2016

Advanced numerical modeling of semiconductor material properties and their device performances

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

ADVANCED NUMERICAL MODELING OF SEMICONDUCTOR MATERIAL PROPERTIES AND THEIR DEVICE PERFORMANCES

by

HANQING WEN

B.S., Peking University, 2011
M.S., Boston University, 2014

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader

Enrico Bellotti, PhD
Professor of Electrical and Computer Engineering
Professor of Materials Science and Engineering

Second Reader

Roberto Paiella, PhD
Professor of Electrical and Computer Engineering
Professor of Materials Science and Engineering

Third Reader

Luca Dal Negro, PhD
Associate Professor of Electrical and Computer Engineering
Associate Professor of Materials Science and Engineering

Fourth Reader

Ronald W. Knepper, PhD
Professor of Electrical and Computer Engineering

学而不思则罔，思而不学则殆

孔子

Acknowledgments

First, I want to express my deepest thanks to my advisor and mentor, Professor Enrico Bellotti, who have guided me through the art of numerical simulation, and given me the tools to explore this wonderful area.

Special thanks go to my defense committee members: Prof. Ronald Knepper, Prof. Roberto Paiella, Prof. Luca Dal Negro and Prof. Malay Mazumder for their time and efforts of reviewing and improving this dissertation. I would also like to give my sincerest thanks to Prof. Francesco Bertazzi of Politecnico di Torino, Prof. Monuko du Plessis of University of Pretoria and Dr. Roger DeWames for their constant help and invaluable advise on developing the numerical models in this work.

I would like to acknowledge the funding agencies of this work: University of Pretoria in Republic of South Africa for the work of SiGe LED; U.S. Army Research Laboratory through the Collaborative Research Alliance for the work of IR material properties and National Science Foundation for the 3D Monte-Carlo modeling.

Hearty thanks also give to my friends, colleagues, and former members of the Computational Electronics Group: Y. Zheng, W. He, B. Tai, Dr. Schuster, Dr. Pinkie, Dr. Matsubara, Dr. Sengupta, Ms. Shishehchi, Mr. Kyrtsos, Mr. Glasmann and Mr. Dominici for their kind help and strong support at every step during my PhD study.

With my earnest thanks, I am deeply grateful for the love and understanding from my parents throughout all these years. I would never get this degree without your unconditional support.

At last, I want to give my eternal gratitude to Chen. Without your constant love, timely encouragement and inspiration, none of these would have been possible.

Hanqing Wen

ADVANCED NUMERICAL MODELING OF SEMICONDUCTOR MATERIAL PROPERTIES AND THEIR DEVICE PERFORMANCES

HANQING WEN

Boston University, College of Engineering, 2016

Major Professor: Enrico Bellotti, PhD
Professor of Electrical and Computer Engineering
Professor of Materials Science and Engineering

ABSTRACT

With the renewed concept of "Materials by Design" attracting particular attentions from the engineering communities in recent years, numerical methods that can reliably predict the optical and electrical properties of materials is highly preferable. Since the growth or the synthesis of a "designed" material and the ensuing devices is usually prohibitively expensive and time-consuming, numerical simulation tools that predict the properties of a proposed material together with its device performance before production is especially important and cost-effective. Furthermore, as the technology advances, semiconductor devices have been pushed to operate at their material limits, which requires a thorough understanding of the materials' microscopic processes under different conditions. Therefore, developing numerical models that are capable of investigating the semiconductor properties from material level to device level is highly desirable.

This dissertation develops a suite of numerical models in which optical absorption and Auger recombination in semiconductor materials are studied and simulated

together with their device performances. In particular, Green's function theory with full band structures is employed to investigate the material properties by evaluating the broadening of the electronic bands under the perturbation of phonons. As a result, both direct and phonon-assisted indirect processes are computed and compared among different materials. Drift-diffusion model and a 3D Monte-Carlo model are subsequently used to simulate the device characteristics with the obtained material parameters. This work first determines the full band structures for Si, Ge, α-Sn, HgCdTe, InAsSb and InGaAs alloys from EPM model, and then investigated the materials' minority carrier lifetime for IR detector applications. Finally device level simulations using drift-diffusion and 3D Monte-Carlo models are demonstrated. In particular, two issues of developing 3D Monte-Carlo device simulation models, namely the use of unstructured spatial meshes and elimination of particle-mesh forces, are discussed, which are crucial in simulating modern semiconductor devices having complex geometry and doping profiles.

Contents

List of Tables

List of Figures

List of Abbreviations

2D	.....	Two-Dimensional
3D	.....	Three-Dimensional
AR	.....	Auger Recombination
BC	.....	Boundary Condition
BLB	.....	Beattie-Landsberg-Blakemore
BTE	.....	Boltzmann Transport Equation
BZ	.....	Brillouin Zone
CB	.....	Conduction Band
CdTe	.....	Cadmium Telluride
CIC	.....	Could-in-Cell
DD	.....	Drift-Diffusion
DFT	.....	Density Functional Theory
EPM	.....	Empirical Pseudo-potential Method
ESWIR	.....	Extended Short Wavelength Infrared
FDTD	.....	Finite-Difference Time-Domain
FEM	.....	Finite Element Method
GaAs	.....	Gallium Arsenide
GaN	.....	Gallium Nitride
GaSb	.....	Gallium Antimonide
GeSn	.....	Germanium Tin
GR	.....	Generation-Recombination
HBT	.....	Heterojunction Bipolar Transistor
HgCdTe	.....	Mercury Cadmium Telluride
HgTe	.....	Mercury Telluride
HH	.....	Heavy-Hole
InAs	.....	Indium Arsenide
InAsSb	.....	Indium Arsenide Antimonide
InGaAs	.....	Indium Gallium Arsenide
InGaSb	.....	Indium Gallium Antimonide
InP	.....	Indium Phosphide
InSb	.....	Indium Antimonide
IR	.....	Infrared
IW	.....	Irreducible Wedge

LED	.....	Light-Emitting-Diode
LH	.....	Light-Hole
LHS	.....	Left Hand Side
LM	.....	Lattice-Matched
LWIR	.....	Long Wavelength Infrared
MC	.....	Monte-Carlo
MOSFET	.....	Metal-Oxide-Semiconductor Field-Effect Transistor
MWIR	.....	Mid Wavelength Infrared
PA	.....	Phonon-Assisted
RHS	.....	Right Hand Side
RR	.....	Radiative Recombination
RTA	.....	Relaxation Time Approximation
SL	.....	Superlattice
SLS	.....	Strained Layer Superlattices
SO	.....	Spin-Orbit
SRH	.....	Shockley-Read-Hall
SWIR	.....	Short Wavelength Infrared
T2SL	.....	Type-2 Superlattices
VB	.....	Valence Band

Chapter 1

Introduction

1.1 Numerical Modeling and Their Applications in Semiconductor Research

1.1.1 Motivations

The concept of "Materials by Design" which emphasizes the selection and design of materials based on a defined application first emerges in early 1990s (McCauley, 2013) and revives in recent years as a result of the rapid development of large scale multi-core computer systems. By definition, "Materials by Design" aims at constructing the materials from atomic level, using various elements and novel structures to meet some specific requirements in the application. To realize this goal, despite the non-trivial first-principle computations of basic material parameters, for example the band structures and phonon dispersions, one also needs a fast and reliable way to check the macroscopic properties of the material such as the optical gains or carrier recombination rates in optoelectronic devices. Allowing for the fact that the cost of growth or synthesis of the designed material and the ensuing devices is usually prohibitively high and the extremely long time to use *ab initio* method to investigate the material and device properties, developing time-efficient numerical models as necessary design verification tools that are capable of investigating the semiconductor properties from material level to device level is highly desirable.

As the optoelectronic devices ranging from laser sources to infrared detectors have increasingly become the factors shaping the way of communication, information

processing, object tracking and surveillance, among all the material properties the optical- and electrical-related parameters of semiconductors under different doping, temperature and injection conditions are of great interest. For example, significant efforts have been devoted to study germanium (Ge) and silicon (Si) optical properties as a result of the ongoing activities to develop light emitter, micro-display arrays and laser structures compatible with standard silicon fabrication processes. Particular attention has been given to understand how the optical properties of these materials change under the effect of strain and doping levels (Boucaud et al., 2013; Virgilio et al., 2013; Soref et al., 2012). On the experimental side, attempts to develop Ge-based lasers have leveraged the fact that using the combination of tensile strain and high doping, a significant enhancement of the direct radiative recombination rate could be obtained (S*ü*ess et al., 2013; Liu et al., 2010; Capellini et al., 2013), and the optimum strain to fulfill the lasing conditions in Ge has also been studied theoretically (Aldaghri et al., 2012). More exotic approaches based on membranes are also being explored (Kurdi et al., 2010; Boztug et al., 2013) which have demonstrated possible ways of achieving direct band-gap in Ge. At the same time Si light emitters (du Plessis et al., 2013; Venter et al., 2013b) based on avalanche processes have made the development of micro-displays possible (Venter et al., 2013a). However, due to the difficulty of modeling the second-order processes, most of these theoretical models used in the investigations of absorption and radiative recombination in silicon and germanium are based on the second-order perturbation theory (SOPT), which suffers a series of convergence issues.

Another important application of studying semiconductor materials is the design of photon detectors, in which direct band gap materials are widely employed to study the infrared (IR) radiation due to their high absorption coefficient. Unfortunately, most of the direct band gap materials are binaries or ternaries and suffer severe Shockley-

Read-Hall (SRH) carrier recombination effect due to the bad material qualities which in turn greatly reduces the carrier lifetime in the material. Over the years, many efforts have been devoted to improve the growth techniques of these narrow-gap materials (Wang et al., 1980; Sporken et al., 1992; Maranowski et al., 2001) and recently the engineering communities have reached a point where other intrinsic carrier recombination mechanisms in the materials, e.g. radiative and Auger recombination, start becoming the limiting factors of the device performances. Unlike the extrinsic recombination mechanisms such as SRH and dislocation, these intrinsic recombinations impose fundamental limits on the carrier lifetime which degrades the device quantum efficiency, contributes to the noise and are difficult to reduce. A thorough theoretical study of these effects is therefore necessary and instructive.

On the device simulation level, increasingly complex geometry and doping profiles have been introduced to the novel device architecture to further improve the functionality and reduce the power, noise and size of the device (Aldegunde and Kalna, 2015). Many numerical models have been developed to handle these requirements. (Hess, 1991; Ram-Mohan, 2002; Taflove and Hagness, 2005; Lindberg et al., 2014) Among them, finite difference time-domain (FDTD) method is widely used in the device simulation involving electromagnetic field, where as the drift-diffusion and Monte-Carlo method are employed to calculate the electronic characteristics of the semiconductor structures. Usually, for devices operating at low or moderate bias conditions where no breakdown takes place, drift-diffusion (DD) model is accurate enough to predict the carrier and electric field distribution in the devices. However, as the device dimension shrinks, much higher electric field may appear in the depletion region and the resulting impact ionization phenomenon is not easy to formulate in the framework of DD model except for the simplest analytical solutions. On the other hand, Monte-Carlo method, as an alternative way of device simulation, is suitable for

investigating devices under high electric field where the exact carrier distribution can be approximated by the average value of a large number of random trails (Hess, 1988). Traditional 2D or 3D Monte-Carlo simulation model usually employed tensor meshes to discretize the real space. Yet, when the geometry and doping profiles changes dramatically (for example in the case of simulating FinFet (Aldegunde and Kalna, 2015)), tensor mesh becomes more and more inaccurate and inefficient when defining the boundary conditions and refining the mesh density. Consequently it is in great demand to develop a 3D Monte-Carlo simulation model with unstructured mesh to be able to compute the electrical characteristics for the most advanced device structures in today's technology.

With these considerations in mind, establishing a suite of comprehensive numerical models which can be used to investigate both novel material and complex device properties is essential to reduce the research cost and further boost the development of semiconductor industry.

1.1.2 Goals

The present work intends to contribute to this effort by achieving three main objectives. First, we intend to develop an appropriate numerical model to study the optical absorption/gain coefficient, radiative and Auger recombination properties that can be applied to both direct- and indirect-gap semiconductor materials and goes beyond the conventional second order perturbation theory. Although, SOPT has been employed to analyze the indirect processes for many years (Bude et al., 1992), it is difficult to properly handle the divergences when the energy of transitions is not known a priori, introducing additional uncertainties to the theory (Takeshima, 1981). As a result, it may not be possible to apply this approach to all the possible cases consistently. The proposed model, however, is based on the Green's function formalism (Mahan, 1993), in which we employ the spectral density functions and electron self-energies

to describe the broadening of the states involved in the transitions due to different physical processes, therefore avoiding the divergence difficulties in SOPT naturally (Takeshima, 1982a; Takeshima, 1984b). Moreover, by including more terms in the electron self-energy, even higher order processes (e.g. two phonon-assisted transition) could be incorporated into the same theoretical framework of Green's function.

Second, we intend to develop a numerical approach suitable to compute the aforementioned properties using the full band structures with the strain effect included. This approach eliminates the approximations introduced by the analytical models based on parabolic or non-parabolic bands, makes it possible to analyze the behavior of the material when a significant portion of carriers populate the states far above the band edge (Bude et al., 1992; du Plessis et al., 2013), and naturally includes the effects of valence bands warping and anisotropy which is particularly important when evaluating the matrix elements of the radiative process.

Finally, we want to develop a 3D Monte-Carlo simulation model incorporating tetrahedron meshes to discretize the simulation space to accommodate the increasingly complicated geometry and doping structures of modern semiconductor devices. Specific problems concerning the search in the unstructured mesh and elimination of particle-mesh forces will be studied and addressed.

This dissertation is organized as follows: Chapter 2 outlines the theories of the numerical models used in this work, in which their implementations are also discussed; Chapter 3 describes the EPM electronic structures for the relevant semiconductor materials investigated in this dissertation; Using the electronic structure obtained from Chapter 3, Chapter 4 presents the optical properties of strained and relaxed Ge and Si using Green's function theory, in which both direct and phonon-assisted indirect optical transitions are investigated. Chapter 5 studies the minority carrier lifetime in IR materials which is determined by calculating the radiative and Auger

recombination rates in both strained and relaxed material; Chapter 6 computes the electrical field and carrier distribution in reverse biased Si LED structures using drift-diffusion model; Chapter 7 discusses a 3D Monte-Carlo model implemented with an unstructured tetrahedron mesh where two relevant issues of using such unstructured mesh, namely, searching in the tetrahedron mesh and the elimination of particle-mesh forces in the mesh are studied; finally, Chapter 8 concludes the whole dissertation and delivers the concluding remarks.

1.2 Recombination Mechanisms in Semiconductor Materials

One of the topic in this dissertation is to investigate the recombination processes which change the number of free carriers in the semiconductor materials. Basically, three types of carrier generation-recombination (GR) mechanisms are considered to be important and usually included in the device simulation: radiative, Auger and Shockley-Read-Hall recombinations. Among them, the first two mechanisms are determined by the doping levels and band structures of the materials which are intrinsic and independent of materials quality. Consequently, these two processes define the upper bound of the minority carrier lifetime and impose a fundamental limit of bulk material performances. The last one, which was originally proposed and studied by Shockley, Read and Hall, is mediated by the localized states or recombination centers in the material. As a result, the SRH recombination rate highly depends on the material quality and impurity levels which can be controlled during the material growth and synthesis. Sections below are detailed descriptions and classifications of the three GR processes.

1.2.1 Radiative Recombination

The radiative recombination happens when an electron in the conduction band loses its energy and fills a electron vacancy (hole) in the outer shell of an atom. The energy

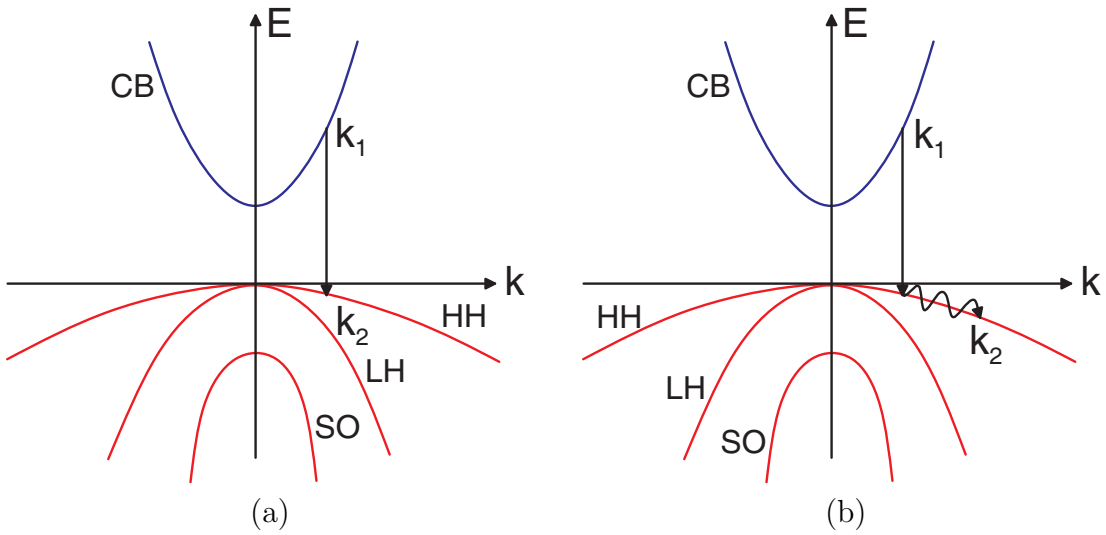


Figure 1·1: Schematics of radiative recombination processes for (a): direct recombination, (b): phonon-assisted indirect recombination.

lost by the electron is carried by an emitted photon to conserve the total energy of the system. As is shown in Fig. 1·1(a), since the momentum of the emitted photon is much smaller than the quasi-momentum of electrons in the crystal, in order to conserve the total crystal momentum, the recombined electron and hole must have similar $\mathbf{k}$ which is indicated by a vertical transition in the figure. Due to the fact that except for one electron and one hole, no other particles in the system participate in the recombination, this vertical transition is classified as direct process.

In a similar case which is shown in Fig 1·1(b), the initial electron at state $\mathbf{k}_1$ recombines with a hole at state $\mathbf{k}_2$ and emits a photon to conserve the total energy. However, one should notice that in this case, the crystal momentum for state $\mathbf{k}_1$ and $\mathbf{k}_2$ is different and the extra momentum is carried by a phonon. As a result, this process is phonon-assisted radiative recombination, which, as is shown in Chapter 2, Section 2.2, is a second order process. Due to the fact that the matrix element for electron-phonon interaction is small, with the same density of states, the probability of PA RR is generally much weaker than the direct RR. However, as will be shown in

Chapter 4, since the extra phonon could greatly increase the density of states for the final states, the overall PA RR rate could be comparable to the direct RR rate. As to the reverse process where an electron in the valence band absorb a photon and transit to conduction band leaving the system with one free electron and one free hole, it is the usual optical absorption.

1.2.2 Auger Recombination

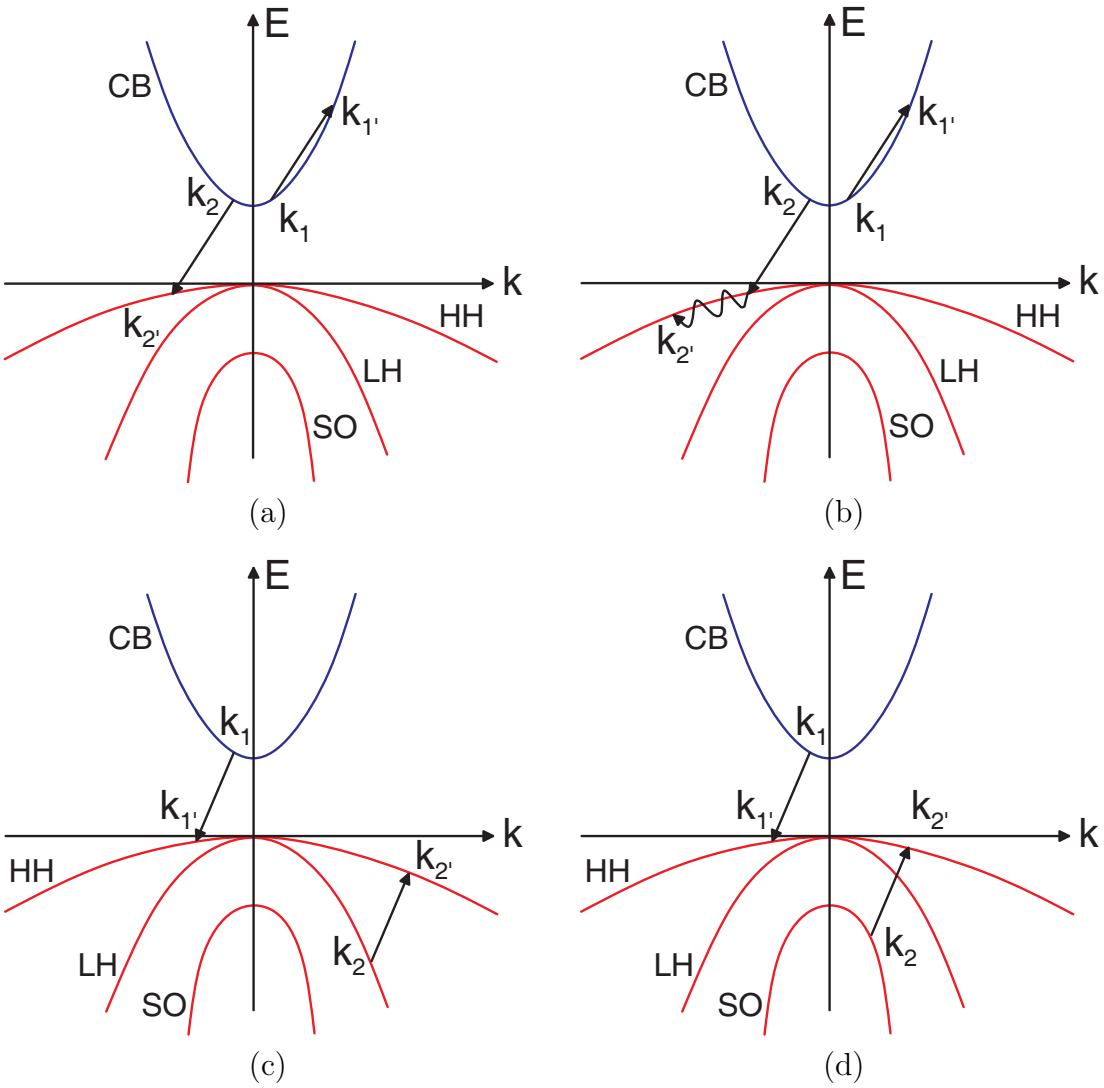


Figure 1·2: Schematics of Auger processes for (a): direct Auger-1 (CHCC) recombination, (b): phonon-assisted indirect Auger-1 (CHCC) recombination, (c): direct Auger-7 (CHLH) recombination and (d): direct Auger-S (CHSH) recombination.

The Auger recombination process involves more than one electron but no radiation. As shown in Fig 1·2(a), an electron in the conduction band can transfer its energy to another free electron and recombine with a hole in the HH band. Since two electrons

in the conduction band are needed to initiate the process, it can be understood that the Auger recombination is only important if the electron density is high enough, which indicates a heavily doped semiconductor or highly injected sample. Similar to the case of radiative recombination, Fig. 1·2(a) shows a direct Auger process in which no extra particle is needed to conserve the total crystal momentum. In contrast, from PA indirect Auger process, as shown in Fig. 1·2(b), a phonon is involved which scattered the a hole in the heavy-hole band to an intermediate state and recombines with an electron from the conduction band.

Depending on the type of carriers involved and their initial bands in the recombination, the Auger processes can generally be classified into the following four types.

- Auger-1 (CHCC) process, shown in Fig 1·2(a). Two electrons in CB initially interact and one electron $\mathbf{k}_2$ transfers certain amount of energy to electron at $\mathbf{k}_1$ and then recombines with a hole at $\mathbf{k}_{2'}$. The other electron, originally at $\mathbf{k}_1$ absorbs the energy and transit to a higher state $\mathbf{k}_{1'}$ in CB.

