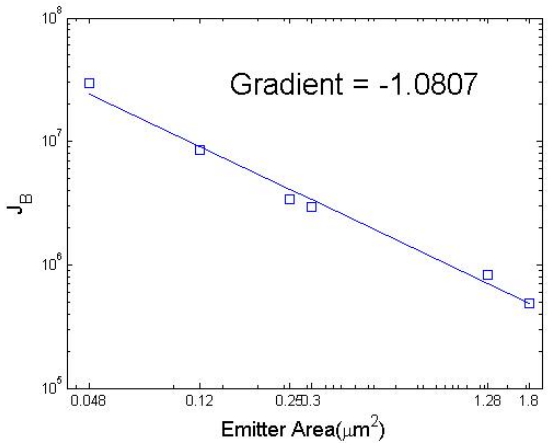


Figure 9.16: J_B versus A_E.

  - Automatic fitting;
  - Identifying each noise component, $1/f$ noise, GR noises and shot noise;
  - Lending itself to investigating the parameters in the noise model.

- Disadvantages

  - Relatively long calculation time;
  - Sometimes giving meaningless fitting parameters;
  - Can not guarantee finding the most efficient fitting parameters as shown in table (9.6).

Whatever happens, this practical fitting method shows a strong potential that it can be further developed to overcome these disadvantages, and to be employed by other applications.

Chapter 10

Summary and Conclusions

From a device perspective, the Low-frequency noise in the fourth generation $375\ GH_Z$ SiGe HBTs was measured in this work. Investigating the power spectrum density characteristics of LFN is the most important issue we have worked on in this thesis, especially by employing the automatic fitting procedure.

Some elementary semiconductor concepts were introduced in chapter 2. From intrinsic semiconductor to extrinsic semiconductor, and from diode to bipolar junction transistor, we systematically illustrated the operation mechanism of semiconductor transistors. In the meantime, each current component in the transistor under active mode has been represented schematically.

The following chapter then introduced the advantages and the fabrication of SiGe HBT whose primary difference from BJT in fabrication technology is that the base is unevenly doped by introducing Ge into Si for achieving higher speed.

In chapter 4, we investigated the origin of the thermal noise, shot noise, generation-recombination noise and $1/f$ noise respectively, especially the last three noises which are supposed to be the dominating contributors of the LFN spectrum. Consequently, we got the LFN power spectrum density model in terms of them.

We introduced some spectrum estimators(Periodogram and Multitaper) in chapter 5. The significance that we represent the LFN by using PSD was illustrated as well.

In the experimental part, we firstly established a LFN measurement system. This system was built based on the purpose that constructing a automatic, accurate and efficient measurement circumstance without any interferences.

By using this measurement system, we measured several sequences of PSD, described in chapter 7. By analyzing these PSDs, we therefore can investigate a variety of properties of LFN of SiGe HBTs, such as the dominant noise source, the current and

geometrical scaling dependence, the variation related to the fabrication and emitter area of devices, etc. These highly scaled devices show unpredictable LFN behaviour, and therefore are very difficult to model using existing theory.

We put our special efforts in developing the fitting procedure in chapter 9. This procedure realizes the expectation to automatically fit any power spectrum density of LFN. Impressively, it shows all of the noise components($1/f$ noise, GR noise and shot noise) in its final fitting parameters as shown in many figures of this work and visually help us understand the components of the PSD. Furthermore, it also shows its unique ability to estimate the parameters of the noise model so that we can investigate the principle of them compared to previous works. However, this method is not perfect yet. For instance, it sometimes gives meaningless final parameter(negative value) and big variation. The initial condition sensitivity, which generates many different final parameter combinations, is another disadvantage of it. In any case, this method gives us a new way to fit PSD and can be further developed to be engaged to other relevant works.

Future Work

The study about the new generation SiGe HBTs is far from finished yet. Due to the limited number of device and tense time schedule, we can just do this far. The automatic fitting procedure, which has been introduced in this work, shows some interesting points for studying the LFN and is therefore worth to be further perfected such as using more reasonable GR noise distribution, different noise model function for establishing the model database and looking into the non-linear fitting procedure itself, etc.

Chapter 11

Appendix

11.1 The Flow Chart of The Automatic Fitting Method

The flow chart in figure 11.1 shows the core work mechanism of the fitting procedure.

In the initialization part, all the initial information of this procedure require is the test power spectrum density itself and the device emitter-area, from which we can approximately estimate the $1/f$ noise level, GR noise level and shot noise level. The model database then can be established in terms of these noise level and various combinations of GR noise centers.

The non-linear fitting part contains some control elements to assure the operation of this procedure. Essentially, we need the fitting result has both meaningful fitting parameters and as small as possible variation from the original PSD. In order to achieve this purpose, we have many iterations in each of which the $1/f$ noise and GR noise level have been slightly adjusted reversely because of the fact that only one noise source dominate in a certain frequency range.

Inevitably, this procedure can not find any fitting curve even a one with meaningless fitting parameters in some extreme cases, such as the very noisy device where the cut-off frequency occurred in the fitting frequency range.

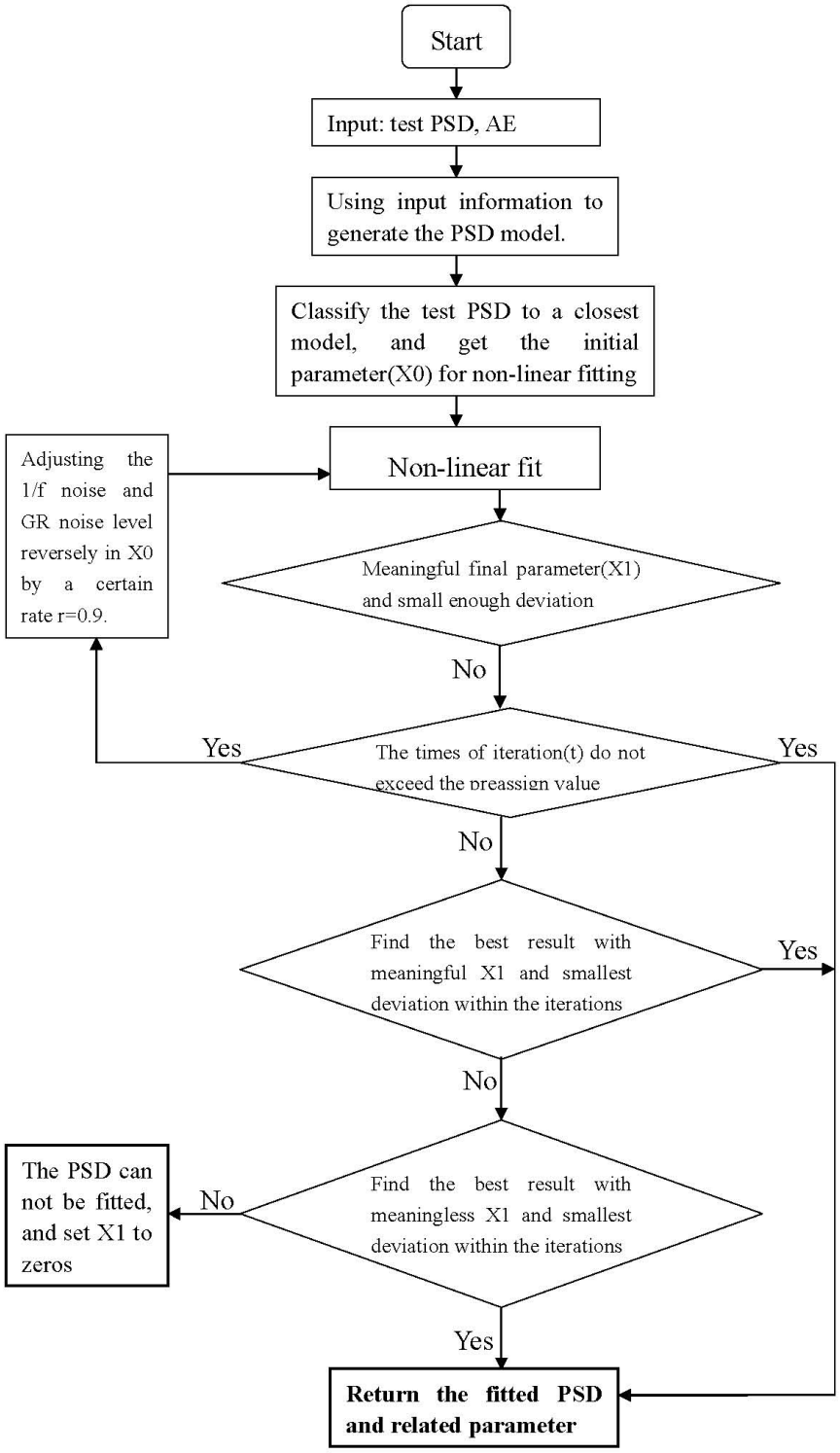


Figure 11.1: Practical fitting method flow chart.