- Auger-7 (CHLH) process, shown in Fig. 1·2(c). One electron in CB recombines with a hole in HH band and transfer the emitted energy to a valence electron in LH band to HH band. (a hole transit from HH to LH.)

- Auger-S (CHSH) process, shown in Fig. 1·2(d). One electron in CB recombines with a hole in HH band and transfer the emitted energy to a valence electron in SO band to HH band. (a hole transit from HH to SO.)

- CHHH process (not plotted). One electron in CB recombines with a hole in HH band and transfer the emitted energy to a valence electron in lower HH band to higher HH band. (a hole transit from higher HH to lower HH state.)

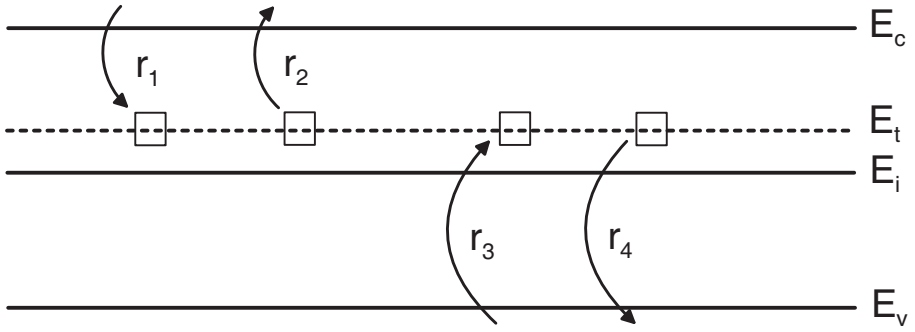


Figure 1·3: Free carriers interacts with localized states by four processes: r_1: an electron is captured in the localized states, r_2: an electron is emitted from localized states, r_3: a hole is captured, and r_4: a hole is emitted. The localized state shown is accepter type at energy E_t within the band gap. E_c represents the conduction band edge, E_v is the valence band edge and E_i is the intrinsic Fermi level of the material.

It should be noted that the reverse process of Auger recombination is impact ionization, which is important in the device breakdown region.

1.2.3 Shockley-Read-Hall Recombination

Recombination of free electrons and holes mediated by the localized states or recombination centers are called SRH recombination given by the initials of the three scientists who firstly studied it (Shockley and Read, 1952; Hall, 1952). As shown in Fig. 1·3, free electron and hole can recombine through a intermediate localized state which relaxes the requirement of crystal momentum conservation (Muller and Kamins, 2003). In reality, localized states always present in the energy gap due to the defects and impurity atoms in the crystal. As a result, SRH recombination rate is closely related to the material quality and doping levels which is extrinsic to the material itself and can be controlled by improving the material growth technologies.

Fig. 1·3 shows a localized state with energy E_t in the forbidden energy gap. Free electrons from conduction band can be captured by an empty localized state through thermal emission which is indicated in process r_1. Similar process can happen to free

holes as well where an electron in the localized state recombines with a free hole in the valence band (process r_4). This process is equivalent to a free hole transit to the localized state. Since process r_1 and r_4 do not need to happen simultaneously, the collective effect of r_1 and r_4, thus, is electron-hole recombination mediated by a localized state.

To formulate the SRH process, it is convenient to define the capture cross section of the localized state σ_n which describes the effectiveness of the localized state in trapping an electron. Assume the electron has thermal velocity v_{th}, the probability of process r_1 is:

$$r_1 = nN_t(1 - f(E_t))v_{th}\sigma_n \tag{1.1}$$

where n is the electron density, $f(E)$ is carrier occupation probability. At thermal equilibrium, $f(E)$ is Fermi-Dirac distribution.

The reverse process of r_1, where an electron is emitted back to the conduction band has probability:

$$r_2 = N_t f(E_t)v_{th}e_n \tag{1.2}$$

with

$$e_n = v_{th}\sigma_n n_i \exp\left(\frac{E_t - E_i}{k_B T}\right). \tag{1.3}$$

Similarly, the hole capture and emission can be written as:

$$r_3 = N_t f(E_t)pv_{th}\sigma_p \tag{1.4}$$

$$r_4 = N_t(1 - f(E_t))e_p \tag{1.5}$$

with

$$e_p = v_{th}\sigma_p n_i \exp\left(\frac{E_i - E_t}{k_B T}\right). \tag{1.6}$$

Combining all the thermal emission and capture processes in Eq. (1.1), (1.2), (1.4), (1.5), the overall net SRH recombination rate is derived as:

$$\begin{aligned} U &= R_{sp} - G_{sp} = r_1 - r_2 \\ &= \frac{pn - n_i^2}{\tau_n \left[p + n_i \exp\left(\frac{E_i - E_t}{k_B T}\right)\right] + \tau_p \left[n + n_i \exp\left(\frac{E_t - E_i}{k_B T}\right)\right]} \end{aligned} \tag{1.7}$$

where

$$\tau_n = (N_t v_{th} \sigma_n)^{-1} \tag{1.8}$$

$$\tau_p = (N_t v_{th} \sigma_p)^{-1}. \tag{1.9}$$

In a special case where $\sigma_p = \sigma_n = \sigma_0$, define $\tau_0 = (N_t v_{th} \sigma_0)^{-1}$, the net SRH recombination rate is:

$$U = \frac{pn - n_i^2}{\tau_0 \left[p + n + 2n_i \cosh\left(\frac{E_t - E_i}{k_B T}\right)\right]}. \tag{1.10}$$

1.3 3D Semiconductor Device Simulation

As the technology advances, geometry and doping profiles of modern semiconductor devices have become increasingly complex (Schuster and Bellotti, 2012; Kodama et al., 2008). Traditionally, in order to reduce the computational resources used by the simulation program to accommodate large device structures in the numerical calculation, 2D model which only simulates the cross section of the devices is widely employed (Hess, 1991). Though 2D model is capable of predicting most important electrical characteristics of a device which has the same cross section in one direction,

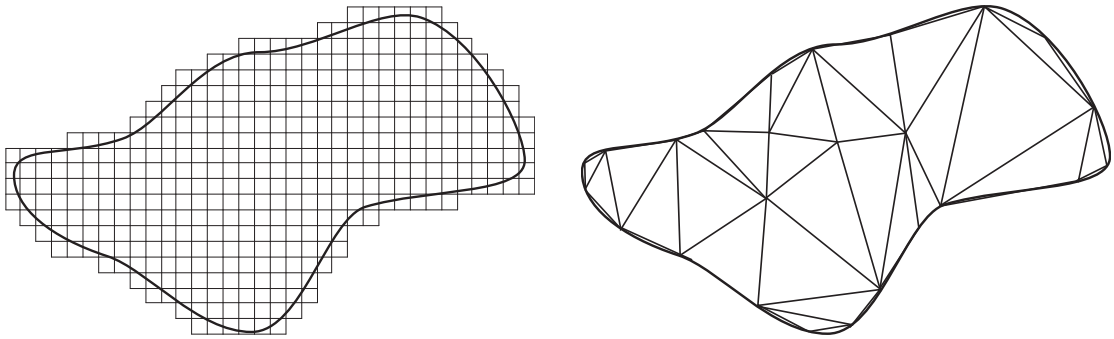


Figure 1·4: (a): Homogeneous tensor mesh and (b): triangular mesh used to discretize an irregular curve domain. 2D meshes are plotted in order to ease the visualization.

as the dimension of the device shrinks, real boundary conditions perpendicular to the 2D cut plane will begin to affect the device properties as well (Fichtner et al., 1983). Furthermore, to achieve higher performances, such as higher blocking voltage for power switches (Kizilyalli et al., 2013), higher gain for HBT, and lower fabrication cost, the geometry of the devices are necessary be designed in 3D fashion, which prevent the use of the simple 2D simulation model. Therefore, developing a 3D semiconductor device simulation model is crucial for studying the state-of-the-art device structures.

To implement 3D simulation model, a proper way of discretizing the simulation space is required. Two methods are widely adopted. One is to use tensor meshes which can be indexed and stored by using the position of its geometrical center. In fact, for device composed of cuboid material regions with its boundary and interfaces aligned with the Cartesian coordinate axes, it is convenient to employ the tensor meshes for the simplicity of searching and locating elements and particles (Dollfus et al., 2004). Moreover, by using the tensor mesh, finite difference method can be used to solve the device transport equations, such as Poisson equation and drift-diffusion equation, which is easier to implement than using the FEM scheme.

Another way of discretizing the simulation space is to use unstructured mesh, in

which the 3D simplex (tetrahedron) element is adopted most. Many advantages on the numerical simulation can be obtained by using the tetrahedron mesh compared to the simple tensor mesh. First is the coverage of curved or irregular simulation boundary. As shown in Fig. 1·4(a), for a curved boundary even though a large number of mesh nodes are used in the tensor mesh, the resulting homogeneous tensor mesh still cannot reproduce the boundary very well. On the contrary, if a 2D simplex (triangular) mesh is employed, as is shown in Fig. 1·4(b), the boundary can easily be covered with a small number of triangular elements. Second, when there is a need to refine the mesh locally, for example to refine the mesh at the channel area of a MOSFET, it is always desirable to keep the mesh density in other area (e.g. source and drain region) low in order to reduce the computational load. However, for a tensor mesh, since all the elements must be conformal and aligned, mesh refinement cannot be done locally, which increases the number of unnecessary elements in the simulation region. The tetrahedron mesh, instead, is suitable to be locally refined, where a large number of elements can be added to a small region without affecting the mesh density of the surrounding area (Wen and Bellotti, 2014). Consequently, in this work a tetrahedron mesh is utilized to discretize the simulation domain in the 3D device simulation.

As a final remark for this section, we want to point out that the advantages of using unstructured mesh does not come for free. Many issues, such as locating, searching and storing the unstructured mesh, need to be solved before the 3D simulation can be done efficiently for a large number of mesh elements. Moreover, in the case of Monte-Carlo simulation, the particle-mesh forces also need to be handled properly otherwise unphysical movement of particles can happen inevitably. Chapter 7 will be devoted to solve these problems concerning the 3D models.

Chapter 2

Theory of Numerical Modeling

To summarize the theories of the various numerical models employed in this dissertation, this chapter reviews the theory of EPM, direct and phonon-assisted indirect radiative and Auger recombination processes in the framework of Green's function formalism and device simulation theory including drift-diffusion model and Monte-Carlo method. The numerical implementations of these models is also presented.

2.1 Empirical Pseudopotential Method

Since the full band structures are crucial in order to include the band nonparabolicity effect in the calculation, a reliable model of band structures is highly desirable. In the literature, electronic structure calculations for both elemental and compound materials have been carried out either within the Density Functional Theory (DFT) framework or using the empirical pseudopotential method (EPM), whereas in the former calculation scheme, all-electron approaches and the state-of-the-art GW approaches (Geller et al., 2001; Malone and Cohen, 2013) have produced many high-quality band structures. On the other hand, the EPM scheme, though less accurate than the *ab initio* techniques, is also widely used in the community due to its simplicity, high computational efficiency and the flexibility that the resulting electronic structures can be optimized to closely match the experimental and/or the *ab initio* results. As shown in Fig. 2·1, the pseudo-wavefunction of electron from EPM can largely reproduce the real electron wavefunction in the region far from the atom cores while smooth the function near the cores. This

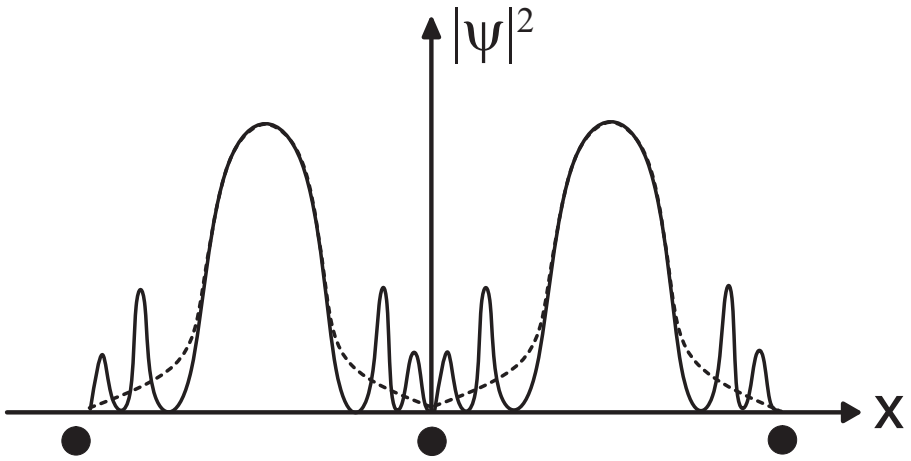


Figure 2·1: Schematics of electron wavefunctions obtained by using realistic atomic potential (solid line) and pseudo-potential (dashed line). Dots in the figure represent the position of atom cores. The pseudo-wavefunction can reproduce the realistic electron wavefunction in the region far from the atomic cores and varies smoothly near the cores, which could greatly reduce the number of plane waves needed to expand it.

will reduce the number of plane waves needed to expand the wavefunctions while maintain the relatively high accuracy of the resulting band structures. Furthermore, since the number of plane waves used to expand the electron wavefunctions in EPM is finite, all the calculations involving the overlap integral between two electronic states can be greatly simplified. As a result, we intend to compute the band structures using EPM.

2.1.1 Fitting of Pseudopotentials

The key information needed in the EPM model is the screened effective atomic potential for each atomic species in the material. The determination of these atomic potentials in standard approaches has relied on the fitting of form factors in order to reproduce a small set of band features which can be known experimentally or from *ab initio* results. The effective potentials are then adjusted using the n-dimensional downhill simplex

method to minimize the error between the theoretical and target results. For elemental semiconductors, such as Si or Ge, the form factors correspond to the matrix elements of the potential operators, whereas for the compound materials, the form factors are a combination of the composing atomic potentials averaged by the corresponding atomic molar fraction. Although, the form factors only need to be known on a discrete number of wavevectors, the knowledge of the functional dependence (Goano et al., 2000) of the screened atomic potential on the wavevector is actually paramount to obtain an electronic structure that can accurately reflect the effect of strain (Kim and Fischetti, 2010) and carrier-phonon interaction (Bertazzi et al., 2009). In this chapter, we have determined the screened atomic potentials for both elemental (Si and Ge) and compound materials (HgCdTe, InGaAs, InAsSb) so that the calculated electronic structures reproduce not only the correct energy transitions, but also the effective masses of electrons and holes and the hydrostatic deformation potentials. Besides the screened atomic potentials, a spin-orbit correction is also considered in our EPM model since the effect is significant in all the materials we considered. The crystal Hamiltonian is therefore given by:

$$H = \frac{\hbar^2}{2m_0}\nabla^2 + V^{\mathrm{LOC}}_{\mathbf{G},\mathbf{G}'} + H^{\mathrm{SO}}_{\mathbf{G},\mathbf{G}'}. \tag{2.1}$$

in which the second term in Eq. (2.1) is the spin-orbit part of the Hamiltonian. The relativistic correction for a state with wavevector $\mathbf{k}$ is introduced as a perturbation term, and the spin-orbit matrix elements are given by (Weisz, 1966; Bloom and Bergstresser, 1968; Walter et al., 1970):

$$\begin{aligned}\langle \mathbf{K}_{i,\nu} | \boldsymbol{H_{so}} | \mathbf{K}_{j,\nu'} \rangle &= \boldsymbol{\sigma_{\nu\nu'}} \cdot (\mathbf{K}_i \times \mathbf{K}_j) \\ &\times \left\{ \lambda^A \sin(\mathbf{G}_i - \mathbf{G}_j) \cdot \tau - i\lambda^S \sin(\mathbf{G}_i - \mathbf{G}_j) \cdot \tau \right\}. \end{aligned} \tag{2.2}$$

where $\boldsymbol{\sigma_{\nu\nu'}}$ is the Pauli matrix; $\mathbf{G}_i$ and $\mathbf{G}_j$ are the reciprocal lattice vectors; $\mathbf{K}_{i,j} =$

$\mathbf{G}_{i,j} + \mathbf{k}$; and τ is the atomic position vector. The symmetric and antisymmetric spin parameters λ^S and λ^A are given by:

$$\lambda^S = \frac{(\lambda_1 + \lambda_2)}{2} \tag{2.3}$$

$$\lambda^A = \frac{(\lambda_1 - \lambda_2)}{2}. \tag{2.4}$$

where λ_1, λ_2 are defined as:

$$\lambda_1 = \mu \boldsymbol{B}_{nl}^{(1)}(K_i)\boldsymbol{B}_{nl}^{(1)}(K_j) \tag{2.5}$$

$$\lambda_2 = \gamma\mu \boldsymbol{B}_{nl}^{(2)}(K_i)\boldsymbol{B}_{nl}^{(2)}(K_j). \tag{2.6}$$

The parameter γ is set to be the ratio of the spin-orbit energy of the free cation and anion atoms (Herman et al., 1963) ($\gamma = 1$ for elemental materials and non-unity value for compound materials), and μ has to be adjusted to obtain the correct spin-orbit energy for the material. The values of $\boldsymbol{B}_{nl}^{(i)}(k)$ are then computed for all the atomic species i in the crystal according to:

$$\boldsymbol{B}_{nl}^{(i)}(k) = C \int_0^\infty \boldsymbol{j}_l(kr)\boldsymbol{R}_{nl}^{(i)}(r) r^2 \mathrm{d}r \tag{2.7}$$

where $\boldsymbol{j}_l(kr)$ is the spherical Bessel function of order l, $\boldsymbol{R}_{nl}^{(l)}(r)$ is the atomic wavefunction corresponding to the quantum numbers n, l, and C is a constant (Walter et al., 1970) that satisfies:

$$\lim_{k\to 0} \frac{\boldsymbol{B}_{nl}^{(i)}(k)}{k} = 1 \tag{2.8}$$

We note that in Eq. (2.7), the formulation used for $\boldsymbol{R}_{nl}^{(l)}(r)$ is crucial to facilitate the evaluation of $\boldsymbol{B}_{nl}^{(i)}(k)$. In the present work, $\boldsymbol{R}_{nl}^{(l)}(r)$ is expanded through a set of suitable basis functions obtained with the Roothaan-Hartree-Fock method (Clementi and Roetti, 1974), and consequently, $\boldsymbol{B}_{nl}^{(i)}(k)$ can be calculated according to an

analytical expression (Masovic and Vukajlovic, 1983).

Strain effect on the band structures can also be added to the EPM model. For relaxed materials, form factors at discrete reciprocal lattice vectors $q = |\mathbf{G} - \mathbf{G}'|$ $= \sqrt{3},\ \sqrt{4},\ \sqrt{8},\ \sqrt{11},\ \sqrt{16}$ (in unit of $2\pi/a_0$) are sufficient to determine the band structures for zincblend material. However, when strain is applied, the reciprocal lattice vector will be deformed accordingly and the fitted potential in the neighborhood of $q = \sqrt{3}, \sqrt{4}$ etc. will determine the band structure instead (Kim and Fischetti, 2010). As a result, to get a proper band structure that is accurate for all strain (compressive and tensile) conditions, not only the form factors at specific q, but also the values in its neighborhood should be adjusted, which can be achieved by fitting the form factors and their first derivatives at the same time (Kim and Fischetti, 2010). Furthermore, to maintain the continuity of the atomic potential under strain and therefore a well-defined first derivative of the potential, cubic spline with natural boundary condition is employed to represent the effective atomic potential $V^{\mathrm{LOC}}_{\mathbf{G},\mathbf{G}'}(q)$. Detailed fitting parameters and the resulting bands for both elemental and compound materials are presented in Chapter 3.

2.1.2 Virtual Crystal Approximation

When several simple materials (elemental or binary materials) form a compound alloy, positions of the composing cation or anion in the material structure are usually randomly filled by the cation and anions of the composing materials, which causes disorder of the atomic structure. To accommodate this phenomenon in the alloy, virtual crystal approximation (VCA) (Lee et al., 1990; Bellaiche and Vanderbilt, 2000) is usually used to simplify the short-range disorder effects in the solid. By approximating the disorder of actual components with a virtual atom that interpolates the behavior of the composing atoms, the band structures of disordered ternary materials can be calculated at the same cost as the ordered one at the expense of loosing the fine

details of disorder effects (Dargam et al., 1997). Using VCA, the full band structures of $A_xB_{1-x}C$ alloy, for example, are fitted starting from the screened atomic potentials of binary material AC and BC. The atomic volume in the original VCA theory derived in Ref. (Lee et al., 1990) together with the disorder of the mixed local potentials are represented by another fitting parameter P_{dis} in the final form of the effective potential in the ternary alloy (Kim and Fischetti, 2010), in which the potentials for cation and anion are averaged between the two binary materials separately:

$$V_{\text{ABC}}^{\text{cation}}(q) = xV_{\text{AC}}^{\text{A}} + (1-x)V_{\text{BC}}^{\text{B}} - P_{\text{dis}}x(1-x)(V_{\text{AC}}^{\text{A}} - V_{\text{BC}}^{\text{B}}), \tag{2.9}$$

$$V_{\text{ABC}}^{\text{anion}}(q) = xV_{\text{AC}}^{\text{C}} + (1-x)V_{\text{BC}}^{\text{C}} - P_{\text{dis}}x(1-x)(V_{\text{AC}}^{\text{C}} - V_{\text{BC}}^{\text{C}}). \tag{2.10}$$

The lattice constant of ABC alloy is usually linearly depended on the molar fraction of the two composing material without bowing:

$$a_{0,\text{ABC}} = xa_{0,\text{AC}} + (1-x)a_{0,\text{BC}}. \tag{2.11}$$

We noted that for most of the alloys, in which the spin-orbit coupling does not strongly affect the electron energy in its composing materials, Eq. (2.9) and (2.10) generally give good agreement with the measured band parameters. However, if the spin-orbit coupling in the composing materials is strong enough to change its band structure, the disorder effect of spin-orbit coupling in the alloy should be considered in the same way as the atomic potentials. Equations for this disorder effect are:

$$\lambda^s(q) = x\lambda_{AC}^s(q) + (1-x)\lambda_{BC}^s - P_{\text{dis}}^{so}x(1-x)(\lambda_{AC}^s - \lambda_{BC}^s), \tag{2.12}$$

$$\lambda^a(q) = x\lambda_{AC}^a(q) + (1-x)\lambda_{BC}^a - P_{\text{dis}}^{so}x(1-x)(\lambda_{AC}^a - \lambda_{BC}^a). \tag{2.13}$$

Here, λ^s and λ^a is the symmetric and antisymmetric spin parameters defined in Eq. (2.3) and (2.4). The spin disorder parameter P_{dis}^{so} is also a fitting parameter which

will be adjusted so that the band gap energy of the alloy under study matches to the experimental data.

2.2 Theory of Radiative Processes

In this section the theoretical model for radiative recombination process using equilibrium Green's function theory is presented. In particular, both direct process and phonon-assisted indirect process are derived within the same theoretical framework, in which the direct radiative process including optical absorption and radiative recombination is expressed as a special case of phonon-assisted indirect process.

As was pointed out in Ref. (Takeshima, 1982a), the main motivation of using the Green's function approach lies in trying to avoid the numerical and convergence issues arising from conventional SOPT (Bude et al., 1992). However, a number of investigations that have employed the Green's function theory were mainly focused on computing the recombination rates of Auger process mediated by the coulomb interaction (Takeshima, 1982a; Takeshima, 1984b; Bertazzi et al., 2012). In the case of radiative recombination processes, dipole interactions where the electromagnetic field interacts with electrons and holes needs to be taken into account instead. Specifically, in this work PA indirect photon absorption and emission processes are studied. Although the derivation of suitable expressions to compute PA and impurity-assisted Auger recombination rates using quasi-equilibrium Green's function formalism (Matsubara method) have already been presented by Takeshima (Takeshima, 1982b), little or no information is available on the analogous formula for the PA radiative recombination/absorption processes, except for an expression reported in Ref. (Takeshima, 1984b) (Eq. 2.4). In the work of Kwong *et al.* (Kwong et al., 2009), though the phonon-assisted process was studied along with the electron-hole plasma interaction using quasi-equilibrium Green's function, only a phenomenological expression was

employed to model the PA process and a more comprehensive approach to calculate the PA absorption is still missing. On the other hand, Bardyszewski and Yevick (Bardyszewski and Yevick, 1989) proposed an alternative way to describe the optical recombination in the context of electron-hole plasma interaction using non-equilibrium Green's function (Keldysh method) which gives good agreement with the experimental data but requires the inclusion of dynamic screening in the model. This leads to a significant increase of computational complexity for the full band calculation while only minor correction to the optical gain can be achieved. Since our main focus is on the PA processes and the full band structure effects, in the present work we employed a Thomas-Fermi static screening and ignored the additional correction due to the electron-plasma interaction.

2.2.1 Phonon-assisted Indirect Optical Absorption

We start from the description of the electron-photon interaction that lies at the heart of the absorption or radiative recombination processes. The Hamiltonian of the system is given by:

$$H_{\text{tot}} = H_{\text{elec}} + H_{\text{ph}} + H_{\text{e-photon}} \tag{2.14}$$

where H_{elec}, H_{ph} and $H_{\text{e-photon}}$ are the Hamiltonian for electrons, photons and electron-photon interaction respectively. Furthermore, we assume that the description of electron-phonon interaction, electron-impurity or dislocation scatterings, have been included in H_{elec}. As a result, the number operator for electrons in band l, at wavevector $\mathbf{k}$, spin σ: $\hat{N}_e = \sum_{l,\mathbf{k},\sigma} C^{\dagger}_{l,\mathbf{k},\sigma} C_{l,\mathbf{k},\sigma}$ will commute with all the terms in Eq. (2.14) except for $H_{\text{e-photon}}$ which is given by:

$$H_{\text{e-photon}} = -\frac{e}{m}\sqrt{\frac{\hbar}{2\omega V \epsilon}} \sum_{i,j} |f_{ij}|^2 A_{\mathbf{k}_p}(t) C^{\dagger}_i C_j \tag{2.15}$$

where

$$f_{ij} = \langle i \,|\hat{\mathbf{e}} \cdot \mathbf{P}|\, j \rangle \,, \quad i,j \in \{(l, \mathbf{k}, \sigma)\} \tag{2.16}$$

$$A_{\mathbf{k}_p}(t) = a_{\mathbf{k}_p} \mathrm{e}^{-i\omega t} + a^{\dagger}_{-\mathbf{k}_p} \mathrm{e}^{i\omega t} \tag{2.17}$$

are the matrix elements for dipole interaction (electron-photon interaction) and vector potential for the electro-magnetic field. Using standard density operator theory on the the electron number operator $\hat{N}_e$, it is possible to obtain the following expression where the effect of phonon or impurity scattering on the electronic bands is combined with the common dipole interaction:

$$\begin{aligned}\alpha_{12}(\hbar\omega_{\mathrm{ph}}) = \frac{2\pi}{n_r c_0 V \omega_{\mathrm{ph}} \epsilon_0} \sum_{\mathbf{k}} |P_{1,2}(\mathbf{k})|^2 \times \int \mathrm{d}E_1' \int \mathrm{d}E_2' \left(\Theta(E_1') - \Theta(E_2')\right) \\ \times \delta(\mu_c - \mu_v + E_2' - E_1' - \hbar\omega_{\mathrm{ph}}) \times \mathrm{Im}G^R_{l_1}(\mathbf{k}, E_1') \mathrm{Im}G^R_{l_2}(\mathbf{k}, E_2').\end{aligned} \tag{2.18}$$

In Eq. (2.18), α_{12} is the total absorption coefficient between band 1 and 2 for photon energy of $\hbar\omega_{\mathrm{ph}}$. $\Theta(E)$, n_r, c_0, ϵ_0, m_0, e, V are the distribution function of electron, refractive index of the material, speed of light in vacuum, vacuum permittivity, electron mass, electron charge and volume of crystal respectively. $P_{1,2}(\mathbf{k})$ is the matrix element for dipole interaction defined in Eq. (A.8), and $\Theta(E)$ is the Fermi factor defined as $\Theta(E) = 1/[1 + e^{E/(k_B T)}]$, where the energy E is measured from the corresponding quasi-Fermi level. The term $\mathrm{Im}G^R_{l_i}(\mathbf{k}, E_i'), i = 1, 2$ is the spectral density function for band i, which represents the broadening of that band as the result of all the relevant scattering processes. This broadening can be evaluated using Matsubara Green's function (Mahan, 1993), and the detailed derivation of this equation can be found in Appendix A. Though the formalism shown in Appendix A can take into account any broadening mechanisms, in this work only the broadening due to electron-phonon scattering is considered.

2.2.2 Direct Optical Absorption

One way to obtain the direct optical absorption coefficient is to use the Fermi's Golden Rule, and treated the electron-photon interaction as a first order perturbation to the single electron system in the periodic potential of the crystal (Chuang, 2009). However, if we have derived the phonon-assisted indirect optical absorption coefficient, it is intuitive to treat the direct process as a special case of the indirect transition. In fact, as is described in Chapter 2, Section 2.2, the spectral density function in Eq. (2.18) represents the broadening of electronic bands under the perturbation of phonon. In the case of direct absorption, no phonon participates in the process and the broadening of the band can be assumed to be a Dirac δ-function. Consequently, by setting all the spectral density function in Eq. 2.18 to $\delta(E_l(\mathbf{k}) - E_l')$, with $E_l(\mathbf{k})$ being the electron energy in band l, and integrating the equation over E_1' and E_2', it is easy to get:

$$\alpha_{12}^{dir}(\hbar\omega_{\mathrm{ph}}) = \frac{2\pi}{n_r c_0 V \omega_{\mathrm{ph}} \epsilon_0} \sum_{\mathbf{k}} |P_{1,2}(\mathbf{k})|^2 \left(\Theta[E_1(\mathbf{k})] - \Theta[E_2(\mathbf{k})]\right) \times \delta(E_2(\mathbf{k}) - E_1(\mathbf{k}) - \hbar\omega_{\mathrm{ph}}). \tag{2.19}$$

As in the previous derivation, we have ignored the momentum of photon $\mathbf{k}_{ph}$ in Eq. 2.19, which is much smaller than the electron momentum.

2.2.3 Radiative Recombination Rate and Carrier Lifetime

Radiative Recombination Rate

The total radiative recombination rate relates to the optical absorption coefficient $\alpha(\hbar\omega)$ by (Chuang, 2009):

$$R_{21} = \frac{n_r^2}{\pi^2 \hbar^3 c_0^2} \int_0^\infty \frac{\alpha_{21}(E) E^2}{\exp(E/k_B T) - 1} \, \mathrm{d}E \tag{2.20}$$

where $\alpha_{21}(E)$ is the energy dependent absorption coefficient from band 1 to band 2,

n_r the refractive index, and $\hbar$ and c_0 are the reduced Planck constant and the speed of light respectively, k_B is the Boltzmann constant.

Substituting the obtained optical absorption coefficient Eq. (2.18) into Eq. (2.20) and taking the derivation of the equation, the radiative recombination rate per unit volume per energy interval from band 2 to band 1 is then computed as (Chuang, 2009):

$$\left.\frac{\mathrm{d}R_{21}}{\mathrm{d}E}\right|_{\hbar\omega_{\mathrm{ph}}} = \frac{2n_r\omega_{\mathrm{ph}}}{\pi\hbar c_0^3 v\epsilon_0}\sum_{\mathbf{k}}|P_{1,2}(\mathbf{k})|^2 \times \int \mathrm{d}E_1' \int \mathrm{d}E_2'\Theta(E_2')\left(1-\Theta(E_1')\right)$$
$$\times\, \delta(\mu_c - \mu_v + E_2' - E_1' - \hbar\omega_{\mathrm{ph}}) \times \mathrm{Im}G_{l_1}^R(\mathbf{k}, E_1')\mathrm{Im}G_{l_2}^R(\mathbf{k}, E_2'). \qquad (2.21)$$

Moreover, as was mentioned in the Appendix A, when computing Eqs. (2.18) and (2.21), we have ignored the contribution of PA inter-band transition to the self-energy. As a result, the spectral density functions for the two bands $\mathrm{Im}G_{l_i}^R(\mathbf{k}, E), i = 1, 2$ are independent. Therefore, by using the two spectral density functions in Eqs. (2.18) and (2.21), one can actually analyze two-phonon events, where an electron is first scattered in the lower (upper) band, absorbs (emits) a photon and is eventually scattered by another phonon reaching its final state in upper (lower) band. Alternatively, if only one of the electron spectral density functions is used in Eqs. (2.34) and (2.35), leaving the others to be a delta function, one phonon event is considered instead. Consequently, if one considers all possible cases, direct, one-phonon and two-phonon PA events can all be investigated. As will be shown in Section 4.2.2, the case of two-phonon processes are evaluated for Si, which is not possible with the traditional SOPT method.

As a final note we want to point out that the matrix element that describes the dipole interaction in Eq. (A.8) is a quantity averaged over all polarizations $\hat{\mathbf{e}}$. When the polarization effect needs to be studied, we explicitly evaluate the matrix elements for a specific polarization state.

Minority Carrier Lifetime

By integrating Eq (2.21) over all possible photon energy, we can obtain the total radiative recombination rate R_{21}^{tot} between band 2 and band 1 of a material under specific doping concentration and lattice temperature. Determining the total RR rate is the key step of finding the minority carrier lifetime and RR coefficient in the material which are defined as:

$$R_{21}^{tot} = (B_n + B_p)(np - n_i^2), \tag{2.22}$$

$$\tau_r = \frac{n_i^2}{R_{21}^{tot}(n_0 + p_0 + \Delta)}. \tag{2.23}$$

In this equation, n_i, n_0 and p_0 are the material's intrinsic carrier concentration, concentrations of electrons and holes at equilibrium respectively. Δ is the excess carrier concentration which fulfills $\Delta \ll$ the majority carrier concentration in order to to maintain the validity of the definition of low injection condition. In the following chapters, Eq. (2.18), (2.21) and (2.23) will be used extensively to evaluate and compare the carrier lifetime among different materials and operating conditions.

2.3 Theory of Auger Processes

The theoretical aspects of both direct (also called "pure-collision") and PA indirect Auger recombination processes have been studied and developed since 1960s when Beattie (Beattie and Landsberg, 1959) firstly proposed an analytical expression, known as BLB formula, to compute the Auger coefficient under the approximation of parabolic band structure and non-degenerate statistics. Though the BLB formula is a convenient way to estimate the Auger lifetime in direct band gap material and therefore is still widely adopted in the device simulation nowadays, the calculated rates, on the other hand, heavily depends on the value of overlap integral factor $|F_1F_2|$, which

is historically chosen between 0.1 to 0.3 without any clear guidelines. While recent calculations(Grein et al., 2008) based on a 14-band $k \cdot p$ band structure indicated that the overlap integral factor should be closer to the lower bound of its supposed range to match the numerical results, the actual value of $|F_1 F_2|$ remains a fitting parameter. As a result, to eliminate the ambiguity of evaluating Auger coefficient, a first-principle systematic approach to compute the recombination process is highly desirable.

At present, perhaps the most rigorous Auger recombination theory is the one that was developed by Takeshima (Takeshima, 1982b) and Bardyszewski *et al.* (Bardyszewski and Yevick, 1985) who incorporated both direct and indirect (e.g. phonon-assisted and impurity assisted) processes into the Green's function formalism to enhance the accuracy of the model. Although a parabolic band approximation was used in these early works due to the limitation of the computing infrastructure, the formalism itself is amenable to the use of a full band structure of the material being studied, which is particularly important to increase the accuracy of the calculation when the carriers' energy levels are far from the Γ point.(Wen and Bellotti, 2015b) It is worthy of note that, besides the Green's function theory, second-order perturbation theory (SOPT) is also widely employed in studying the indirect Auger process (Krishnamurthy et al., 1997). However, due to the undefined behavior of the SOPT equation when a virtual state overlaps with a real one, the theory's numerical accuracy has been questioned (Takeshima, 1982b; Bude et al., 1992), especially when a full band structure is used. To avoid this difficulty, in this work we followed the procedure proposed by Takeshima (Takeshima, 1982b; Takeshima, 1984a) and implemented the Green's function theory of Auger recombination with full band structures obtained from EPM. A detailed derivation of the expression for the recombination rates was presented in Ref.(Takeshima, 1982b) and will not be repeated here. Instead, the final expressions are presented below for completeness.

2.3.1 Phonon-assisted Indirect Auger Recombination

The total Auger recombination rate per unit volume is given by:

$$\begin{aligned} R_{\mathrm{AR}} = & \frac{2\pi}{\hbar}\frac{V^3}{(2\pi)^9}[1-e^{(\mu_v-\mu_c)/k_BT}] \\ & \times \int\int\int\int \mathrm{d}\mathbf{k}_1\mathrm{d}\mathbf{k}_2\mathrm{d}\mathbf{k}_{1'}\mathrm{d}\mathbf{k}_{2'}\,|M_{ee}|^2\,\delta(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_{1'}-\mathbf{k}_{2'}) \\ & \times \int \mathrm{d}E_1\mathrm{d}E_2\mathrm{d}E_{1'}\mathrm{d}E_{2'}\Theta(E_1)\Theta(E_2)[1-\Theta(E_{1'})][1-\Theta(E_{2'})] \\ & \times \mathrm{Im}G^R_{l_1}(\mathbf{k}_1,E_1)\mathrm{Im}G^R_{l_2}(\mathbf{k}_2,E_2)\mathrm{Im}G^R_{l_{1'}}(\mathbf{k}_{1'},E_{1'})\mathrm{Im}G^R_{l_{2'}}(\mathbf{k}_{2'},E_{2'}). \end{aligned} \tag{2.24}$$

where $\mathrm{Im}G^R_{l_i}(\mathbf{k}_i,E_i)$ is the spectral density function representing the energy broadening at $\mathbf{k}_i$ in Brillouin Zone (BZ) around energy E_i at band l_i. The specific form of the expression can be found in Chapter 2, Section 2.4. M_{ee} is the matrix element for Auger process which involves four electron/hole states(Laks and Neumark, 1990):

$$\begin{aligned} M_{ee} = & \int\int \mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2\phi^*_{\mathbf{k}_1}(\mathbf{r}_1)\phi^*_{\mathbf{k}_2}(\mathbf{r}_2)\nu(\mathbf{r}_1-\mathbf{r}_2) \\ & \phi_{\mathbf{k}_{1'}}(\mathbf{r}_1)\phi_{\mathbf{k}_{2'}}(\mathbf{r}_2) + \text{exchange term}, \end{aligned} \tag{2.25}$$

with

$$\phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{G}}[A_u(\mathbf{k}+\mathbf{G})+A_d(\mathbf{k}+\mathbf{G})]e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{2.26}$$

$$\nu(\mathbf{r}) = \sum_{\mathbf{G}}\int_{\mathrm{BZ}}\frac{\mathrm{d}\mathbf{q}}{(2\pi)^3}\nu(\mathbf{q}+\mathbf{G})e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}, \tag{2.27}$$

$$\nu(\mathbf{q}+\mathbf{G}) = \sum_{\mathbf{G}'}\epsilon^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})\frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}'|^2}. \tag{2.28}$$

Here, V is the volume of the crystal and $\phi_{\mathbf{k}}(\mathbf{r})$ is the Bloch wavefunction whose plane wave expansion coefficients are $A_u(\mathbf{k}+\mathbf{G})$ and $A_d(\mathbf{k}+\mathbf{G})$ for spin up and spin down states respectively (Harrison et al., 1999). $\epsilon^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ is the dielectric function of the material and $\nu(\mathbf{r})$ is the screened Coulomb interaction. The exchange term in

Eq. (2.25) can be obtained by swapping the indices 1' and 2' in the direct term. When using plane waves to expand the electron wavefunction $\phi_{\mathbf{k}}(\mathbf{r})$, the overlap integral of the four Bloch states involve a four-fold summation over the reciprocal vector $\mathbf{G}$. Due to the large number of plane waves used in the EPM, direct summation over $\mathbf{G}$ is very time consuming. Since the interaction matrix element needs to be evaluated very frequently during the calculation of Auger rate, we followed the approach of Laks *et al.*(Laks and Neumark, 1990) and reduced the four-fold summation into a two-fold one which greatly decreased the computing requirement. It should be noted that since Auger process is mediated by the electron-electron interaction, the screening of the electric field due to electron gas plays a vital role in modeling the interaction. Therefore, in this work we employed a wavevector-dependent dielectric function $\epsilon(\mathbf{q})$ to account for the spatial screening effect:

$$\epsilon(\mathbf{q}) = 1 + \frac{c_1}{|\mathbf{q}|^{c_2} + c_3}, \tag{2.29}$$

with the parameters c_1, c_2 and c_3 fitted to the numerical results from the random phase approximation (Walter and Cohen, 1972). We want to point out that a more rigorous way of modeling the screening effect in Auger process is to further include the frequency dependence in the dielectric function. However, due to the significant increase of computing time, the dynamic screening effect is hard to incorporate with full band structure and as a result, only wavevector-dependent screening is considered in this work.

It should be noted that due to the complexity of the energy integral in Eq. (2.24), when calculating the PA Auger rate, only one band is broadened. Since the heavy hole (HH) band usually has the largest broadening compared to that of the conduction bands, light hole band and spin-orbit band, only the spectral function for one of the final state in HH $\mathrm{Im}G^R_{HH}(\mathbf{k}_i, E_i)$ will be calculated in the Auger process.

Similar to the RR rate but using the total AR rate instead, Auger coefficients C_n, C_p and minority carrier lifetimes τ_{AR} under low injection condition can be obtained as:

$$R_{\mathrm{AR}} = (C_n n + C_p p)(np - n_i^2), \tag{2.30}$$

$$\tau_{\mathrm{AR}} = \frac{n_i^2}{R_{\mathrm{AR}}(n_0 + p_0 + \Delta)}, \tag{2.31}$$

where n_i, n_0, p_0 and Δ are the intrinsic, electron and hole concentrations at thermal equilibrium and injected carrier concentration respectively.

To facilitate the calculation of the overlap integral in the BLB formula, the corresponding n-type Auger coefficient C_n is cited as the following (Bellotti and D'Orsogna, 2006):

$$\begin{aligned} C_n = & \frac{\left(\frac{m_e*}{m_0}\right)|F_1F_2|^2}{7.6\times10^{-18} n_i^2 \epsilon_\infty^2 \left(1+\frac{m_e*}{m_h*}\right)^{1/2}\left(1+2\frac{m_e*}{m_h*}\right)} \\ & \times \left(\frac{E_g}{k_BT}\right)^{-3/2} \exp\left[-\frac{1+2\frac{m_e*}{m_h*}}{1+\frac{m_e*}{m_h*}}\cdot\frac{E_g}{k_BT}\right] \end{aligned} \tag{2.32}$$

with m_e*, m_h*, ϵ_∞, E_g, $|F_1F_2|$ being the electron and hole effective masses, high-frequency dielectric constant, band gap energy and overlap integral factor respectively. Except for $|F_1F_2|$, the empirical formulas of these quantities as a function of composition ratio and temperature can be found in Ref. (Lopes et al., 1993; Schuster et al., 2012; Ioffe,).

2.3.2 Direct Auger Recombination

The equation for direct AR rate can be determined in the same way of RR rate by evaluate Eq. (2.24) in the special case where no broadening of bands are considered.

Let the spectral density function $\mathrm{Im}G^R_{l_i}(\mathbf{k}_i, E_i) = \delta(E(\mathbf{k}_i) - E_i)$, with $i = 1, 2, 1', 2'$ and $E(\mathbf{k}_i)$ being the band energy at $\mathbf{k}_i$, Eq. (2.24) can be simplified as:

$$\begin{aligned} R^{dir}_{\mathrm{AR}} = & \frac{2\pi}{\hbar}\frac{V^3}{(2\pi)^9}[1 - e^{(\mu_v - \mu_c)/k_B T}] \\ & \times \int\int\int\int \mathrm{d}\mathbf{k}_1 \mathrm{d}\mathbf{k}_2 \mathrm{d}\mathbf{k}_{1'} \mathrm{d}\mathbf{k}_{2'} \left|M_{ee}\right|^2 \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{1'} - \mathbf{k}_{2'}) \\ & \times \delta\left(E(\mathbf{k}_1) + E(\mathbf{k}_2) - E(\mathbf{k}_{1'}) - E(\mathbf{k}_{2'})\right) \\ & \times \Theta(E(\mathbf{k}_1)\Theta(E(\mathbf{k}_2)[1 - \Theta(E(\mathbf{k}_{1'})][1 - \Theta(E(\mathbf{k}_{2'})]. \end{aligned} \tag{2.33}$$

Consequently, the direct AR rate and its related minority carrier lifetime can be found using Eq. (2.33).

2.4 Electron Self-Energy under Phonon Perturbation

In order to evaluate Eqs. (2.18), (2.21) and (2.24), the specific form of spectral density function and electron Green's function $\mathrm{Im}G^R_{l_i}$ need to be known. To evaluate the spectral density function, the equation of motion for the electrons under both the periodic potential of the lattice and the scattering with phonons needs to be solved. Moreover, in heavily doped semiconductor materials, or in the case of low or high injections condition, the number of electrons in conduction band could be very large, in which case the electron-electron interaction would be significant as well. This many-body problem, however, could be sufficiently described by the total longitudinal dielectric function of the system. As a result, the electron Matsubara functions in Eq. (A.32) and (A.33) is expanded by using Dyson's equation with a suitable screened dielectric constant model and electron-phonon interaction potential. As a consequence, one can get a closed form of Matsubara Green's functions $g_{2,3}$ and $g_{4,1}$ (Eq. (A.32) and (A.33)) when only the first order term in the expansion is retained. Since the procedure is well-established and can be found in most textbooks on Green's function

theory (see Appendix A and Chapter 5 of Ref. (Mahan, 1993) and Ref. (Bardyszewski and Yevick, 1985)), we only present the results here. Specifically, we have:

$$\mathrm{Im}G^R_{l_i}(\mathbf{k},E) = -\frac{1}{\pi} \times \frac{\mathrm{Im}\Sigma_i(\mathbf{k},E)}{\left[E - E_i(\mathbf{k}) - \mathrm{Re}\Sigma_i(\mathbf{k},E)\right]^2 + \left[\mathrm{Im}\Sigma_i(\mathbf{k},E)\right]^2} \tag{2.34}$$

and,

$$\begin{aligned}\mathrm{Im}\Sigma_i(\mathbf{k},E) = -\pi \sum_\nu \int \frac{\mathrm{d}\mathbf{k}'}{(2\pi)^3} \left|g_\nu(\mathbf{k}'-\mathbf{k})\right|^2 B_{ii}(\mathbf{k},\mathbf{k}') \\ \times \left\{\left[1 + P(\hbar\omega_\nu) - \Theta(E-\hbar\omega_\nu)\right]\delta(E - \hbar\omega_\nu - E_i(\mathbf{k}')) \right. \\ \left. + \left[P(\hbar\omega_\nu) + \Theta(E+\hbar\omega_\nu)\right]\delta(E + \hbar\omega_\nu - E_i(\mathbf{k}'))\right\},\end{aligned} \tag{2.35}$$

in which $B_{ii}(\mathbf{k},\mathbf{k}')$ is the overlap integral between the periodic part of the Bloch function of the two electronic states $\mathbf{k}$ and $\mathbf{k}'$ in band i. ω_ν, $P(\hbar\omega_\nu)$ are the angular frequency and distribution function of phonon respectively. $\Sigma_i(\mathbf{k},E)$ is the retarded electron self-energy for band i, at wavevector $\mathbf{k}$ and energy E, obtained by summing all the possible phonon scatterings from each initial electron state $\mathbf{k}'$ in Brillouin zone. The real part of $\Sigma_i(\mathbf{k},E)$ which gives the shift of electron energy under the effect of phonon-electron interaction has already been included in the band from the non-local EPM model, thus no other effort is needed. The imaginary part of Σ_i, on the contrary, is at the center of our interest since it is directly related to the broadening of each band in the presence of phonons. As shown in Eq. (2.18), all the effects of phonon-electron interaction have been contained in the spectral density function A_i, and computing $\mathrm{Im}\Sigma_i(\mathbf{k},E)$ represents the most challenging part of the model. In evaluating the spectral function, we have not included the realistic phonon dispersion. To describe the optical phonons we have employed a constant frequency obtained at $\mathbf{q} = 0$. For the acoustic phonon branches we have used a linear dispersion and the sound velocity. In the case of non-equilibrium conditions, as is in the calculation

of excess carrier lifetime, quasi-Fermi distributions are used for electrons and holes with respective quasi-Fermi levels. In Eq. (2.35), g_ν is the matrix element for different electron-phonon interaction mechanisms. We have employed four most important forms in our calculation, which are given by:

$$|g_1(\mathbf{q})|^2 = \frac{\Xi_d^2 \hbar\omega_{ac}(\mathbf{q})}{2c_l} \frac{q^4}{(q^2+\lambda^2)^2} \tag{2.36}$$

$$|g_2(\mathbf{q})|^2 = \frac{e^2 \hbar\omega_{ac}(\mathbf{q})}{4\epsilon} K_{AV}^2 \frac{q^2}{(q^2+\lambda^2)^2} \tag{2.37}$$

$$|g_3(\mathbf{q})|^2 = \frac{\hbar D^2 v_s^2}{2\omega_{op}(\mathbf{q})\bar{c}} \frac{q^4}{(q^2+\lambda^2)^2} \tag{2.38}$$

$$|g_4(\mathbf{q})|^2 = \frac{e^2 \hbar\omega_{op}(\mathbf{q})}{2\epsilon^*} \frac{q^2}{(q^2+\lambda^2)^2} \tag{2.39}$$

where c_l and c_t are the longitudinal and transverse elastic constant respectively, $\bar{c}$ is the average elastic constant $\bar{c} = c_l/3 + 2c_t/3$, v_s is the sound velocity and ϵ^* is the effective inverse dielectric constant $\epsilon^{*-1} = \epsilon_\infty^{-1} - \epsilon_0^{-1}$. The phonon occupation factor $P(\hbar\omega_\nu)$ is given by the Bose-Einstein distribution, which is $P(\hbar\omega_\nu) = 1/[1 + e^{\hbar\omega_\nu/(k_B T)}]$. The strengths of these scattering are defined as Ξ_d, the effective deformation potential; D, the optical coupling constant and K_{AV}, average electromechanical coupling coefficient. For the elemental material, such as Si and Ge, which are centrosymmetric crystal, $K_{AV} = 0$. For the compound material systems that do not posses this centrosymmetric property, the piezoelctric scattering is non-zero and its contribution to the self-energy should be included.