Bibliography

Adel S. Sedra, K. C. S. (1998). *Microelectronic Circuits(Fourth Edition)*. Oxford University Press, Inc.

A.Gruhle and C.Mahner (1997). Low $1/f$ noise sige hbt's with application to low phase noise microwave oscillators. *Electronic Letters*, 33:2050–2052.

Chen, X. Y., Johansen, J. A., Salm, C., and van Rheenen, A. D. (2001). On low-frequency noise of polycrystalline Ge_xSi_{1-x} for sub-micron CMOS technologies. *Solid-State Electronics*, 45(11):1967–1971.

Cressler, J. D. and Niu, G. (2002). *Silicon-Germanium Heterojunction Bipolar Transistors*. Artech House, Inc, 685 Canton Street Norwood, MA 02062.

D'yakonova, N. V., Levinshtein, M. E., and Rumyantsev, S. L. (1991). Temperature-dependence of the low-frequency noise in structurally perfect GaAs and after destructive compression. *Soviet Physics Semiconductors*, 25(2):217–218.

F.Capasso (1987). Band-gap engineering: from physics and materials to new semiconductor devices. *Science*, 235:172–176.

Hanssen, A. (2003). *Introduction to STATISTICAL SIGNAL PROCESSING AND DATA ANALYSIS*. Physics Department, University of Tromsø.

H.A.W.Markus and T.G.M.Kleinpenning (1995). Low-frequency noise in polysilicon emitter bipolar transistors. *IEEE Transactions on Electron Devices*, 42(4):720–727.

Hooge, F. N. (1969). $1/f$ noise is no surface effect. *Physics Letters A*, 29(3):139–140.

Hooge, F. N. (1994). $1/f$ noise sources. *IEEE Transactions on Electron Devices*, 41(11):1926–1935.

Hooge, F. N. (2003). On the additivity of generation-recombination spectra. Part 2: $1/f$ noise. *Physica B: Condensed Matter*, 336(3-4):236–251.

Hooge, F. N. and Vandamme, L. K. J. (1978). Lattice scattering causes $1/f$ noise. *Physics Letters A*, 66A(4):315–316.

Iyer, S. S. and et al. (1987). Silicon-germanium base heterojunction bipolar transistors by molecular beam epitaxy. *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pages 874–876.

Jin, Z. (2004). *Low-Frequency Noise in SiGe BiCMOS Technology*. PhD thesis, Shool of Electrical and Computer Engineering Georgia Institute of Technology, Georgia, USA.

Jin, Z., Cressler, J. D., Niu, G., and Joseph, A. J. (2002). Low-frequency noise variation in scaled SiGe HBTs. In *Proceedings of IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pages 224–227, Monterey, CA, USA.

Jin, Z., Cressler, J. D., Niu, G. F., and Joseph, A. J. (2003). Impact of geometrical scaling on low-frequency noise in SiGe HBTs. *IEEE Transactions on Electron Devices*, 50:676–682.

Johansen, J. A., Jin, Z., Cressler, J. D., and Joseph, A. J. (2003). Geometry-dependent low-frequency noise variations in 120 GHz f_T SiGe HBTs. In Ponchak, G. E., editor, *Digest of Papers, Topical Meeting on Silicon Monolitic Integrated Circuits in RF Systems (SiRF)*, pages 57–59, Grainau, Germany.

Jones, B. K. (1994). Low-frequency noise spectroscopy. *IEEE Transactions on Electron Devices*, 41(11):2188–2197.