As to the electron-phonon interaction matrix element, g_1 represents acoustic phonon scattering, while g_2 is the piezoelectric scattering. The third and fourth terms are for the non-polar optical scattering and polar optical scattering in which we adopted the simple Frohlich Hamiltonian to describe the interaction between electrons and optical phonons. The electron-electron and phonon-phonon scattering effect is included by means of the inverse screening length λ, and the Thomas-Fermi model is applied as a

first order correction. Unlike the Auger recombination, which itself is the result of this screened Coulomb interaction requiring a more rigorous description of λ (Takeshima, 1982a), in the context of radiative recombination, the Thomas-Fermi model is usually a good approximation under the conditions of low injection. In this way, by including the effect of electron-electron and phonon-phonon scattering into the screening length, we can treat each electron as a 'single' electron surrounded by phonons, thus the Frohlich Hamiltonian, which describes a system containing single electron and phonons, is a suitable description. Eq. (2.35) is then integrated over the whole Brillouin zone on a band-dependent tetrahedral mesh (as is discussed in next section). By inserting the resultant spectral density function A_i into Eq. (2.18), the absorption coefficient can be obtained recognizing the fact that the delta function will vanish after the integration of E_2' in Eq. (2.18), thus avoiding the complicated integration on the equi-energy surface.

2.5 3D Device Simulation Models

In the section, models used in the 3D semiconductor device simulations will be briefly summarized. In particular, the Boltzmann Transport Equation (BTE), the governing equation for carrier transport in semiconductors, is discussed first. Two derived simulation models, namely the drift-diffusion model and Monte-Carlo model, which approximate the BTE on different assumptions, are described subsequently.

2.5.1 Boltzmann Transport Equation

Within the framework of the semi-classical transport theory, the BTE describes the motion of carriers in phase space and can be considered as a continuity equation for the carrier distribution function $f(\mathbf{r}, \mathbf{k}, t)$ in phase space. The distribution function is defined by three momentum coordinates $\mathbf{k}$ and three spatial coordinates $\mathbf{r}$ at time t. In bulk material, where the force field $\mathcal{E}$ does not depend on the wavevector $\mathbf{k}$ and the carrier velocity $\mathbf{v}$ is independent of space coordinate, the BTE can be written as

in the form of:

$$\frac{\partial f(\mathbf{r},\mathbf{k},t)}{\partial t} = -\mathbf{v}\cdot\nabla f(\mathbf{r},\mathbf{k},t) - \frac{\mathcal{E}}{\hbar}\cdot\nabla_k f(\mathbf{r},\mathbf{k},t) + \left.\frac{\partial f(\mathbf{r},\mathbf{k},t)}{\partial t}\right|_{coll}. \tag{2.40}$$

Due to the fact that BTE treats electrons and holes as point particles with definite position and momentum at the same time, this transport theory is valid only in the classical domain. Carriers are also considered to be uncorrelated so that a single particle distribution function can be defined. The collision term in Eq. (2.40) is given by:

$$\left.\frac{\partial f(\mathbf{r},\mathbf{k},t)}{\partial t}\right|_{coll} = \int \{f(\mathbf{r},\mathbf{k}',t)[1-f(\mathbf{r},\mathbf{k},t)]\,S(\mathbf{k}',\mathbf{k}) - f(\mathbf{r},\mathbf{k},t)[1-f(\mathbf{r},\mathbf{k}',t)]\,S(\mathbf{k},\mathbf{k}')\}\,d\mathbf{k}' \tag{2.41}$$

where $S(\mathbf{k},\mathbf{k}')$ is the scattering rate from the state $\mathbf{k}$ to $\mathbf{k}'$ due to different microscopic scattering mechanisms, such as phonon, impurity, dislocation scatterings, and GR processes. We noted that the collision integral in Eq. (2.41) contains the quantum description of the system, since the scattering rates are evaluated using quantum mechanical techniques. As a result, the BTE model used in this work is actually a semi-classical model. Moreover, it is clear that the description of the band structure also enters in the BTE through the velocity computed as $\mathbf{v} = \nabla_k E(\mathbf{k})/\hbar$, where the momentum-energy relation for the particles is given by $E(\mathbf{k})$.

The task of solving BTE is challenging due to the presence of many obstacles. Among them are the complex description of scattering rates $\left.\frac{\partial f}{\partial t}\right|_{coll}$, the nonlinear response, and the boundary conditions of devices. Some analytical techniques are available in the literature. For example, assumptions can be made about the form of the distribution function, and it is possible to use the drifted Maxwellian formulation (Hansch, 1991) to obtain the carrier distribution functions. Another possibility relies on the functional expansion of the distribution function in terms of spherical

harmonics (Ventura et al., 1992). Although useful, these analytical models only provide a limited answer to the study of transport in realistic semiconductors since analytical band structures must be used.

On the other hand, when applying the relaxation time approximation (RTA) and using the method of momentum (Hess, 1988), it is possible to separate the BTE into two decoupled equations and the drift-diffusion model can be derived, which is shown in the following section. Monte Carlo method is another powerful tool in determining the solutions of BTE. Rather than directly solving the BTE, the distribution function is built through direct simulation of the motion of an ensemble of particles in phase space. The ensemble is subjected to the action of external forces $\mathcal{E}$ and scattering phenomena. Detailed simulation procedures can be found in Section 2.5.3.

2.5.2 Drift-Diffusion Model

To avoid evaluating the collision term $\frac{\partial f}{\partial t}\big|_{coll}$ in Eq. (2.40), which requires the solving of Schrodinger equation for different microscopic processes, it is convenient to apply the RTA to obtain an overall effect of all the scattering mechanisms. (Hess, 1988) The RTA assumes that the action of inter-particle collisions relaxes the non-equilibrium distribution function f to its value at equilibrium f_0 in a characteristic time τ, which effectively transforms the BTE into a mathematically tractable differential equation:

$$\left.\frac{\partial f}{\partial t}\right|_{coll} = -\frac{f - f_0}{\tau}. \tag{2.42}$$

When further simplify the non-equilibrium distribution function f with its averaged value by applying the method of momentum and integrating over momentum space (Hess, 1988; Brennan, 1999):

$$\int \Theta \frac{\partial f}{\partial t} \mathrm{d}\mathbf{k} = -\int \Theta \left(\mathbf{v} \cdot \nabla f\right) \mathrm{d}\mathbf{k} - \int \Theta \left(\frac{\mathcal{E}}{\hbar} \cdot \nabla_k f\right) \mathrm{d}\mathbf{k} - \int \Theta \frac{f - f_0}{\tau} \mathrm{d}\mathbf{k}, \tag{2.43}$$

Eq. (2.40) can be decoupled into several parts with proper choice of the momentum function Θ. First let $\Theta = 1$. Each term in Eq. (2.43) becomes:

$$\int \Theta \frac{\partial f}{\partial t} \mathrm{d}\mathbf{k} = \frac{\partial}{\partial t} \int f(\mathbf{r}, \mathbf{k}, t) \mathrm{d}\mathbf{k} = \frac{\partial n}{\partial t}, \tag{2.44}$$

with n being the carrier density in real space, and

$$\int \Theta \frac{f - f_0}{\tau} \mathrm{d}\mathbf{k} = \frac{n - n_0}{\tau}, \tag{2.45}$$

$$-\int \Theta \left(\mathbf{v} \cdot \nabla f\right) \mathrm{d}\mathbf{k} - \int \Theta \left(\frac{\mathcal{E}}{\hbar} \cdot \nabla_k f\right) \mathrm{d}\mathbf{k} = \nabla \cdot (n\mathbf{v}). \tag{2.46}$$

In the derivation above, partial integration is used: $\mathbf{g} \cdot \nabla h = \nabla \cdot (h\mathbf{g}) - h \nabla \cdot \mathbf{g}$. Eq. (2.43) therefore is simplified as:

$$\frac{\partial n}{\partial t} = \nabla \cdot (n\mathbf{v}) - \frac{n - n_0}{\tau}. \tag{2.47}$$

Eq. (2.47) is the well-known continuity equation.

Further evaluating the second momentum of Eq. (2.43) by setting $\Theta = \mathbf{v}$, following the similar partial integration procedure, one can get:

$$\mathbf{J}_n = e\mu_n n\mathcal{E} + eD_n \nabla n, \tag{2.48}$$

with

$$\mu_n = \frac{e\tau}{m} \tag{2.49}$$

$$D_n = \mu_n \frac{k_B T}{e} \tag{2.50}$$

$$\mathbf{J}_n = -en\mathbf{v}, \tag{2.51}$$

in which e is the absolute value of electron charge, and m is the electron mass.

The final piece of the drift-diffusion model is the Poisson equation relating the

system's free and fixed charge to the electrostatic potential created by its displacement:

$$-\nabla(\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})) = e(N_D + p - n - N_A). \tag{2.52}$$

Here, the electrostatic potential is ϕ while N_D and N_A represent the concentration of positively and negatively charged donor and acceptor atoms, respectively.

Eq. (2.47) - (2.52) represent the complete drift-diffusion model which can be used to simulate carrier and electric field distribution in a self-consistent manner.

2.5.3 Monte-Carlo Model

Monte-Carlo simulation is another popular method of solving the BTE. The central idea of Monte-Carlo simulation is to simulate the motion of a particle ensemble in a device structure under specific external field and boundary conditions. When it comes to the scattering process, instead of using RTA to simplify the collision term in Eq. (2.40), quantum mechanical models for different scatterings are used to evaluate the scattering rates as functions of particle energy. Due to the probabilistic nature of scattering event in the real semiconductor devices, the simulated scattering behavior of particles is determined in a stochastic way, where both the type of scattering and the final state of particles in the momentum space are chosen randomly (Jacoboni and Lugli, 1989). The whole Monte-Carlo simulation model, therefore, can be divided into two parts: one is the motion of particles in the real space which is drifted by the local electric field, the other is the scattering processes which happened in the momentum space. The following two sections will discuss these two topics respectively.

Free Flight and Self-Scattering

The wavevector $\mathbf{k}$ of a carrier (electron or hole) changes continuously during the free flight due to the presence of electric field, which is:

$$\hbar\dot{\mathbf{k}} = -e\mathcal{E}. \tag{2.53}$$

One assumption of MC simulation is that between two consecutive collisions (scattering), particles can fly freely in the electric field generated by the local charge distribution. The free flight makes particles move in the real space while the scattering events change the particle velocity, thus move the particle in the momentum space. Apparently, the free flight time is crucial to correctly simulate the particles' trajectory which will determine the both carrier distributions and the microscopic current in the devices.

The free flight time can be assessed through the probability of the scattering events. Let $P[\mathbf{k}(t)]$ be the probability that an electron in the state $\mathbf{k}$ undergoes a collision in the time interval of $\mathrm{d}t$. The probability of this electron that had a collision at $t = 0$ yet has not subject to a subsequent collision at time t can be written as: $\exp\left[-\int_0^t P[\mathbf{k}(t')]\mathrm{d}t'\right]$. As a result, the probability of an electron subject to a collision at time t is given by:

$$\mathscr{P}(t) = P[\mathbf{k}(t)] \exp\left[-\int_0^t P[\mathbf{k}(t')]\mathrm{d}t'\right]. \tag{2.54}$$

The scattering probability for each scattering mechanism $P[\mathbf{k}(t)]$ can be obtained by using quantum mechanical theory (Ridley, 1999), which can only be evaluated numerically if a full band structure is adopted. Furthermore, the integration of $P[\mathbf{k}(t')]$ over time interval $[0t]$ on the exponential makes Eq. (2.54) even harder to use. Rees (Rees, 1968; Rees, 1969) has proposed a simpler way to avoid solving Eq. (2.54)

directly, in which the maximum value of $P[\mathbf{k}(t)]$ in momentum space is assumed to be $\Gamma \equiv 1/\tau_0$. We use this maximum probability Γ to replace the $P(\mathrm{k})$ and Eq. (2.54) can be converted to:

$$\mathscr{P}(t) = \frac{1}{\tau_0} \exp(-t/\tau_0). \tag{2.55}$$

By introducing Rees's method, we actually have increased the total probability of $\mathscr{P}(t)$ and the realistic scattering probability $P(\mathbf{k})$ is replaced by a constant value $\Gamma \geq P(\mathbf{k}), \forall\, \mathbf{k}$. To maintain the physical correctness of Eq. (2.54), the extra probability that goes into Eq. (2.55) can be attributed to a fictitious "self-scattering" where the particle will scatter with itself and maintain its original velocity (momentum $\mathbf{k}$) after the scattering.

To obtain the total free flight after one real scattering event, integrate both sides of Eq. (2.55) from 0 to t_r. The LHS $\int_0^{t_r} \mathscr{P}(t')\mathrm{d}t' \equiv r$ gives the probability of a free flight that lasts t_r time. The result $t_r = -\tau_0 \ln(1-r)$ can be further simplified to

$$t_r = -\tau_0 \ln(r), \tag{2.56}$$

since r is assumed to be uniformly distributed within [0 1]. It is worthy of note that the assumption of uniformly distributed r is the consequence of the constant scattering rate Γ used in Eq. (2.55). At the end of each free flight where a scattering is supposed to happen, self-scattering will be further checked by using another uniform random number r'. The comparison of r' with the normalized scattering probability at $\mathbf{k}$, $P(\mathbf{k})/\Gamma$, will indicate either 1: a real scattering event if $r' \leq P(\mathbf{k})/\Gamma$ or 2: self-scattering if otherwise. Therefore, Eq. (2.54) can be properly handled, in which a real scattering only happens with a probability of $P(\mathbf{k})$.

As a final remark, one should notice that when using the obtained free flight time t_r in the simulation, different particles generally have different flight time which makes it difficult to track the properties of each particle at the same simulation time.

Consequently, it is advisable to discard the concept of free flight time and evaluate Eq. (2.54) after each small constant time step using the self-scattering method (Rees, 1968; Rees, 1969). This will effectively eliminated the need to compute the free flight time at the expense of more times of checking for the self-scattering event.

Scattering Processes

The scattering rate as a function of particle's initial $\mathbf{k}$ and final states $\mathbf{k}'$ in momentum space with a perturbation Hamiltonian H_p can be obtained using first-order perturbation and Fermi's Golden Rule:

$$S(\mathbf{k},\mathbf{k}') = \frac{2\pi}{\hbar}\left|\langle\psi_{\mathbf{k}}|\, H_p\, |\psi_{\mathbf{k}'}\rangle\right|^2 \delta[E(\mathbf{k}) - E(\mathbf{k}')] \tag{2.57}$$

for elastic processes, and

$$S(\mathbf{k},\mathbf{k}') = \frac{2\pi}{\hbar}\left(N_q + \frac{1}{2} \pm \frac{1}{2}\right)\left|\langle\psi_{\mathbf{k}}|\, H_p\, |\psi_{\mathbf{k}'}\rangle\right|^2 \delta[E(\mathbf{k}) - E(\mathbf{k}') \mp \hbar\omega_q] \tag{2.58}$$

for phonon scattering processes. In Eq. (2.58) $\hbar\omega_{\mathbf{q}}$ is the phonon energy, N_q is the phonon occupation number at the lattice temperature T, and the upper and lower signs correspond to emission and absorption, respectively. The total rate $1/\tau(\mathbf{k})$ out of the Bloch state $\psi_{\mathbf{k}}$ for each scattering mechanism can be evaluated numerically by integrating over the Brillouin zone using the full-band description

$$\frac{1}{\tau(\mathbf{k})} = \frac{V}{(2\pi)^3}\int S(\mathbf{k},\mathbf{k}')\,\mathrm{d}\mathbf{k}' \tag{2.59}$$

The Fermi's Golden Rule is derived under the assumption that the scattering event is instantaneous. This assumption breaks down in a few important cases. In case of very fast transients, where the dynamics of the system has to be studied on a time scale comparable to the scattering rate, the rates have to be derived in a different way. In case of very high scattering rates, the initial state may decay appreciably

by the time the scattering is completed, and the determination of the rate has to be done with this taken into account. Lastly, in case of very high electric fields, the field can transfer significant amount of energy to the particle during the scattering events (intracollisional field effect) which changes the assumption of the Golden Rule. Different attempts have been made to include such phenomena in transport simulators but no unified treatment is presently available. In cases where these effects can be neglected, the first order transition rate provides a good description of the scattering events. Each scattering process can then be described by specializing Eq. (2.57), (2.58) using an appropriate expression for the perturbation Hamiltonian.

Here, the final expressions for polar scattering, deformation potential scattering, impurity scattering and dislocation scattering derived using Fermi's Golden Rule are presented for completeness. The detailed derivation can be found in Ref. (Ridley, 1999).

- Polar optical phonon scattering rates, described by using the Fröhlich formulation:

$$S(\mathbf{k},\mathbf{k}') = \frac{(2\pi e)^2\omega_{op}}{4\pi V}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_0}\right) \times\left\{N_q+\frac{1}{2}\pm\frac{1}{2}\right\}\frac{|I(\mathbf{k},\mathbf{k}')|^2}{q^2}\,\delta[E(\mathbf{k})-E(\mathbf{k}')\mp\hbar\omega_{\mathrm{LO}}] \tag{2.60}$$

  where ϵ_∞, ϵ_0 are the optical and static dielectric functions, $I(\mathbf{k},\mathbf{k}')$ is the overlap integral between the initial $\mathbf{k}$ and final $\mathbf{k}'$ Bloch state, $\hbar\omega_{\mathrm{LO}}$ is the longitudinal optical phonon energy and V the volume of the crystal. A common approximation is the use of a constant value for the polar optical energy associated with the longitudinal optical mode. The dependence of the matrix element on the inverse square of the phonon wave vector, and the angular dependence of the overlap integral makes this scattering mechanism anisotropic.

- Non-polar acoustic phonon scattering rates:

$$S(\mathbf{k},\mathbf{k}') = \frac{\pi D_a^2}{V\rho u}\left\{N_q + \frac{1}{2} \pm \frac{1}{2}\right\} I(\mathbf{k},\mathbf{k}') q\, \delta[E(\mathbf{k}) - E(\mathbf{k}') \mp \hbar\omega_{\mathbf{q}}] \tag{2.61}$$

where D_a is the isotropic coupling constant, ρ is the density of the semiconductor, $\omega_{\mathbf{q}}$ is the frequency of the phonon with wavevector $\mathbf{q}$, and $\mathbf{k}' = \mathbf{k} \mp \mathbf{q} + \mathbf{G}$ is the final electron wavevector which is mapped into the first BZ by adding a vector $\mathbf{G}$ of the reciprocal lattice. The acoustic phonon dispersion has been approximated by:

$$\hbar\omega_{\mathbf{q}} = \begin{cases} \hbar\omega_{\max}\left[1 - \cos\left(|\mathbf{q}|a/4\right)\right]^{1/2}, & |\mathbf{q}| \le 2\pi/a \\ \hbar\omega_{\max}, & |\mathbf{q}| > 2\pi/a \end{cases} \tag{2.62}$$

- Non-polar optical phonon scattering rates:

$$S(\mathbf{k},\mathbf{k}') = \frac{\pi D_{TK}^2}{V\rho u}\left\{N_q + \frac{1}{2} \pm \frac{1}{2}\right\} I(\mathbf{k},\mathbf{k}')\, \delta[E(\mathbf{k}) - E(\mathbf{k}') \mp \hbar\omega_{\mathrm{TO}}] \tag{2.63}$$

- Ionized impurity scattering rates, derived using the Brooks-Herring model:

$$S(\mathbf{k},\mathbf{k}') = \frac{32\pi^3 Z^2 N_I e^4}{\hbar V 4\pi\epsilon^2}\frac{|I(\mathbf{k},\mathbf{k}')|^2}{(\beta^2 + q^2)^2}\, \delta[E(\mathbf{k}) - E(\mathbf{k}')] \tag{2.64}$$

where β is the screening length, which depends on the free carrier density, and N_I is the density of ionized impurities of charge Ze.

- Piezoelectric scattering rates:

$$S(\mathbf{k},\mathbf{k}') = \frac{e^2 K_{av}^2 k_B T}{8\pi^2\epsilon\hbar}\frac{q^2}{(q^2 + \beta^2)^2}\delta[E(\mathbf{k}) - E(\mathbf{k}')] \tag{2.65}$$

Piezoelectric coupling constant has a complex directional dependence. Longitudinal and transverse modes are lumped together by defining an average

electromechanical coupling K_{av} such that

$$K_{av}^2 = \frac{\langle e_L^2 \rangle}{\epsilon c_L} + \frac{\langle e_T^2 \rangle}{\epsilon c_T} \tag{2.66}$$

where c_T, c_L are the transverse and longitudinal elastic constants and

$$\langle e_L^2 \rangle = \frac{1}{7} e_{33}^2 + \frac{4}{35} e_{33}(e_{31} + 2e_{15}) + \frac{8}{105}(e_{31} + 2e_{15})^2 \tag{2.67a}$$

$$\langle e_T^2 \rangle = \frac{2}{35}(e_{33} - e_{31} - e_{15})^2 + \frac{16}{105} e_{15}(e_{33} - e_{31} - e_{15}) + \frac{16}{35} e_{15}^2 \tag{2.67b}$$

- Dislocation scattering rates (Bertazzi et al., 2007):

$$\frac{1}{\tau(\mathbf{k}_\perp)} = \frac{N_{dislo}\, q^4 m^* \lambda^4}{\epsilon c^2 \hbar^3} \frac{1 + 2k_\perp^2 \lambda^2}{(1 + 4k_\perp^2 \lambda^2)^{3/2}} \tag{2.68}$$

  where λ is the Debye length, $\mathbf{q} = \mathbf{k}_\perp - \mathbf{k}_\perp{}'$, and $\mathbf{k}_\perp$, $\mathbf{k}'_\perp$ are the incoming and outgoing wavevectors in the direction perpendicular to the growth axis.

2.6 Numerical Implementation

2.6.1 Finite Element Method

FEM is a numerical method for obtaining approximate solutions of system equations over a discretized space. Unlike the Rayleigh-Ritz variational method which tries to find the solution over the entire calculation space, for FEM solutions are computed by minimizing the equation error on each individual elements in a averaged sense (Ram-Mohan, 2002). Due to the fact that FEM imposes less restriction on the form of governing equations and solutions, it is more flexible to use than the variational method especially for large systems with coupled equations.

In general, FEM can be divided into five parts (Huebner et al., 1995):

- Discretize the continuum. The first step is to discretize the continuum simulation space into elements with some specific shape of elements. When choosing the

type of elements to represent the simulation space, one should consider both the shape and symmetry of the space and the nature of the expected solutions. For example in 3D simulation, if the boundary of simulation space is aligned with the Cartesian coordinate axes, it is usually convenient to use brick tensor element in the space discretization. However, if the boundary is irregular with many curved faces, tetrahedron or triangular prism are usually employed. This step also defines the interpolation function that can be used to evaluate the unknowns inside each elements.

- Select interpolation functions. As is stated above, the interpolation functions are used to represent the variation of the field variable over the element. The field variable is the unknown solution to the problem which can be scalar, vector or high-dimensional tensor. In general, ploynomials are preferred due to the ease of evaluation, integration and differentiation to them. The nature of the expected solutions may also imposes restrictions on the interpolation function. For example, if the physical properties of the solution has C^1 continuity, its interpolation function should take this into consideration. Otherwise unphysical solution may be obtained.

- Find the element properties and assemble the system equations. Once the type of element and interpolation functions have been determined, one can choose the way to solve the equations. Usually the weighted residual approach (weak form) is used for its localized error. The connectivity of elements are then used to relate different elements together to form a group of discretized equations.

- Impose the boundary conditions. In order to obtain a definite solution of the system, proper boundary conditions need to be included in the matrix equations. The way to apply the boundary conditions will be briefly introduced

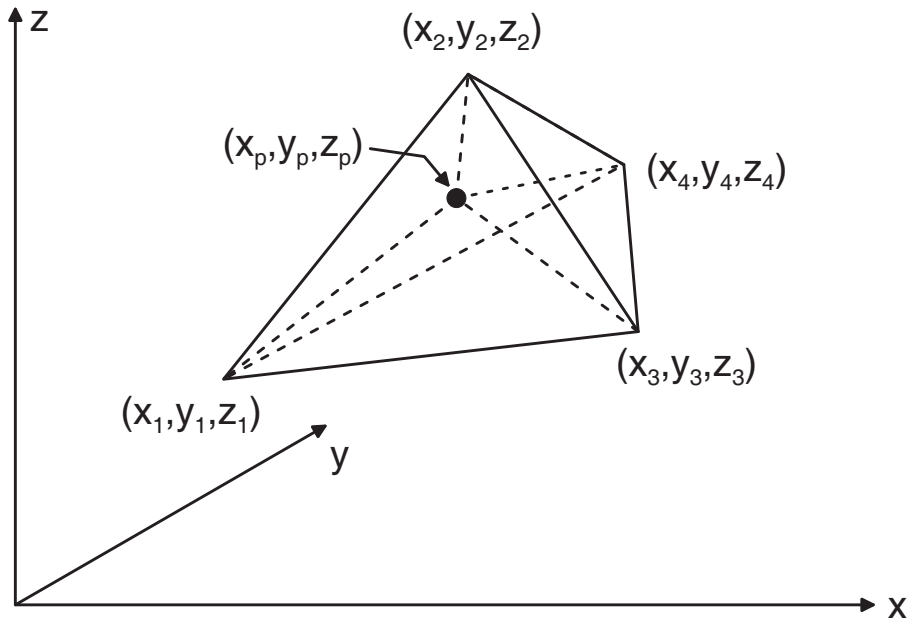


Figure 2·2: A tetrahedral element with its four nodes labeled in a right-hand coordinates. The point (x_1, y_1, z_1) is within the element.

in Section 2.6.3.

- Solve the system equations and make additional calculations if required. Depending on the type of the problem, a set of linear or non-linear equations need to be solved to obtain the solutions of the original problem. Many algorithms have been developed to solve these matrix equations. A good review can be found in Ref. (Hesthaven and Warburton, 2008).

Natural Coordinates in 3D

In this work, tetrahedron element together with linear interpolation function is used to solve the Poisson equation in the MC simulation. As a result, here we only introduce the linear interpolation function suitable for tetrahedron elements using 3D natural coordinates.

The linear interpolation functions in 3D for a quantity ϕ can be represented as:

$$\phi(x, y, z) = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 z. \tag{2.69}$$

in which α_i it some quantities relate to both element shape and the individual $\phi_i, i = 1, 2, 3, 4$ at each vertices. To separate the geometry information of the mesh with the unknown quantity ϕ, Eq. (2.69) can be rewrite as:

$$\phi(x, y, z) = N_1\phi_1 + N_2\phi_2 + N_3\phi_3 + N_4\phi_4. \tag{2.70}$$

with N_i defined as natural coordinates. Since the quantity ϕ is arbitrary, in particular, let ϕ equals to x, y, z, 1, we can obtain four equations which can further be written in the matrix form:

$$\begin{pmatrix} 1 \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \end{pmatrix} \tag{2.71}$$

Here, x_i, y_i, z_i, with $i = 1, 2, 3, 4$ is the coordinates of the four vertices of the tetrahedron element, as shown in Fig. 2·2.

In order to use Eq. (2.70) to interpolate any functions linearly, one only need to inverse the matrix in Eq. 2.72 and get:

$$\begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ z \end{pmatrix} \tag{2.72}$$

in which the coefficients of matrix a_i, b_i, c_i, d_i, with $i = 1, 2, 3, 4$ are functions of tetrahedron vertices coordinates, and are only depended on the geometry of the mesh. When it comes to the derivation of the arbitrary function ϕ, it is easy to derive the

following equations using chain rule:

$$\frac{\partial \phi}{\partial x} = \sum_{i=1}^{4} \frac{\partial \phi}{\partial N_i} \frac{\partial N_i}{\partial x} \tag{2.73}$$

$$\frac{\partial \phi}{\partial y} = \sum_{i=1}^{4} \frac{\partial \phi}{\partial N_i} \frac{\partial N_i}{\partial y} \tag{2.74}$$

$$\frac{\partial \phi}{\partial z} = \sum_{i=1}^{4} \frac{\partial \phi}{\partial N_i} \frac{\partial N_i}{\partial z}. \tag{2.75}$$

We noted that since the coefficients a_i, b_i, c_i, d_i is independent on the problems to be solved, once the discretization of the simulation space is done, these interpolations coefficients can be calculated in advance. When interpolating the field variables in the device simulation phase, Eq. (2.70) can be used directly.

2.6.2 Discretization of 3D Momentum Space

One difficulty in numerically evaluating the Green's function and calculating scattering rates is related to the integration of the delta functions in Eq. (2.19), (2.33), (2.35) and (2.60) to (2.68). which involves the computation of a two-dimensional integral on a constant-energy surface in momentum space (Wen and Bellotti, 2014). Since the full band structures are employed, the lack of an analytical expression for the energy dispersion $E(\mathbf{k})$ makes this integration even more challenging. In this work, we adopted the tetrahedron method where the electron energy $E(\mathbf{k})$ and other $\mathbf{k}$-dependent quantities are linearly interpolated throughout each tetrahedron using the corresponding values at the four vertices. This approach has proven to be efficient and accurate (Lehmann and Taut, 1972; Jungemann and Meinerzhagen, 2003), and, as described in the previous section, can also be applied to real space integration. Although higher order interpolation schemes, such as quadratic interpolation, can provide even more accurate results (MacDonald et al., 1979), their complicated implementation and computational load make the linear interpolation method more

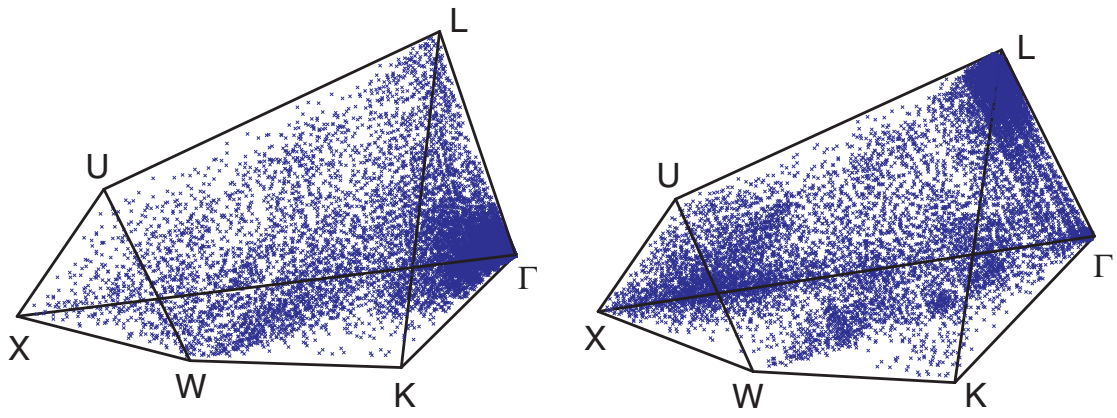


Figure 2·3: Generated mesh in the irreducible wedge of BZ of zincblend structure for (a): direct band gap material and (b): indirect band gap material. In (a), region around Γ sees highest density of mesh nodes, whereas in (b), the density of mesh nodes is highest around L valley. Points in the figure represent the geometry center of each tetrahedron element to ease the visualization.

appealing. To reduce the error introduced by the numerical implementation, while maintaining a reasonable computation time, instead of using a uniform mesh, we employed a band-adaptive non-uniform mesh (Jungemann and Meinerzhagen, 2003) by considering the specific features of each band involved in a transition, where a dedicated tetrahedron mesh is produced to optimize the energy interpolation.

For example, in the work of evaluating the optical and Auger properties of HgCdTe, a tetrahedron mesh with 22,887 elements in the irreducible wedge (IW) is used to perform the integration of Eqs. (2.18), (2.21) and (2.24), and another mesh with 8,422 tetrahedrons is employed in the evaluation of Eq. (2.35) due to the different types of integrands in these equations. For other materials, such as Si, Ge, InGaAs and InAsSb, separate sets of meshes are generated for each material respectively to accommodate their full band structures. As shown in Fig. 2·3, meshes for direct band gap material (a, e.g. HgCdTe, InGaAs and InAsSb) and indirect band gap material (b, e.g. Si and Ge) are demonstrated, which has the highest node density around the band edge.

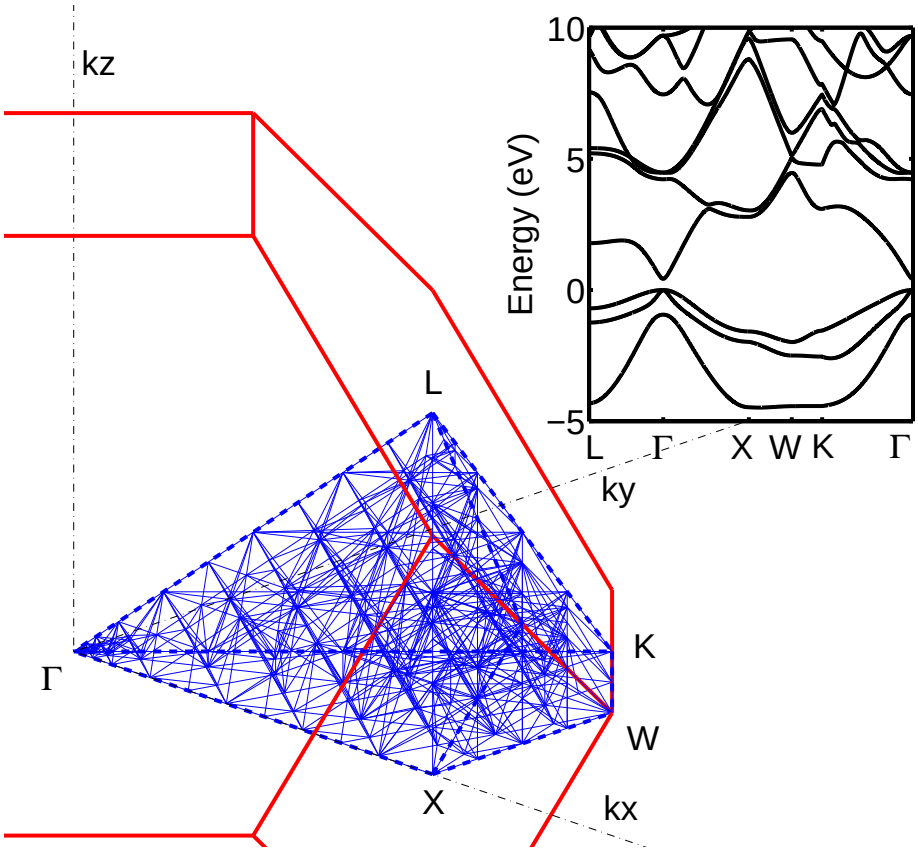


Figure 2·4: The tetrahedral mesh for the first conduction band of $Hg_{0.6}Cd_{0.4}Te$ with 376 vertices in the IW. The boundaries of Brillouin zone and IW are depicted as bold lines. The inset of the figure is the corresponding band structures from the non-local EPM calculation, which is used in the mesh generation and later calculation.

As a demonstration, Fig. 2·4 shows a simplified tetrahedron mesh generated for $Hg_{0.6}Cd_{0.4}Te$ with 376 vertices in the IW. We want to point out that in the realistic calculation, meshes with much higher density is adopted to reduce the interpolation error. For instance, the calculation of density of states involving 4 valence bands and 8 conduction bands, which is necessary to evaluate the realistic quasi-Fermi levels for specific doping/injection conditions, are evaluated on a 35,299 tetrahedron mesh. The convergence of the integration has been checked by using meshes of different sizes. For example, a mesh of 32,560 tetrahedrons in IW provides essentially the same result as

the mesh of 22,887 tetrahedrons with a maximum difference always below 5%.

2.6.3 Discretization of Poisson Equation

As shown in Section 2.5.3, electric field in the semiconductor devices is essential in the MC simulation, since it determines the particles' velocity and consequently their distribution in the real space. To obtain the electric field, Poisson equation (Eq. (2.52)) need to be solved, and the gradient of the resulting electrostatic potential gives the electric field. Due to the fact that the simulation domain is discretized using tetrahedron elements in this work, FEM becomes the natural choice of solving Poisson equation. To include the space charges carried by the particles, Eq. (2.52) can be rewrite as:

$$-\nabla(\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})) = \sum_p q_p\delta(\mathbf{r}-\mathbf{r}_p) + \rho(\mathbf{r}), \tag{2.76}$$

with

$$\phi|_{\partial\Omega_{Di}} = \phi_D + V_{ext,i}, \tag{2.77}$$

$$\nabla_{\mathbf{n}}\phi|_{\partial\Omega_N} = 0. \tag{2.78}$$

as the boundary conditions. Here, q_p is the charge carried by each particles, $\mathbf{r}_p$ is the position of the particle and $\rho(\mathbf{r})$ represents the continuum charges in the space such as the doping concentration. Eq. (2.77) and (2.78) gives the Dirichlet BC and Neumann BC respectively, in which the electrostatic potential at the ith ohmic contact region $\partial\Omega_{Di}$ is the sum of Fermi level ϕ_D in the device at equilibrium and the applied bias voltage $V_{ext,i}$.

Following the discussion in Section 2.6.1, with the linear interpolation function N_i and the potential defined at each node in the mesh ϕ_i, the electrostatic potential in

the entire simulation domain Ω can be represented as:

$$\phi(\mathbf{r}) = \sum_{j=1}^{K} N_j(\mathbf{r})\phi_j(\mathbf{r}) \tag{2.79}$$

assuming there are K nodes in the discretized mesh. One needs to notice that in the framework of FEM, $N_j(\mathbf{r})$ fulfills:

$$N_j(\mathbf{r}) = \begin{cases} \text{nonzero}, & \text{if } \mathbf{r} \in \text{ element that contains node } j \\ 0, & \text{if } \mathbf{r} \notin \text{ element that contains node } j \end{cases} \tag{2.80}$$

The ϕ_i in Eq. (2.79) is the solution of Poisson equation we are seeking since with that information, the potential at any point $\mathbf{r}$ can be found through Eq. (2.79).

Using the Ritz-Galerkin method (Ram-Mohan, 2002), and combine Eq. (2.76), (2.79) and (2.80), one can get the weak form of Poisson equation:

$$-\epsilon \sum_{j=1}^{K} \phi_j \int_{\Omega} T(\mathbf{r}) \nabla^2 N_j \, \mathrm{d}\mathbf{r} = \sum_{p} \int_{\Omega} T(\mathbf{r}) q_p \delta(\mathbf{r} - \mathbf{r}_p) \, \mathrm{d}\mathbf{r} + \int_{\Omega} T(\mathbf{r}) \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r}. \tag{2.81}$$

Here, the testing function in Galerkin's method is defined as $T(\mathbf{r})$ and the error of introducing this arbitrary function into the rigorous Poisson equation is zero when averaged on all the elements.

For all $i = 1, 2, 3...K$, let $T(\mathbf{r}) = N_i(\mathbf{r})$, which is the standard procedure of Galerkin's method, Eq. (2.81) can be converted into:

$$\epsilon \sum_{j=1}^{K} \phi_j \int_{\Omega} \nabla N_j \cdot \nabla N_i \, \mathrm{d}\mathbf{r} = \sum_{p} q_p N_i(\mathbf{r}_p) + \sum_{j=1}^{K} \int_{\Omega} \rho_j N_j(\mathbf{r}) N_i(\mathbf{r}) \, \mathrm{d}\mathbf{r}, \quad i = 1 \text{ to } K. \tag{2.82}$$

In this equation, the usual partial integration technique is used and Neumann BC Eq. (2.78) is applied to eliminate the surface integral over $\partial\Omega$. The continuum charge distribution also has been linearly interpolated as $\rho(\mathbf{r}) = \sum_{j=1}^{K} \rho_j N_j(\mathbf{r})$, and the exact form of the last term in Eq. (2.82) is derived in Appendix B.

Eq. (2.82) is ready to be assembled into the matrix form for a linear solver to solve numerically:

$$\mathbf{A}\Phi = \mathbf{b}, \tag{2.83}$$

with

$$A_{ij} = \int_{\Omega} \nabla N_i \cdot \nabla N_j \, \mathrm{d}\mathbf{r}, \quad i, j = 1, 2, 3, \ \ldots K \tag{2.84}$$

$$\Phi = (\phi_1, \phi_2, \ \ldots, \phi_K)^T. \tag{2.85}$$

We noted that the RHS of Eq. (2.83) has already included the Dirichlet BC (Eq. (2.77)) and the RHS of Eq. (2.82). When applying the Dirichlet BC to Eq. (2.83), besides setting the corresponding potential ϕ_i in Eq. (2.85) to the fixed value, the cross "benediction" (Ram-Mohan, 2002) to the matrix $\mathbf{A}$ also needs to be performed and added to the RHS. Though we have outlined the basic steps of discretizing and solving Poisson equation using FEM, detailed procedures of calculating $\mathbf{A}$ and $\mathbf{b}$ will be skipped here for the sake of simplicity.

Chapter 3

Full Band Structures and Material Constants for Narrow-Gap Materials

Using the theoretical framework of EPM outlined in Chapter 2, Section 2.1, full band structures for both elemental and compound materials investigated in this work are fitted with strain effect enabled.

3.1 Strain Effect on the Semiconductor Material Properties

Strain appears in the semiconductor material when the material is grown on a lattice-mismatched substrate or subjects to external forces such as high pressure gas (Boztug et al., 2013). Depending on the device fabrication processes, usually, uniaxial (Süess et al., 2013) or biaxial strain (Boztug et al., 2013) can be applied to the material. In this work, only the balanced biaxial strain, which is generally induced by the different lattice constants of semiconductor materials and its substrates, will be investigated for its importance and widespread existence.

Basically, there are two major effects of strain on the material properties that are significant in our calculation. The first is the change of the shape of the BZ for the material. As is shown in Fig. 3·1 in which the BZ of zincblend structure material is plotted together with its irreducible wedge, when a balanced biaxial strain is applied in the xy plane, the irreducible wedge (outlined by green and blue lines) is deformed and becomes three times larger than the one in the relaxed material (blue lines). This can be understood by considering the symmetry of the strained material. For example,

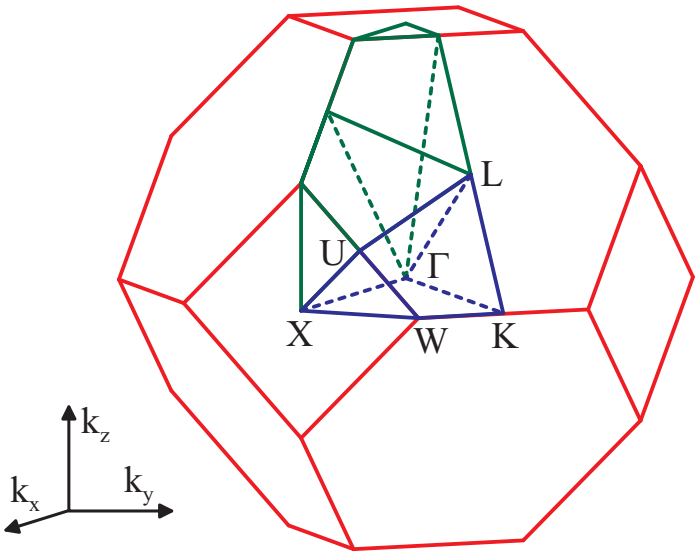


Figure 3·1: BZ of zincblend structure material. The irreducible wedge for relaxed material is shown in blue lines. Green lines represent the additional region that will be included in the IW if biaxial strain is applied in xy plane. Letters in the figure indicate the high-symmetry point in the BZ.

when the material is grown on a substrate (assuming z is the growth direction) that has smaller lattice constant than the material's, atoms in the xy plane will be squeezed, and therefore the length of reciprocal vector of the strained material $\mathbf{G}_x, \mathbf{G}_y$ will be larger than $\mathbf{G}_z$. Consequently, the reflection symmetry regards to the plane $z = y$ and $z = x$ will disappear, and the IW will becomes three times as large as the relaxed IW. In Fig. 3·1, the blue lines show the IW for a relaxed material whereas the green lines plotted the additional region for IW when balanced biaxial strain presents in xy plane of crystal.

The second effect is the change of full band structures. As an example, in Fig. 3·2, EPM band structure of tensilely strained Ge is plotted around Γ point. As is shown, the resulting band gap $(\Gamma_c - \Gamma_v)$ energy reduced to 0.47 eV compared to the relaxed direct gap of 0.9 eV and relaxed indirect gap $(L_c - \Gamma_v)$ of 0.74 eV, effectively turns the Ge into a direct band gap material. Moreover, with tensile strain, LH and HH valence

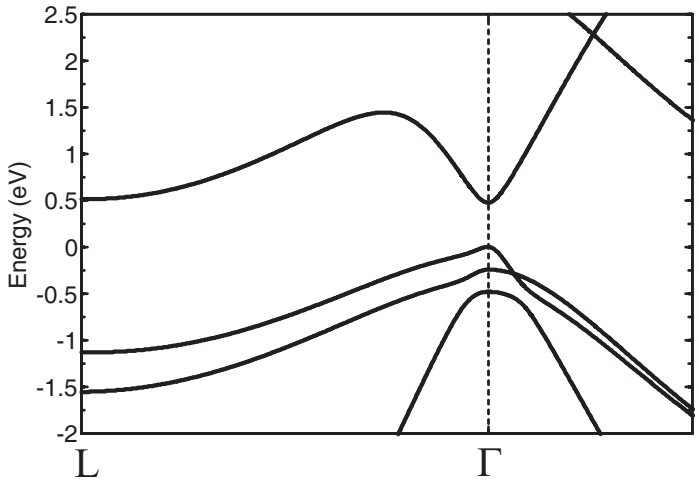


Figure 3·2: EPM band structure of 2.7% tensilely strained Ge around L and Γ valley. Band HH and LH are mixed under the biaxial tensile strain.

band mixed around Γ point, in which energy of LH band becomes higher than HH energy at the vicinity of Γ valley. Since the density of state for LH band is much smaller than HH band, this phenomenon will have significant impact on the recombination processes in the material due to the fact that all the recombination processes depend on the number of available final states in valence band, as is demonstrated in Section 2.2 and Section 2.3.