Joseph, A., Coolbaugh, D., Zierak, M., Wuthrich, R., Geiss, P., He, Z., Liu, X., Orner, B., Johnson, J., Freeman, G., Ahlgren, D., Jagannathan, B., Lanzerotti, L., Ramachandran, V., Malinowski, J., Chen, H., Chu, J., Gray, P., Johnson, R., Dunn, J., Subbanna, S., Schonenberg, K., Harame, D., Groves, R., Watson, K., Jadus, D., Meghelli, M., and Rylyakov, A. (2001). A 0.18 μm BiCMOS technology featuring 120/100 GHz (f_T/f_{max}) HBT and ASIC-compatible CMOS using copper interconnect. In *Proceedings of IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pages 143–146, Minneapolis, MN, USA.

Kasap, S. O. (1997). *Principles of Electrical Engineering Materials and Devices*. Tom Casson.

Kingston, R. H. and McWhorter, A. L. (1956). Relaxation time of surface states on germanium. *Physical Review*, 103(3):534–540.

Kroemer, H. (1957). Theory of a wide-gap emitter for transistors. *Proc. IRE*, 45(11):1535–1537.

L.S.Vempati, J.D.Cressler, F.A.Babcock, R.C.Jaeger, and D.L.Harame (1996). Low-frequency noise in uhv/cvd epitaxial si and sige bipolar transistors. *IEEE Journal of Solid-State Circuits*, 31(10):1458–1467.

McWhorter, A. L. (1955). Surface traps and 1-f noise in germanium. *Physical Review*, 98:1191–1192.

Niu, G. (2005). Noise in sige hbt rf technology:physics, modeling, and circuit implications. *Proceedings of the IEEE*, 93(9):1583–1597.

P.-F.Lu (1987). Low-frequency noise in self-aligned bipolar transistors. *J.App.Phys.*, 62:1335–1339.

Pierret, R. F. (1988). *Volume I Semiconductor Fundamentals(Second Edition)*. Addison-Wesley publishing Company.

Plana, R., Escotte, L., Roux, J. P., Graffeuil, J., Gruhle, A., and Kibbel, H. (1995). $1/f$ noise in self-aligned Si/SiGe heterojunction bipolar transistor. *IEEE Electron Device Letters*, 16(2):58–60.

R. Brederlow, W. Weber, C. D. D. S.-L. and Thewes, R. (2001). Low-frequency noise of integrated poly-silicon resistors. *IEEE Transactions on Electron Devices*, 48(6):1180–1187.

Riedel, K, S. and Sidorenko, A. (1995). Minimum bias multiple taper spectral estimation. *IEEE Transactions on Signal Processing*, (43):188–195.

S. P. O. Bruce, L. K. J. V. and Rydberg, A. (1999). Measurement of low-frequency base and collector current noise and coherence in sige heterojunction bipolar transistors using transimpedance amplifiers. *IEEE Transactions on Electron Devices*, 46(5):993–1000.

Sanden, M. and Deen, M. (2002). Low-frequency noise in advanced si-based bipolar transistors and circuits. *American Scientific Publishers*, pages 235–247.

Sanden, M., Marinov, O., Deen, M. J., and Östling, M. M. (2002). A new model for the low-frequency noise and the noise level variation in polysilicon emitter BJTs. *IEEE Transactions on Electron Devices*, 49(3):514–520.

Schottky, W. (1918). Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern. *Annalen der Physik*, 57:541–567.

Slepian, D. (1978). Prolate spheroidal wave functions, fourier analysis and uncertainty; v: The discrete case. *Bell Syst. Tech.*, (5):1371–1429.

Surdin, M. (1939). Fluctuation of thermionic current and the "flicker" effect. *Journal de Physique et le Radium*, 10:188–189.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096.

van der Ziel, A. (1955). Theory of shot noise in junction diodes and junction transistors. *Proc.IRE*, 43:1639–1646.

van der Ziel, A. (1979). *Advanced Electronic Electron Physics*. Academic Press, New York.

Vandamme, L. K. J. (1989). Bulk and surface $1/f$ noise. *IEEE Transactions on Electron Devices*, 36(5).

W.Neudeck, G. and Pierret, R. F. (1989). *Volume III Semiconductor Fundamentals(Second Edition)*. Addison-Wesley publishing Company.

Ziel, A. V. D. (1989). Formulation of surface $1/f$ noise processes in bipolar junction transistor and in $p - n$ diodes in hooge-type form. *Solid-State Electronics*, 32(1):91–93.