3.2 Fitting Results for Narrow-Gap Materials

This section reports the fitted band structures for all the narrow-gap materials to be investigated in the following chapters. Both pseudopotentials and some fitting parameters for EPM will be demonstrated.

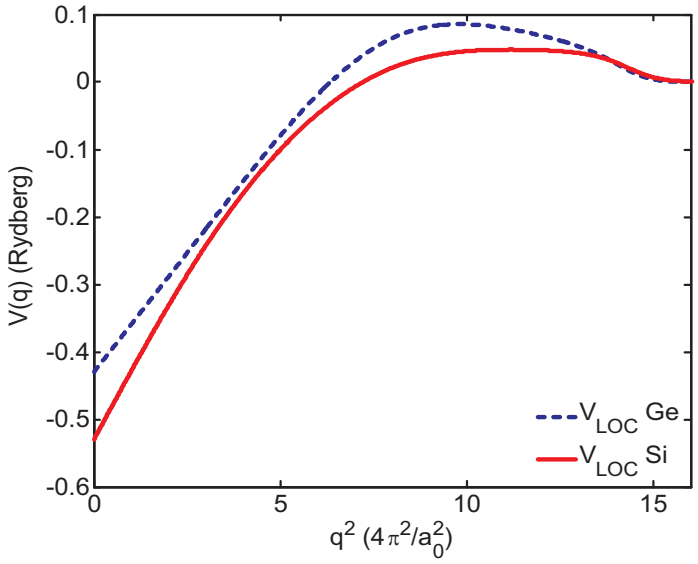


Figure 3·3: Wavevector-dependent screened local pseudo-potentials employed in the calculation of Si and Ge electronic structure. a_0 is the lattice constant.

Table 3.1: Optimized parameters of the local pseudo-potential in EPM model for Si and Ge.

Parameter	Unit	Ge	Si
V_0	Ry	-0.5303	-0.7000
V_3	Ry	-0.2425	-0.2177
V_8	Ry	0.0210	0.06484
V_{11}	Ry	0.0479	0.07982
V_{19}	Ry	0.0	0.0
γ	-	1.0	1.0
μ	-	0.00142	0.0115

3.2.1 Silicon, Germanium, α-Tin and Their Alloys

Silicon and Germanium

The parameters of the local EPM pseudo-potentials and the spin-orbit corrections employed in the calculation of Si and Ge band structures are reported in Table 3.1. The functional dependence of the screened atomic potentials on the wavevector is numerically described by a spline interpolation, which is presented in Fig. 3·3 for a specific case. Notice that the features of the interpolation depend on the behavior

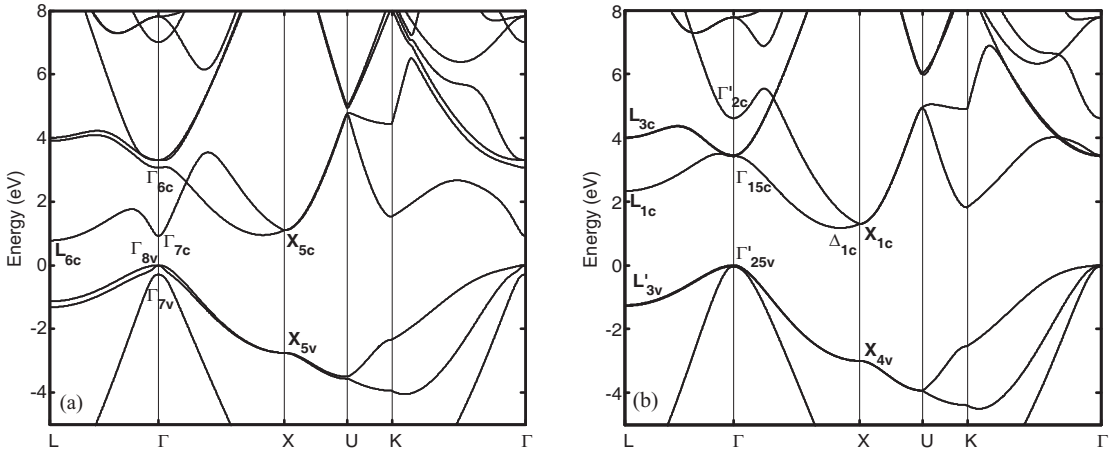


Figure 3·4: Calculated electronic structures of relaxed (a): Ge and (b): Si.

of the electronic structure under strain. When calculating the band structures, the Hamiltonian in Eq. (2.1) is diagonalized using a closed set of 226 spinors (113 plane waves) for a suitable set of $\boldsymbol{k}$-vectors. As it will be explained later, this set of $\boldsymbol{k}$-vectors is optimized for each band in order to capture the details of the band edges while maintaining the datasets of wavefunction expansion coefficients in a manageable size.

The calculated band structures of relaxed Ge and Si using the EPM approach outlined above, are presented in Fig. 3·4. For further comparison, the fundamental gaps and effective masses in different valleys are listed in Table 3.4 and Table 3.3 for Ge and Si respectively from GW calculations, experimental data and this work. The change of the fundamental energy gaps as a function of the applied strain is also calculated and plotted in Fig. 3·5. A strain applied along the [100] direction is considered as the case of the growth of Ge on [100] Si. Fig. 3·5(a) presents the calculated valence and conduction band energies of Ge at the Γ, Δ_{100}, Δ_{001}, and $L_{6,c}$ points in the first Brillouin zone. We notice that for relaxed Ge, the calculated energy separation between $\Gamma_{7,c}$ and $L_{6,c}$ is approximately 0.143 meV, while this number between the minima at Δ_{100} and $L_{6,c}$ is about 0.174 meV. Consequently, the intervalley

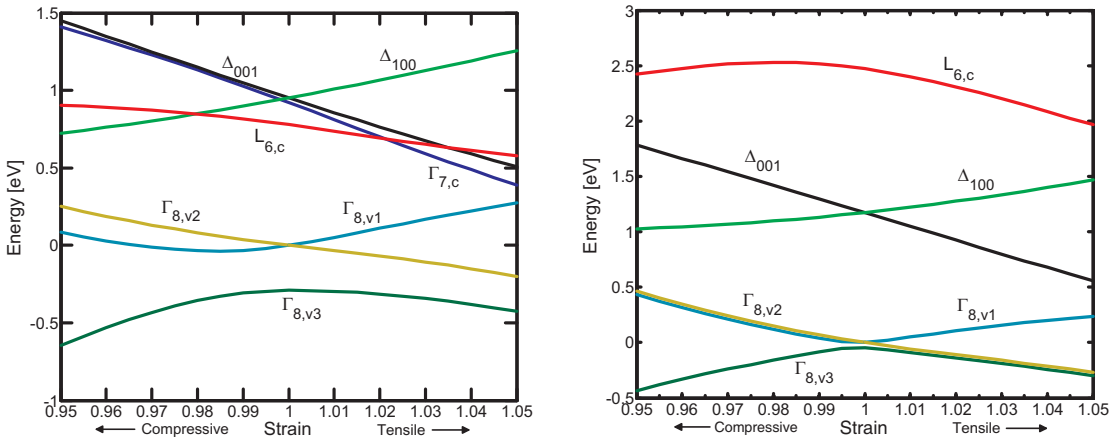
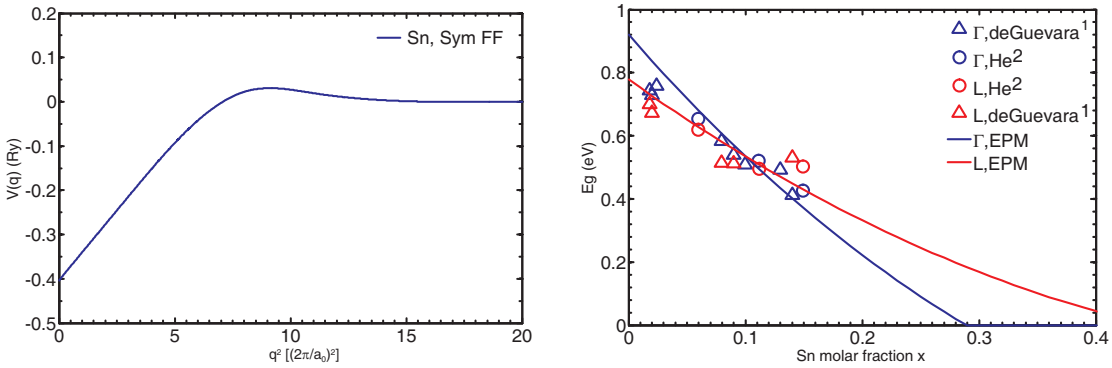


Figure 3·5: Calculated fundamental energy gaps of (a): Ge and (b): Si under compressive and biaxial tensile strain.

energy difference between Δ_{100} and $\Gamma_{7,c}$ is only around 30 meV, which gives rise to the possibility of changing Ge into a direct gap material. Indeed, as is shown in Fig. 3·5(a) and also in the previous example of Fig. 3·2, when applying a tensile strain along [100] direction, the energy of the $\Gamma_{7,c}$ valley drops more rapidly than that of the $L_{6,c}$ valley, and at the tensile strain of 2% the two valleys have the same energy, effectively transiting the material into a direct band-gap material. Fig. 3·5(b) gives the results of the same calculation for Si and one can notice that the minimum energy of the four Δ valleys lying in plane (100) is higher than the two Δ valleys along the [100] direction where a tensile strain is applied. It is also obvious that regardless of the strain magnitude, the $L_{6,c}$ point is always above the Δ valleys and the energy of the Γ point (not shown) never falls below 3 eV.

Besides the band parameters, values of these constants used in calculating electron-phonon interaction using Eq. (2.35), (2.36) to (2.39) can be found in Table 3.5. As was stated in Section 2.2, Thomas-Fermi screening has been adopted in modeling the electron-electron interaction for its simplicity. Therefore, no parameters for dielectric fitting model are listed.



(a) (b)

Figure 3·6: (a): Local pseudo-potentials for α-Sn fitted at 300 K. Cubic spline is used to interpolate the potentials. (b): Energy of L (red line) and Γ (blue line) point computed as functions of α-Sn molar fraction. All the energy references to the top of valence band of the alloy, which is set to zero. Symbols are measured band gap from 1: de Guevara *et al.* (de Guevara et al., 2004) and 2: He *et al.* (He and Atwater, 1997).

α-Sn and GeSn

Pseudo-potentials for Tin at 300 K is fitted in Fig. 3·6(a). Same as Si and Ge, only symmetric atomic potentials exist in this case since the two atoms at the two sites in zincblend structure are exactly the same for elemental materials. We note that at 300 K, α-Sn is metal with its conduction and valance band at Γ point flipped (Madelung, 2004) which makes GeSn alloy also have zero band gap energy if the ratio of α-Sn is large enough. In fact, from Fig 3·6(b) in which the band gap energy for GeSn is computed using VCA, it can be seen that when the ratio of Sn is larger than 29% the alloy becomes a metal. We also noticed that the measured GeSn band gap energy as a function of Sn molar fraction only exists in a limited number of literature and the value recorded varies a lot (Moontragoon et al., 2012; de Guevara et al., 2004; He and Atwater, 1997).

Furthermore, due to the flip of CB and VB at Γ point in α-Sn, the spin-orbit coupling in Sn, which strongly affect the electron energy level in valence band, biases

the band gap energy of the alloys a lot. This behavior is very different from Si and Ge which do not have such band flipping. In order to properly model the spin-orbit coupling in GeSn so that the band obtained from VCA can reproduce the experimental data, we further introduced spin-orbit disorder into the band fitting procedure to accommodate the possible disorder effect during the atom mixing (Kim and Fischetti, 2010). With spin disorder effect included, one could easily fit the band gap energy of GeSn to a satisfactory degree. As is shown in Fig. 3·6(b), the crossover composition ratio of Sn, where GeSn changes from indirect gap material to direct gap material, can be found to be 11%. We noted that in the literature the measured crossover composition varies a lot, from 7% (He and Atwater, 1997) to 17% (Moontragoon et al., 2012). Compared to the result in this work, the obtained crossover point lies in the middle of this range, which is acceptable.

3.2.2 InGaAs, InAsSb and HgCdTe Alloys

Unlike the elemental materials, the effective atomic potentials for compound semiconductors especially for the ternary or quaternary alloys, are not easy to fit from the experimental or first-principle results due to the disorder effects. The disorder effects happens because the crystal may contain fewer atomic species than the fully crystallized structure, or having some crystallographic site occupied by more than one type of atom in different unit cells.

InGaAs

The fitted screened atomic potentials for InAs and GaAs at 300 K as functions of q are shown in Fig. 3·7, where the symmetric and antisymmetric form factors of the binary material are defined as: $V^{\rm sym}_{\rm InAs} = (V_{\rm In}+V_{\rm As})/2$, $V^{\rm asym}_{\rm InAs} = (V_{\rm In}\text{-}V_{\rm As})/2$, in the case of InAs. It is worthy of note that usually, the local potentials for each atomic species at $q = 0$ is set to $-2E_F/3$ with E_F being the Fermi energy at 0 K (Allen and Cardona,

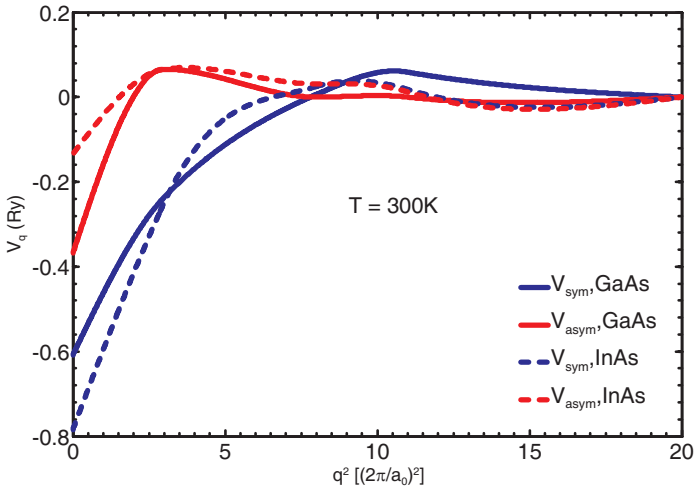


Figure 3·7: Symmetric (blue lines) and antisymmetric (red lines) local pseudo-potentials for GaAs (solid lines) and InAs (dashed lines) fitted at 300 K. The potential at $q = 0$ is also set to be a fitting parameter to get a proper band offsets and workfunction. The potential is cutoff at $q^2 = 20 \times (2\pi/a_0)^2$ by using the natural boundary condition at this point. Here a_0 is the lattice constant of the material.

1983). However, at room temperature since the behavior of local potentials at long wavelength ($q = 0$) is not well defined, and in order to get the correct workfunction and band alignment of the material, we treated $V_q(0)$ also as a fitting parameter. Natural boundary conditions are then applied to $q = 0$ and $q = \sqrt{20}$ to truncate the local potentials to avoid excessive large number of expansion basis plane waves.

Using the screened local potentials for the binary material, we compute the eigenvalues of conduction and valence bands as functions of in-plane lattice constant for strained InAs and GaAs. Figures 3·8(a) and (b) presents the results obtained in this work (solid lines) along with the corresponding energy values (open symbols) calculated by using DFT (Kadantsev and Hawrylak, 2011). It can be appreciated that by properly adjusting the local potentials and their first derivatives, it is possible to obtain energy bands that closely match the first-principle calculations. We can immediately notice that the energy gap decreases under tensile strain for both InAs and GaAs. Furthermore, the crossover of conduction and valence band in InAs is

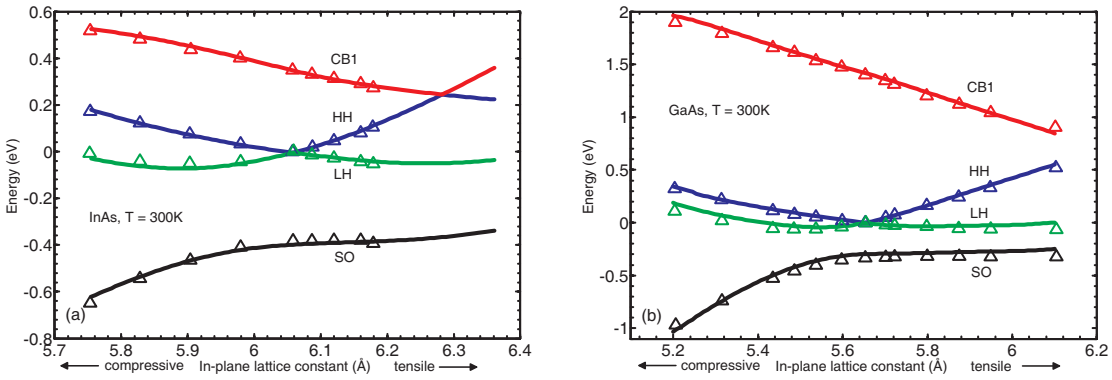


Figure 3·8: Fitted band energy at Γ point for (a) InAs and (b) GaAs as functions of material in-plane lattice constant at 300 K (solid lines). Four bands: first conduction band (CB1), heavy hole band (HH), light hole band (LH) and spin-orbit band (SO) are plotted and compared with the band energy computed from the first-principle calculation (upper triangular symbols) (Kadantsev and Hawrylak, 2011). The strain applied is ±5 % for InAs and ±8 % for GaAs.

predicted to happen under 3.7% tensile strain. A detailed comparison of the band energy and effective masses at different valleys among published results and results from this work is shown in Table 3.2 for the case of relaxed materials.

By utilizing the construction scheme for local potential and lattice constant of a ternary alloy in Section 2.1, and solving the single-electron Schrodinger equation, we fitted the full band structure of $In_xGa_{1-x}As$ by adjusting the disorder parameter P_{dis} according to the measured band gap energy. The results are shown in Fig. 3·9.

In Fig. 3·9, the black solid line and blue dashed line are the band gap energy for relaxed $In_xGa_{1-x}As$ as a function of the Indium molar fraction x from this work and Kim's work (Kim and Fischetti, 2010) respectively. The open circles are data collected from Ref. (Berolo et al., 1973). The band gap energies obtained in this work agree very well with both the experimental measurements and other theoretical calculation. When growing $In_xGa_{1-x}As$ on InP substrate, as a result of the variation of the lattice constant due to different Indium molar fraction values, an in-plane strain is induced

Table 3.2: Comparison of band energy and effective masses between fitted EPM band structure from this work and values from literature at 300 K. Unit for energy and effective mass are eV and rest electron mass m_0 respectively. References are a: Ref. (Madelung, 2004), b: Ref. (Adachi, 2005).

Material	Type	$E^{\Gamma_c-\Gamma_v}$	$E^{L_c-\Gamma_v}$	$E^{X_c-\Gamma_v}$	Δ_{SO}	m_e^{Γ}	m_{HH}^{Γ}	m_{LH}^{Γ}
InAs	This work	0.352	1.567	1.933	0.394	0.0295	0.53	0.036
	Literature	0.354^a	1.55^a	1.90^a	0.41^a	0.023^a	0.35^a	0.028^a
GaAs	This work	1.418	1.79	1.856	0.296	0.070	0.44	0.104
	Literature	1.43^b	1.72^b	1.894^b	0.33^b	0.067^b	0.33^b	0.09^b

in the material. The resulting band gap energy as a function of x is presented by red solid line in Fig. 3·9. As it is can be seen, at $x = 0.53$ which indicates the InGaAs is lattice-matched to the InP substrate, the obtained band energy equals to the energy in relaxed material.

When the Indium molar fraction increases, the lattice constant for InGaAs becomes larger and compressive strain appears which leads to an increase of band gap compared to the relaxed situation. In the opposite case, where the Indium molar fraction is smaller than 53%, an in-plane tensile strain presents in the material and the resulting band gap energy decreases compared to the value for relaxed InGaAs. The detailed full band structure of compressively strained $In_{0.89}Ga_{0.11}As$ is shown in the inset of the Fig. 3·9 (solid lines). The calculated electronic structure of relaxed HgCdTe (Bertazzi et al., 2011; Wen and Bellotti, 2014), with the same cutoff wavelength as that of the strained $In_{0.89}Ga_{0.11}As$, is plotted using dashed lines. It can be computed from the band curvature that the effective mass for electron in HgCdTe ($0.052\, m_0$) is larger than that in InGaAs ($0.043\, m_0$), which usually results in a smaller radiative recombination rate in the material. Since in the SWIR spectral range the dominating intrinsic recombination mechanism is radiative recombination (Wen et al., 2015) at room temperature, it can be expected that HgCdTe has longer intrinsic minority carrier

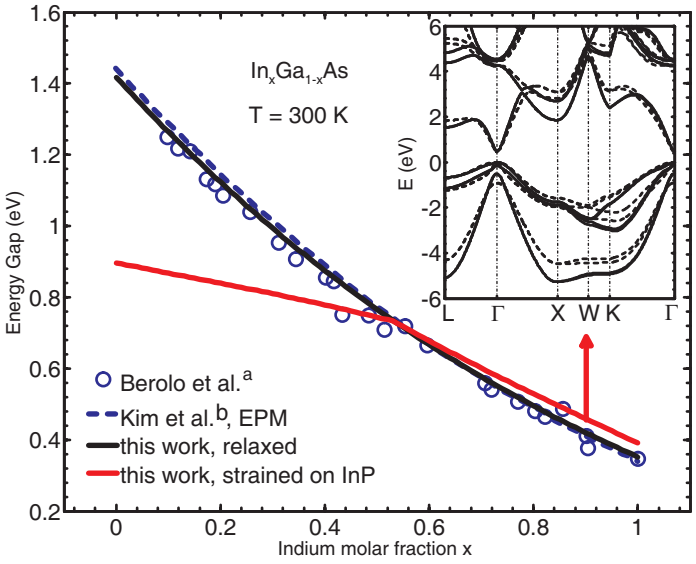


Figure 3·9: Band gap energy for relaxed (black solid line) and strained (red solid line) InGaAs in terms of Indium molar fraction x at 300 K. When calculating the band gap energy for strained material, an InP substrate is assumed which is lattice-matched to $In_{0.53}Ga_{0.47}As$. The blue dashed line is the result from b: Kim's work (Kim and Fischetti, 2010) and the circle symbols are the measured band gap energy data excerpted from a: Ref. (Berolo et al., 1973). The inset of the figure compares the band structure of strained $In_{0.89}Ga_{0.11}As$ (solid lines) and relaxed HgCdTe (dashed lines) which have cutoff wavelength of 2.7 μm.

lifetime compared to the strained InGaAs as will be shown in Chapter 5, Section 5.3.

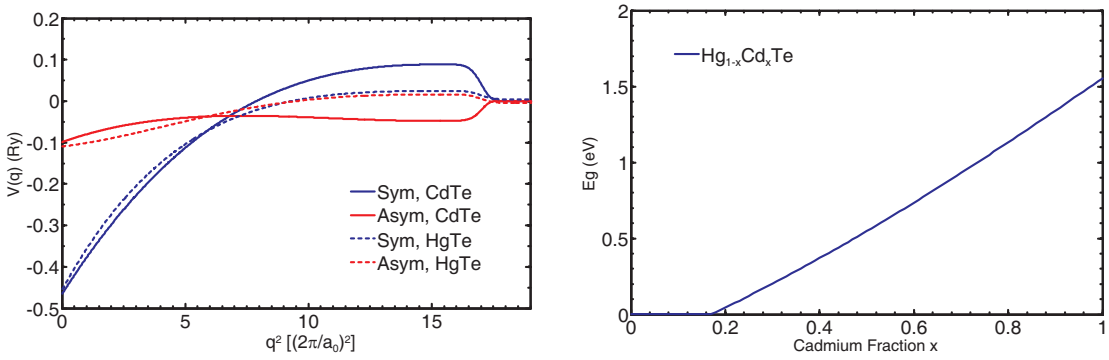
InAsSb

Using the same fitting procedure in Section 3.2.2, results for InSb is shown in Fig. 3·10, which include both (a): form factors and (b): obtained electron energy at Γ point for different bands in terms of in-plane lattice constant. We notice that the fitted form factors of InSb, especially the antisymmetric part, oscillate when the q gets larger, which is in contrast to the fitting of InAs and GaAs. Usually one should expect a smooth variation of the atomic potentials in EPM as is stated in Section 2.1. We speculate that this phenomenon is due to the comparable size of Indium and Antimony atoms, which can be eliminated by further employing non-local pseudo-potential

correction. Nevertheless, the fitted band energy for InSb matches to the experimental data very well as is shown in Fig. 3·10(b).

As to the band structure of $InAs_{1-x}Sb_x$ alloy, Fig. 3·11 gives the fitting results obtained from VCA. Similar to InGaAs, the lattice constant of the alloy is linearly interpolated between values for InAs and InSb, and only potential disorder effect is considered. It is also worthy of note that the when the material is subject to compressive strain (black dashed line), the band gap energy increases compared to the relaxed case which agrees with the theoretical prediction in Chapter 2.

HgCdTe



(a) (b)

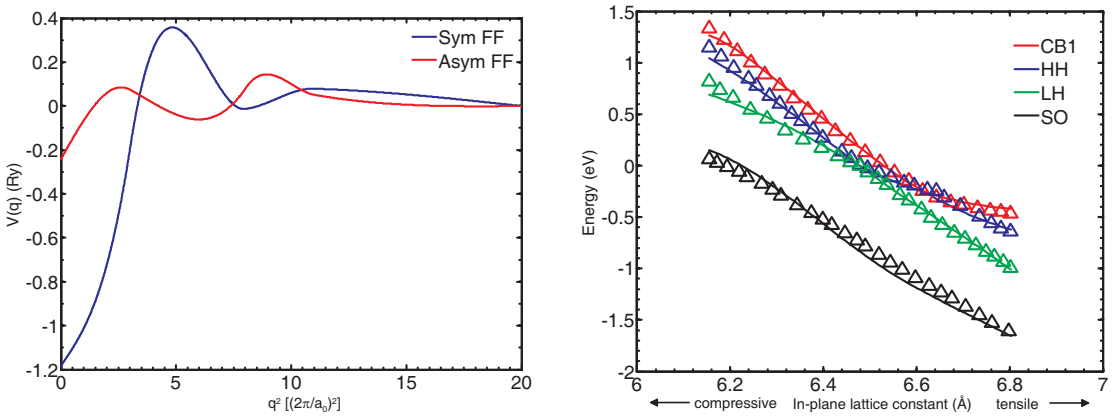
Figure 3·12: (a): Symmetric (blue lines) and antisymmetric (red lines) pseudo-potentials for CdTe (solid lines) and HgTe (dashed lines) fitted at 0 K. Nonlocal corrections to the pseudo-potential is used in CdTe whereas only local pseudo-potential is employed when fitting HgTe. The pseudo-potentials have been cut of at $q = \sqrt{18}\, 2\pi/a_0$. (b): Fitted band gap energy for $Hg_{1-x}Cd_xTe$ as functions of Cd molar fraction x.

Fitting results for HgCdTe is shown in Fig. 3·12. Different from the fitting for InGaAs and InAsSb, here nonlocal correction is applied to the fitting of CdTe and a 7-parameter asymmetric exponential fit is adopted to interpolate the atomic potentials for both HgTe and CdTe. The potentials are forced to cutoff at $q = \sqrt{18}\, 2\pi/a_0$. It can be seen that when nonlocal correction is employed and oscillation of potential disappeared,

and the resulting potentials (form factors) are very smooth. Band gap energy of HgCdTe as a function of Cd fraction is demonstrated in Fig. 3·12(b). Since the HgTe is metal that has negative band gap energy, when the ratio of CdTe is lower in the HgCdTe alloy, the material is effectively a metal that has zero band gap.

3.3 Conclusions of the Chapter

In this chapter, strain effect and detailed electronic structures for elemental and compound materials have been investigated. First the effects of strain in the material on both the shape of Brillouin zone and electronic structures are demonstrated. Then the constructed full band structures of silicon, germanium, α-tin, the GeSn, InGaAs, InAsSb and HgCdTe alloys from the experimental data or results from the first principle calculations are presented. These electronic structures will be further used in the material and device simulations which provides the abstraction of the interaction between electrons and the periodic atomic potentials.



(a) (b)

Figure 3·10: (a): Symmetric (blue line) and antisymmetric (red line) local pseudo-potentials for InSb fitted at 300 K. (b): Fitted band energy at Γ point for InSb as functions of material in-plane lattice constant at 300 K. Four bands are plotted and compared with the band energy from the first-principle calculation (triangle symbols) in Ref. (Kadantsev and Hawrylak, 2011). The strain applied in InSb is ± 5%.

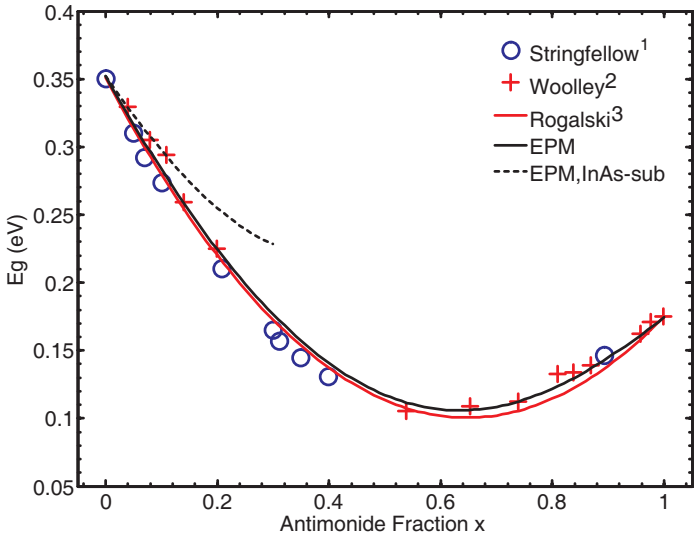


Figure 3·11: Fitted band gap energy of $InAs_{1-x}Sb_x$ alloy in terms of x using VCA. Pseudo-potentials for InAs used in the fitting come from the results presented in Fig. 3·7. Band structures from this work (black solid line) are compared to the measured data from Stringfellow *et al.* circle symbols (Stringfellow and Greene, 1971), Woolley *et al.* cross symbols (Woolley and Warner, 1964) and fitted formula from Rogalski (Rogalski, 1994). The dashed black line shows the change of band gap energy of $InAs_{1-x}Sb_x$ assuming an InAs substrate.

Table 3.3: Energy gaps and effective masses for Si. (a) Values calculated using GW+SO from Ref. (Malone and Cohen, 2013). (b) Experimental values from Ref (Madelung, 2004). m_0 is the electron mass.

Parameter	Unit	This Work	GW+SO[a]	Exp.[b]
Γ_{1v}	eV	-12.41	-11.63	-12.5
Γ'_{25v}	eV	-0.042	-0.05	-0.044
Γ_{15c}	eV	3.36	3.28	3.34
Δ_{1c}	eV	1.17	-	1.17
X_{1c}	eV	1.29	1.43	1.3
L_{1c}	eV	2.23	2.13	2.04
L_{3c}	eV	4.0	4.16	3.9
$m_e^{\Delta\perp}$	m_0	0.2	-	0.19
$m_e^{\Delta\parallel}$	m_0	0.928	-	0.92
m_e^{Γ}	m_0	0.4	-	-
m_{hh}^{Γ} [100]	m_0	0.27	-	0.537
m_{lh}^{Γ} [100]	m_0	0.22	-	0.153
m_{so}^{Γ} [100]	m_0	0.24	-	0.234
m_{hh}^{Γ} [110]	m_0	0.56	-	0.537
m_{lh}^{Γ} [110]	m_0	0.15	-	0.153
m_{so}^{Γ} [110]	m_0	0.24	-	0.234
m_{hh}^{Γ} [111]	m_0	0.69	-	0.537
m_{lh}^{Γ} [111]	m_0	0.14	-	0.153
m_{so}^{Γ} [111]	m_0	0.24	-	0.234

Table 3.4: Energy gaps and effective masses for Ge. (a) Values calculated using GW+SO from Ref. (Malone and Cohen, 2013). (b) Experimental values from Ref (Madelung, 2004). m_0 is the electron mass.

Parameter	Unit	This Work	GW+SO[a]	Exp.[b]
Γ_{6v}	eV	-12.2	-12.53	-12.66
Γ_{7v}	eV	-0.291	-0.32	-0.3
Γ_{7c}	eV	0.92	0.38	0.9
Γ_{6c}	eV	3.3	2.89	3.25
X_{5c}	eV	1.1	1.16	1.16
L_{6c}	eV	0.78	0.54	0.76
L_{3c}	eV	3.9	4.18	4.2
$m_e^{L\perp}$	m_0	0.09	-	0.082
$m_e^{L\parallel}$	m_0	1.45	-	1.57
m_e^{Γ}	m_0	0.047	-	0.038
m_{hh}^{Γ} [100]	m_0	0.286	-	0.284
m_{lh}^{Γ} [100]	m_0	0.061	-	0.0438
m_{so}^{Γ} [100]	m_0	0.132	-	0.095
m_{hh}^{Γ} [110]	m_0	0.544	-	0.376
m_{lh}^{Γ} [110]	m_0	0.054	-	0.0426
m_{so}^{Γ} [110]	m_0	0.132	-	0.095
m_{hh}^{Γ} [111]	m_0	0.736	-	0.352
m_{lh}^{Γ} [111]	m_0	0.054	-	0.043
m_{so}^{Γ} [111]	m_0	0.132	-	0.095

Table 3.5: Material parameters used in the calculation of electron self-energy for electron-phonon interaction. Values of the parameters are obtained or derived from Refs. (Fischetti and Laux, 1996; Madelung, 2004).

Parameter	Unit	Ge	Si
v_l	10^5 cm/s	5.31	9.0
v_t	10^5 cm/s	3.61	5.41
c_l	10^{11} dyn/cm^2	12.85	16.58
c_t	10^{11} dyn/cm^2	6.68	7.96
ϵ_0	-	16.2	11.9
ϵ_∞	-	16.2	11.9
$\hbar\omega_{\rm op}$	meV	37.04	61.2
Ξ_d	eV	8.1	7.1
K_{AV}	-	0	0
$D(L_{6,c})$	10^8 eV/cm	9.05	9.35
$D(\Gamma_v)$	10^8 eV/cm	12.17	13.24

Chapter 4

Optical Absorption and Radiative Recombination Rate in Si and Ge

4.1 Optical Absorption and Gain of Ge Lasing Medium

As a result of the ongoing activities to develop light emitter and laser structures compatible with standard silicon processes, a renewed effort has been devoted to study germanium (Ge) and silicon (Si) optical properties. Particular attention has been paid to understand how the optical properties of these materials change under the effect of strain and doping (Boucaud et al., 2013; Virgilio et al., 2013; Soref et al., 2012). On the experimental side, attempts to develop Ge-based lasers have leveraged the fact that using the combination of tensile strain and high doping a significant enhancement of the direct radiative recombination rate could be obtained (S*ü*ess et al., 2013; Liu et al., 2010; Capellini et al., 2013), and the optimum strain to fulfill the lasing conditions in Ge has also been studied theoretically (Aldaghri et al., 2012). More exotic approaches based on membranes are also being explored (Kurdi et al., 2010; Boztug et al., 2013) which have demonstrated possible ways of achieving direct band-gap in Ge. At the same time Si light emitters (du Plessis et al., 2013; Venter et al., 2013b) based on avalanche processes have made the development of micro-displays possible (Venter et al., 2013a). The present work intends to contribute to this effort by achieving three main objectives. First, we intend to develop an appropriate formalism to study the optical absorption and radiative properties of strained Ge and Si that goes beyond

the conventional second order perturbation theory (SOPT). Although, SOPT has been employed to analyze the indirect processes for many years (Bude et al., 1992), it is difficult to properly handle the divergences when the energy of transitions is not known a priori. As a result, it may not be possible to apply this approach to all the possible cases consistently. The proposed model, however, is based on the Green's function formalism (Mahan, 1993), in which we employ the spectral density functions to describe the broadening of the states involved in the transitions due to different physical processes, therefore avoiding the divergence difficulties in SOPT naturally (Takeshima, 1982a). Second, we intend to develop a numerical approach suitable to compute the aforementioned properties using the full band structure of the material. This approach eliminates the approximations introduced by the analytical models based on parabolic or non-parabolic bands, makes it possible to analyze the behavior of the material when a significant portion of carriers populate the states far above the band edge (Bude et al., 1992; du Plessis et al., 2013), and naturally includes the effects of valence bands warping and anisotropy, which is particularly important when evaluating the matrix elements of the radiative process. Finally, we apply the developed model to study strained Ge and Si, where it will be shown that the model can reproduce the experimental data for the temperature-depended absorption coefficient of relaxed material. Subsequently, strained materials will be further investigated in terms of energy- and temperature-dependent absorption coefficients and radiative recombination rates under different doping and injection conditions.

4.1.1 Absorption Coefficients of Ge

In this section we will discuss the results obtained by applying the method outlined in Chapter 2, Section 2.2 to the calculation of the absorption coefficients and radiative recombination rates in relaxed and strained Ge and Si. A schematic of the important absorption paths in relaxed Ge is illustrated in Fig. 4·1.

We first consider the case of relaxed and strained bulk Ge. Fig. 4·2(a) presents the calculated photon-energy-resolved absorption coefficient in two sets of lines. The first set, on the right side of the plot, is for relaxed bulk Ge, and the second set, on the left side, is for 1.4% tensile strained bulk Ge. The solid lines (red color) represent the calculated absorption coefficient at the temperature of 77 K, while the dashed lines (green color), dash-dot lines (black color) and dash-two-dot lines (purple color) are for the temperatures of 150 K, 225 K, and 300 K respectively. To validate the accuracy of our numerical model, the experimental data of the absorption coefficient measured at 77 K and 300 K (Dash and Newman, 1955) are included as well, which are represented by the open symbols (squares for 77 K and circles for 300 K). As one can see in Fig. 4·2(a), in the case of unstrained intrinsic Ge, the numerical results match very well with the experimental data for both 77 K and 300 K in the whole range of the measured photon spectrum. These comparisons provide a high degree of confidence in the predictive ability of the numerical method and the reliability of our ensuing analysis of the strain effect on the absorption coefficients and radiative recombination rates. As an additional note, we would like to point out that for the results presented in this section, contributions of the transitions from LH and HH to the first conduction band are summed together because of the difficulties of separating one transition from the other in the presence of the band mixing between LH and HH under strain. Furthermore, since the higher conduction bands and lower valence bands are generally less populated, they provide a negligible correction to our results in the range of photon energy considered here. Consequently, they are selectively included in the calculation of optical transitions based on their relative carrier populations.

Comparing the absorption spectra for different lattice temperatures T in Fig. 4·2(a), it is apparent that when T decreases, at the same amount of tensile strain, the shape of the curves barely changes despite the fact that the curves tend to shift toward

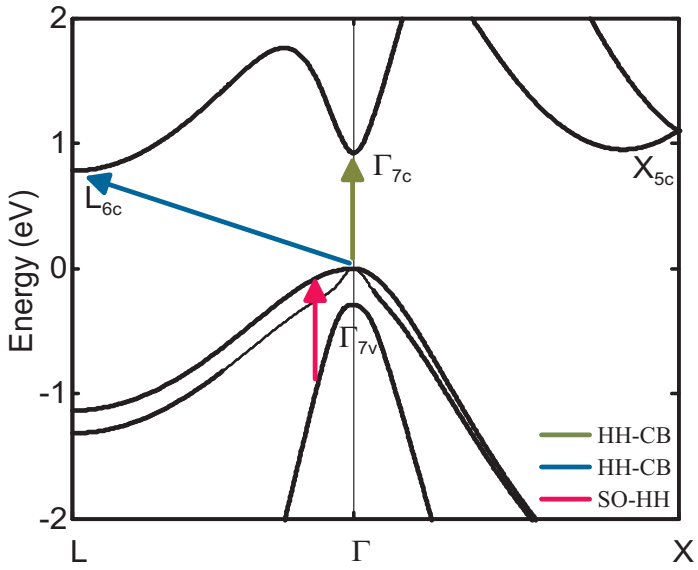
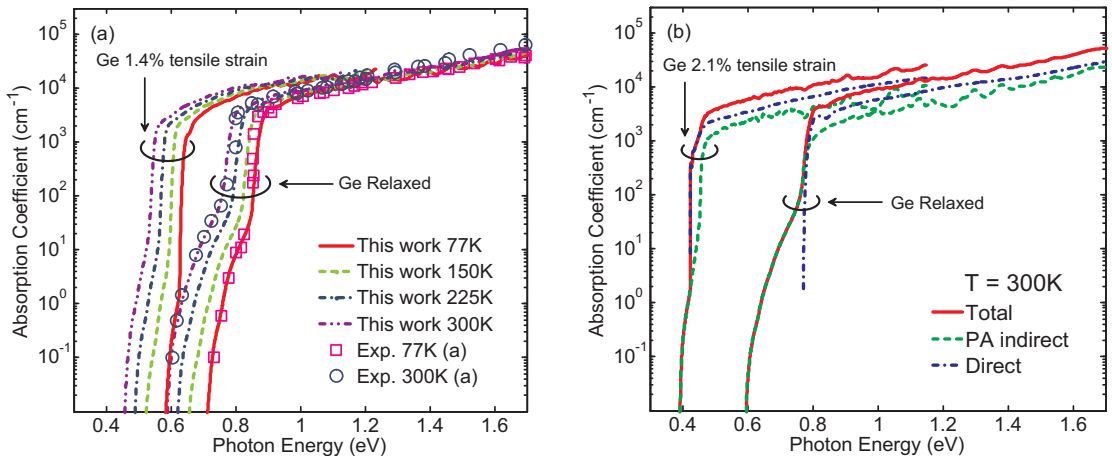


Figure 4·1: Detail of the calculated Ge electronic structure near the band edges including the direct, indirect HH-CB and SO-HH transition paths for electrons. The absorption transition paths are represented by the arrows in the figure while the gain transitions are the reverse processes.

higher photon energy and the tail near the absorption edge shrinks. The shift of curves with temperature is consistent with the experimental data (Madelung, 1996), where the increase of lattice temperature makes both direct and indirect band gaps shrink nearly uniformly. Additional information about the tails of the absorption curves can be obtained from Fig. 4·2(b), in which the contributions to the total absorption coefficient (solid lines) from the PA indirect (dashed lines) and direct (dash-dot lines) processes are plotted in two tensile strain configurations: $\epsilon = 0$ and $\epsilon = 2.1\%$ at 300 K for n-type Ge ($N_d = 10^{19}\text{cm}^{-3}$). As is shown, the PA indirect transition from Γ_v to L_c states predominantly give rise to the tails of absorption curves since the corresponding photon energy is just slightly above the indirect bang gap. When the absorbed photon energy gets larger, the direct transition process starts to be dominant, resulting in a sudden increase in the absorption curves. Back to Fig. 4·2(a), due to the fact that at low temperature, the average number of phonons in the crystal decreases, the



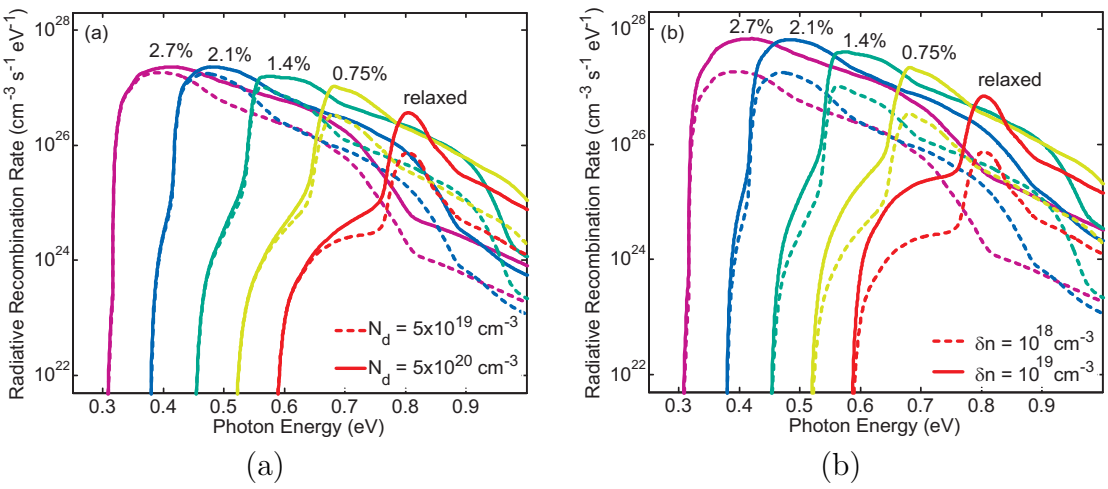
(a) (b)

Figure 4·2: Absorption coefficient spectra of Ge calculated with Green's function theory under different strains, doping and lattice temperatures at thermal equilibrium. (a): Intrinsic Ge with relaxed lattice structure (right group) and $\epsilon = 1.4\%$ biaxial tensile strained structure (left group) are compared with the measured data for intrinsic relaxed Ge at 77 K (open circles) and 300 K (open squares) (Dash and Newman, 1955). Style and color of lines represent the corresponding lattice temperatures. (b): Results of 10^{19}cm^{-3} n-doped Ge at 300 K with the tensile strain $\epsilon = 0$ (right group) and $\epsilon = 2.1\%$ (left group). Contributions from direct process (dash-dot lines) and PA indirect process (dashed lines) are compared with the total absorption coefficient (solid lines).

strength of indirect transitions is reduced as well. In both cases of relaxed ($\epsilon = 0$) and tensile strained material ($\epsilon = 1.4\%$), the indirect part in the total absorption gradually diminishes when T decreases from 300 K to 77 K. In fact, for a tensile strain of $\epsilon = 1.4\%$ and a photon energy equal to the direct band gap, the indirect absorption process only accounts for about 2% of the total absorption at 77 K, whereas the number for the same process is 48% at 300 K. For the unstrained material ($\epsilon = 0$), although the indirect transition is weaker at lower temperature, the difference is not significant due to the relatively large energy separation between L_c and Γ_c valleys making the indirect transition consistently dominant.

From the analysis of the relaxed Ge band structure in Fig. 3·4(a), the phonon-assisted indirect transition generally can be divided into three contributions. The first one is from the Γ_v to L_c which is largely responsible for the absorption at lower photon energy. A second contribution, coming from the Γ_v to X_c transition, which is only relevant when the photon energy is larger than the direct band gap. Finally, the last one, from the region around Γ_v to Γ_c, which is usually negligible due to the small matrix element for phonon scattering (see Eqs. (2.36)-(2.39)). Although the PA transitions from Γ_v to L_c in relaxed Ge is the dominant absorption mechanism for the photons with energy between the fundamental indirect gap and the direct gap, it should be noted, however, that all the possible transitions need to be considered when the absorbed photon energy gets larger than the direct band gap. In fact, as demonstrated by the right group of lines in Fig. 4·2(b), for photon energies higher than the direct gap, even though the direct process is dominant, the indirect process still accounts for $\approx 41\%$ of the total absorption making the PA process never negligible. This is in stark contrast to direct-gap materials, such as HgCdTe or InAsSb (Wen and Bellotti, 2014).

When a tensile strain is added to the system, as in the case of $\epsilon = 2.1\%$ in Fig. 4·2(b), the situation changes. It can be noticed that the tail of absorption spectrum, which is obvious in the curve obtained for relaxed Ge, fades as the strain increases. Indeed, as is shown in Fig. 3·5, when the tensile strain increases, the direct band-gap shrinks faster than the indirect band-gap, making the direct process gradually more important than its indirect counterpart and eventually becoming the dominant one. Specifically, at $\epsilon = 2.1\%$, the energy difference between the L_c and Γ_c valleys almost vanishes and the indirect tail of the absorption spectra completely merges into the direct part. This band-gap reduction also explains the shift of absorption edge to lower photon energy under increasing tensile strain. The same trend can also be observed in Fig. 4·2(a) for



(a) (b)

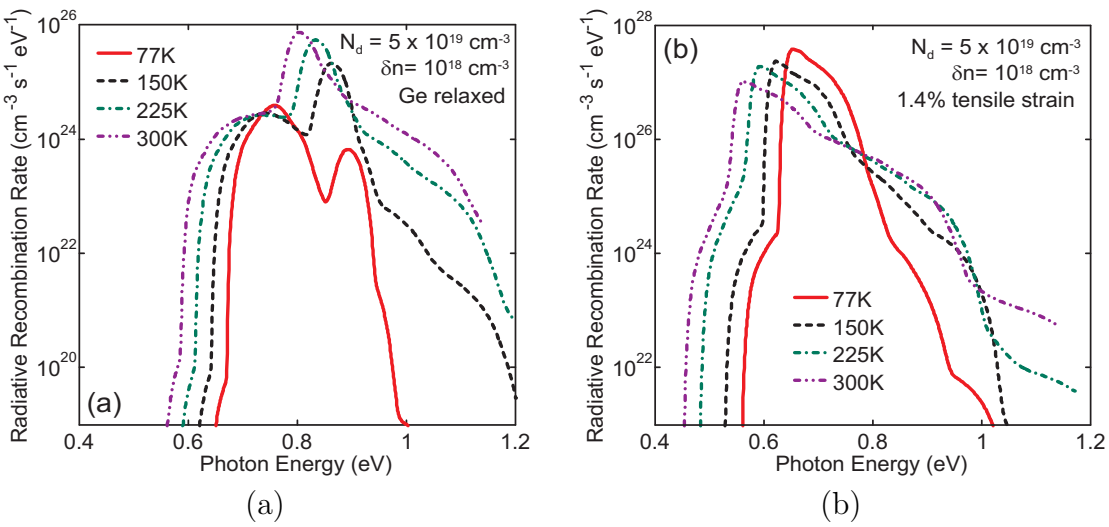
Figure 4·3: Radiative recombination rates of Ge at 300 K are computed under (a): different doping concentrations of $5 \times 10^{19}\text{cm}^{-3}$ (dashed lines) and 10^{20}cm^{-3} (solid lines) with the same injection of 10^{18}cm^{-3}; (b): different injection of 10^{18}cm^{-3} (dashed lines) and 10^{19}cm^{-3} (solid lines) with the same n-doping of $5 \times 10^{19}\text{cm}^{-3}$. Numbers adjacent to each curve represent values of the corresponding biaxial tensile strain.

the strain of $\epsilon = 0$ and $\epsilon = 1.4\%$, where the tail significantly diminished in the latter case at all the lattice temperatures considered.

4.1.2 Radiative Recombination Rates of Ge

Besides the considerable effects of strain on the absorption behavior of Ge, it is also important to investigate its effect on the spontaneous radiative recombination rate spectra $R(\hbar\omega)$, which could provide additional insight in understanding the electroluminescence properties in both relaxed and strained materials. Fig. 4·3(a) presents the computed spontaneous recombination rate $R(\hbar\omega)$ for bulk Ge under different strain conditions. The calculation is carried out under two kinds of n-type doping, $N_d = 10^{19}\text{cm}^{-3}$ (solid lines) and $N_d = 10^{20}\text{cm}^{-3}$ (dashed lines) at a constant injection level of $\delta n = 10^{18}\text{cm}^{-3}$. In Fig. 4·3(b), the same quantity is calculated under a constant n-type doping $N_d = 5 \times 10^{19}\text{cm}^{-3}$ with two different injection levels $\delta n = 10^{18}\text{cm}^{-3}$ (dashed lines) and $\delta n = 10^{19}\text{cm}^{-3}$ (solid lines).

We first consider the case of constant injection presented in Fig. 4·3(a), where, because of the faster reduction of direct band-gap than the indirect one under increasing tensile strain, the shoulder in the curve due to PA indirect processes shrinks and eventually merges into the peak corresponding to the direct recombination process. This is clearly consistent with the previous results obtained for the absorption coefficients. Furthermore, identical reasoning also explains the enhancement effect of increasing doping concentration on the indirect radiative rate among different strain conditions. In fact, for relaxed Ge, due to the higher density of states and lower electron energy in L_c valley compared to Γ_c, a large fraction of electrons ionizing from the dopant will reside in L_c, making the enhancement of the PA indirect rate most significant in this case as is shown in Fig. 4·3(a). When the strain increases, more electrons tend to be in Γ_c valley, effectively lowering the enhancement of PA indirect process. Specifically, when increasing the doping concentration at $\epsilon = 2.1\%$, no visible change of the indirect radiative rate can be observed. Different behavior of increasing doping and injection concentrations can be investigated by comparing the results from Fig. 4·3(a) to that from Fig. 4·3(b), in which we can observe that by increasing the injection level, the PA process gets uniformly enhanced, while adding dopants hardly changes the radiative rate at the indirect absorption edge. This can be understood by considering the carrier distribution in the material. Similar to the aforementioned analysis, when extra donors are added to the system, as is the case of Fig. 4·3(a), more electrons will reside in L_c valley with the available hole concentration remaining the same in Γ_v. Consequently, only a small part of the additional electrons can recombine with holes due to the lack of additional holes in Γ_v. As a result, processes occurring at the indirect absorption edge are not significantly enhanced. When the carrier injection is increased instead, both electrons and holes are added to the system ensuring that the holes will not limit the recombination process. Therefore, the PA indirect processes are enhanced in the



(a) (b)

Figure 4·4: Temperature dependence of the radiative recombination rate in Ge obtained with biaxial tensile strain of (a): 0.0% and (b): 1.4% under the same doping/injection condition, i.e. $5 \times 10^{19}\text{cm}^{-3}/10^{18}\text{cm}^{-3}$.

whole spectrum which gives rise to the differences between Fig. 4·3(a) and (b). For this reason, we can conclude that increasing the carrier injection level is a more efficient way to enhance the radiative recombination rate than using high doping. One should be aware that, however, at high injection other recombination mechanisms, such as Auger recombination and free carrier absorption will become significant which will in turn limit the radiative rate that could be achieved otherwise. It is also noteworthy that in both Figs. 4·3(a) and (b), when the tensile strain exceeds 2.1%, for which Ge becomes into direct band-gap material, there is no further increase of $R(\hbar\omega)$ due to strain.

To conclude the discussion about Fig. 4·3, we compare our results with the calculations performed by Virgilio and coworkers (Virgilio et al., 2013). We noticed that our results are generally in agreement with those presented in Fig. 3 and Fig. 4 of Ref. (Virgilio et al., 2013), in which the radiative recombination rate was obtained using the SOPT. While the peak values and the trend of the recombination rate as a function of the photon energy from the two works are similar, a number of differences

need to be mentioned. In particular, one can notice that sharp corners are present in the curves shown in Fig. 3 and Fig. 4 of Ref. (Virgilio et al., 2013). We speculate that this may be due to the difference between numerical approaches used in the calculations, especially the band structures for Ge. In fact, in our model, instead of considering the two principal transitions Γ_v to Γ_c and Γ_v to L_c separately, where a simple parabolic band approximation can be employed, we use the full band structure with all the possible recombination paths beyond the two major transitions covered. This methodological difference distinguishes our results from those in Ref. (Virgilio et al., 2013) by having a smooth interim region between the two peaks, where processes involving higher energy states in Γ_c, L_c and X_c valleys are included.

The temperature dependence of the radiative recombination rate for both relaxed and strained bulk Ge are examined in Fig. 4·4 with a $5 \times 10^{19}\text{cm}^{-3}$ n-type doping and 10^{18}cm^{-3} carrier injection. Fig. 4·4(a) presents the results for relaxed Ge and Fig. 4·4(b) provides the same information for a tensile strain of $\epsilon = 1.4\%$. As is shown in Fig. 4·4(a), two features representing the PA indirect and direct processes are presented for all four lattice temperatures. As the temperature decreases, the peak corresponding to the direct process is significantly reduced, while the shoulder which accounts for the indirect process practically retains its magnitude. Given that all of the four curves are computed using the same carrier concentration and assuming that the dopants are fully activated at all temperatures of interest, the change in the shape of the radiative rate can be understood by considering the relative carrier distributions between the regions around Γ_c and L_c valleys. When the temperature decreases, the probability of occupation for electrons at higher energies in the conduction bands, particularly around the Γ_c valley, will be reduced significantly according to the quasi-Fermi distribution. This leads to a decrease of the direct radiative recombination rate which drops from a peak value of $7.4 \times 10^{25}\text{cm}^{-3}\text{s}^{-1}\text{eV}^{-1}$ at 300 K down to

$6.5 \times 10^{23}\text{cm}^{-3}\text{s}^{-1}\text{eV}^{-1}$ at 77 K. As to the indirect process, due to the large density of states and low electron energy in the L_c valley, the change in the electron occupation probability will not significantly impact the overall electron population in L_c valley, making the indirect recombination rate remain the same.

As it can be expected, the outcome of the same analysis changes if strain is added to the system. Fig. 4·4(b) presents the calculated radiative recombination rate for $\epsilon = 1.4\%$ tensile strained Ge. We can notice that in this situation the peak corresponding to the direct process, in contrast with the case of relaxed Ge, increases in magnitude as the temperature decreases. This behavior can be traced back to the interplay between the L_c - Γ_c inter-valley energy separation and the position of the electron quasi-Fermi energy. For relaxed Ge, due to the large energy separation between L_c and Γ_c, which is 0.143 eV compared to 0.047 eV for the case of $\epsilon = 1.4\%$, and the higher density of states of L_c valley, the electron quasi-Fermi level at 77 K is 0.07 eV below the direct band-edge (at Γ_c), while in strained Ge, the electron quasi-Fermi level is 0.03 eV above the direct band-edge. This leads to a Γ_c valley that is mostly empty at low temperatures in the former case, and an always partially filled Γ_c valley even at 77 K in the latter case. Additionally, for strained Ge when the temperature is reduced, in order to maintain the same carrier injection more electrons will be concentrated around the Γ_c valley, making the radiative recombination happen in a very small range of photon energies. Consequently, the peak value of the recombination spectrum increases while the width of the peak shrinks. As to the relaxed structure, the lack of electrons in the Γ_c valley at low temperature makes the peak of the direct recombination fall dramatically. We noted that this carrier concentration effect also happened in indirect recombination from L_c, as observed in Fig. 4·4(a), but is much weaker.

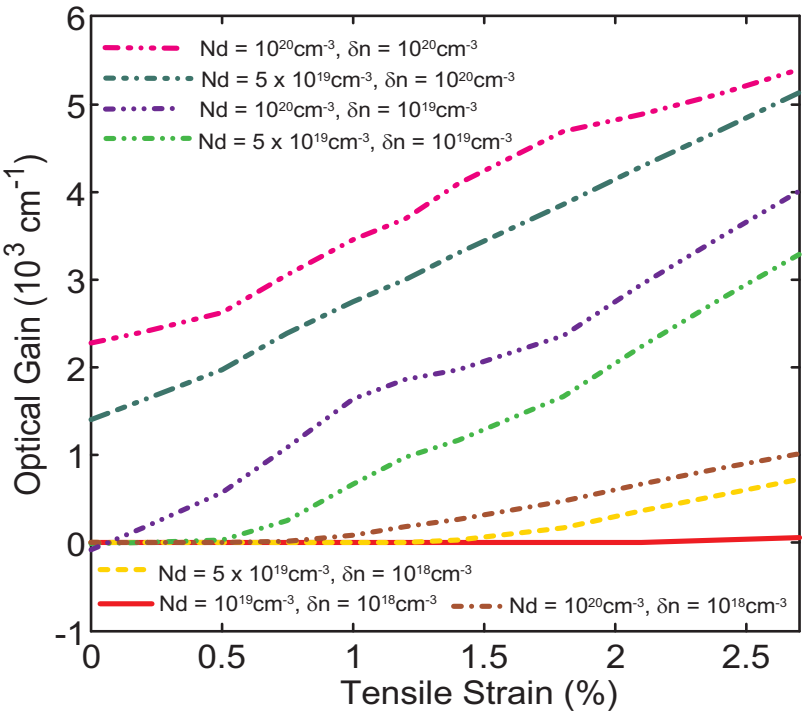


Figure 4·5: Peak values of optical gain coefficient in Ge as a function of strain conditions are calculated under various (doping,injection) levels (see legend). The gain are obtained solely from the transition between CB-HH and CB-LH with no free carrier absorption or spin-orbit to HH absorption included.

4.1.3 Optical Gain of Ge under High Injection Conditions

One of the most compelling reasons to study the tensile strained Ge is the possibility to turn the indirect band-gap material into a direct one, potentially leading to the design of an efficient and CMOS-compatible laser. Therefore, in addition to the absorption coefficient we have also calculated the optical gain of strained Ge at 300 K under various doping/injection conditions. Fig. 4·5 shows the peak gain values as a function of applied tensile strain, where each curve represents a different doping/injection condition. It can be observed that the peak of the optical gain for a given doping/injection concentration increases almost linearly with the tensile strain, and a maximum value of 5400 cm^{-1} is attained at a tensile train of 2.7% for an equal

doping/injection concentration of 10^{20}cm^{-3}. As in the conclusion drawn from Fig. 4·3, here the importance of carrier injection can also be observed. In Fig. 4·5 (solid, dashed and dot-dash), regardless of the strain conditions, a relative minor increase of the gain with doping level is presented when the carrier injection δn is 10^{18}cm^{-3}. On the other hand, for example at $\epsilon = 2\%$, when keeping the doping level at $5 \times 10^{19}\text{cm}^{-3}$ while changing the injection level from $\delta n = 10^{18}\text{cm}^{-3}$ to 10^{20}cm^{-3}, the gain increases by about an order of magnitude. This observation still holds for other strain conditions. It is also interesting to point out that, though minor compared to the gain achieved by increasing the injection, higher doping usually leads to a moderate gain increase at small strain values. This increase is largely due to the reciprocal space filling in the L_c valley which occurs under high doping concentration leading to a marginally higher carrier population in the Γ_c valley even at a small injection (Dutt et al., 2012). Finally, we note that for all of the calculated gain values, the PA process always contributes to 20-30% of the total gain, making it an indispensable part in a credible model of optical processes in Ge.

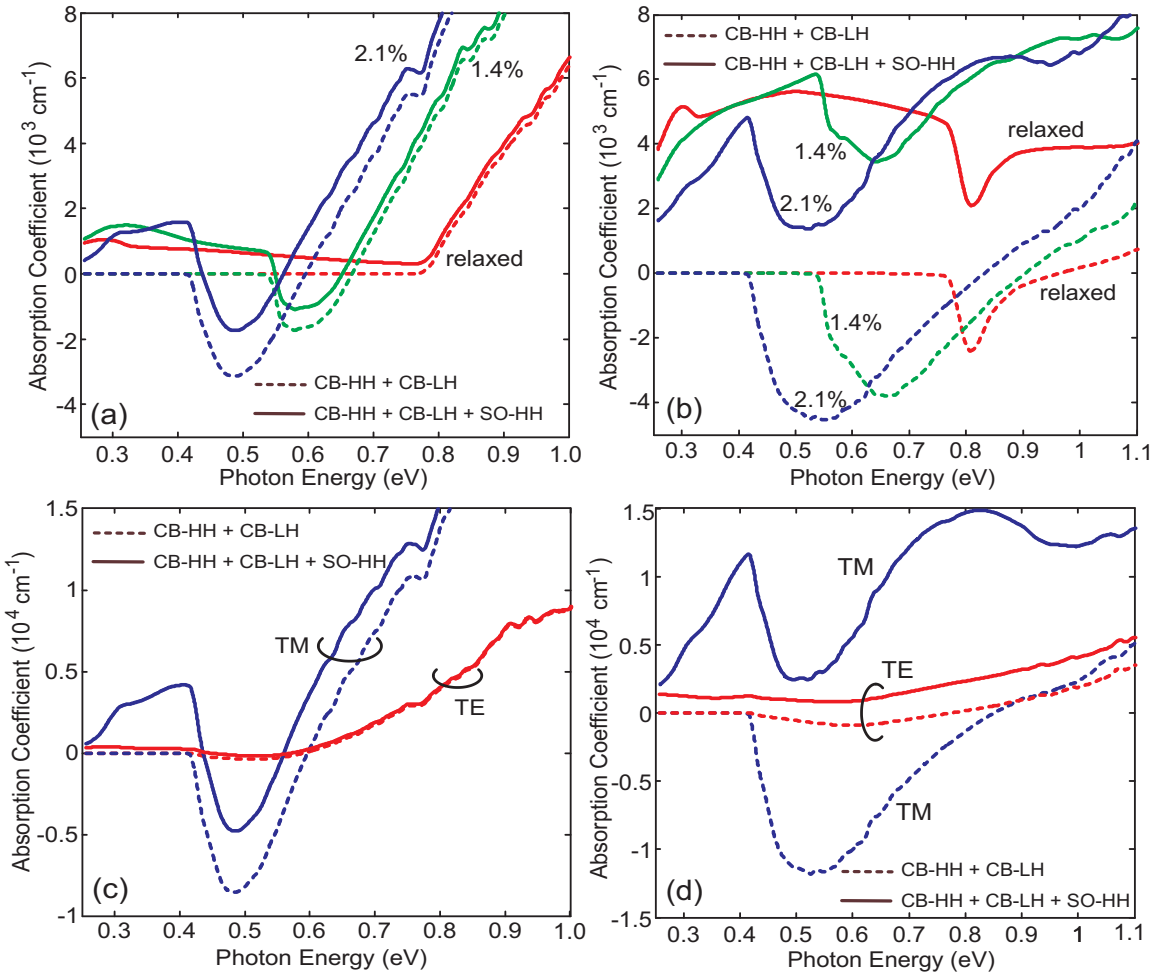


Figure 4·6: Absorption coefficient spectra for unpolarized [(a) and (b)] and polarized [(c) and (d)] light are examined with (doping, injection) of (10^{20}cm^{-3}, 10^{19}cm^{-3}) [for (a) and (c)] and (10^{20}cm^{-3}, 10^{20}cm^{-3}) [for (b) and (d)]. Throughout (a)-(d), dashed lines represent the results from CB-HH and CB-LH transitions, and the solid lines give the total absorption by adding the results from SO-HH process to the preceding quantities. Note that in (a) and (b) different colors stand for different tensile strains, while in (c) and (d) blue and red lines are for TM and TE polarized light respectively.

Our calculations presented so far indicate that highly strained Ge in conjunction with high injection is the best option to attain high optical gain. However, this is probably not the case in reality when all the other injection-induced absorption processes are taken into consideration. As an example, in addition to the previously

calculated gain from the recombination between CB-HH and CB-LH, we also analyzed the effect of the transition from the spin-orbit (SO) band to HH which is believed to be a major source of extra absorption, especially at high injection (Carroll et al., 2012).

Figs. 4·6(a) and (b) present the gain profiles for 1.4% and 2.1% tensile strained Ge calculated from the CB-HH and CB-LH transitions (dashed lines) as well as the results after adding the SO-HH contributions (solid lines). Same quantity for relaxed Ge are also included as a comparison. It can be immediately observed that the large optical gain achieved in the previous calculation is significantly reduced by the absorption process due to the SO-HH transition. For a carrier injection level of 10^{19}cm^{-3} the gain peak value suffers a 50% reduction for both 2.1% and 1.4% tensile strain conditions. As to an injection of 10^{20}cm^{-3}, the reduction is even more dramatic that, as illustrated in Fig. 4·6(b), the optical gain is completely canceled by the strong SO-HH absorption. Indeed, at such a high carrier injection level a large number of free holes will appear at the top of valence band which in turn enhances the hole transitions to the SO band. This effect has already been observed by Süess and coworkers (Ref. (Süess et al., 2013) Fig.5) in the case of a tensile strained Ge microbridge. As shown in Fig. 4·5, it is clear that higher injection will generally provide higher optical gain regardless of strain and doping conditions. However, the effect of SO-HH absorption process shows that there is a soft limit, typically around $5 \times 10^{19}\text{cm}^{-3}$, on the injection level that can be applied before the gain starts decreasing and eventually disappearing. On the other hand, although the SO-HH absorption can be minimized at relatively low carrier injection density, in order to achieve a usable optical gain the injection level cannot be lower than 10^{18}cm^{-3}. In fact, there are other types of absorption processes such as intra-conduction band absorption (ICBA) and intra-valence band absorption (IVBA) (both initial and final states are in the same band) (Liu et al., 2007) that need to be

overcome. As a comment to the SO-HH absorption, we noticed that this process is predominately due to the direct transition, for the indirect absorption only accounted for about 10^{-4} of the total absorption.

We have also investigated the effect of polarization $\hat{\mathbf{e}}$ on the gain profile for a fixed tensile strain of 2.1%, which is demonstrated in Figs. 4·6(c) and (d). A similar effect of SO-HH absorption is also observed as was the case for unpolarized light. In fact, due to this absorption, the optical gain for both TM and TE modes (calculated from Eq. (A.8)) is suppressed at an injection level of 10^{20}cm^{-3}. However, it should be noted that for the injection of 10^{19}cm^{-3}, the TM mode exhibited an optical gain of $4797\,\text{cm}^{-1}$ in contrast to a much lower value of $191\,\text{cm}^{-1}$ for TE mode as the result of transition selection rules (Chuang, 2009).

To conclude this section on the results pertaining to Ge, we would like to point out that, although there is a general agreement on the importance of free-carrier absorption (FCA) in the Ge lasing medium, different definitions of FCA have been used in the literature (Liu et al., 2007; Carroll et al., 2012; Adachi, 2005). In this paper, to facilitate the comparison of the absorption results from this work with others' we have treated the FCA as the combination of inter and intraconduction/valence band absorption. Consequently, the SO-HH transition is actually part of the FCA observed experimentally. As is indicated in Fig. 4·6, the SO-HH absorption is significant between a photon energy of 0.3-0.6 eV, but diminishes at the lower end which corresponds to the spin-orbit split energy. We also found that, the SO-HH absorption at the photon energy corresponding to the peak optical gain calculated using our numerical model, with an injection between 10^{19}cm^{-3} and 10^{20}cm^{-3}, reproduces the absorption cross section model presented in Ref. (Carroll et al., 2012), and the corresponding hole absorption cross sections of $\sigma_h = 4.4 \times 10^{-3}\text{nm}^2$ and $9.43 \times 10^{-3}\text{nm}^2$ are obtained for relaxed and 1.4% strained Ge respectively. However, when the injection level

goes above 10^{20}cm^{-3}, the SO-HH absorption increases in a logarithmic fashion with the injection level rather than linearly. This can be understood by the small density of states at SO band. At high injection (e.g. 10^{20}cm^{-3}), even the SO band starts accumulating holes, which in turn reduces the available electrons that can be excited to the HH band, leading to a reduced absorption. On the other hand, when analyzing the experimental data (Newman and Tyler, 1957; Carroll et al., 2012), the relation of FCA-hole concentration is almost linear even above $p = 10^{20}\text{cm}^{-3}$, suggesting that other type of absorption, namely the IVBA, is dominating. Indeed, the IVBA and ICBA can usually be modeled with the simple Drude model which gives a linear dependence of the absorption coefficient on the carrier concentration (Adachi, 2005). We then conclude that the SO-HH absorption is important when the hole concentration is below 10^{20}cm^{-3}. At higher carrier concentration, the IVBA and ICBA dominate over the SO-HH transition and need to be included in the calculation instead. To our knowledge no direct experimental data or theoretical model for IVBA and ICBA in strained Ge has been established at this time, making it difficult to estimate the overall FCA contribution. However, if one considers the experimental data of relaxed Ge available in the literature (W.G. Spitzer and Logan, 1961; Newman and Tyler, 1957), it is clear that IVBA and ICBA will lead to a significant gain reduction for photon energies below 0.3 eV. Consequently, we speculated that the same could be true for strained Ge as well, which could make reaching lasing conditions in Ge even more challenging. As a final remark, it is worthy of note that at the highest doping and injection conditions, Thomas-Fermi screening may not be adequate to describe the screening effect in the material. Instead, a more rigorous dynamic screening involving electron-plasma interaction should be included in the electron Green's function. As shown by Bardyszewski and coworkers (Bardyszewski and Yevick, 1989), with this additional correction the optical gain from the CB-HH and CB-LH transitions becomes

slightly larger. Nevertheless, our approach provides a baseline value for the gain and captures the most important physical phenomena which are responsible for the optical properties of strained Ge under various doping/injection conditions.

4.2 Optical Absorption of Si

4.2.1 Absoption Coefficients in Si and Their Comparison with Experimental Data

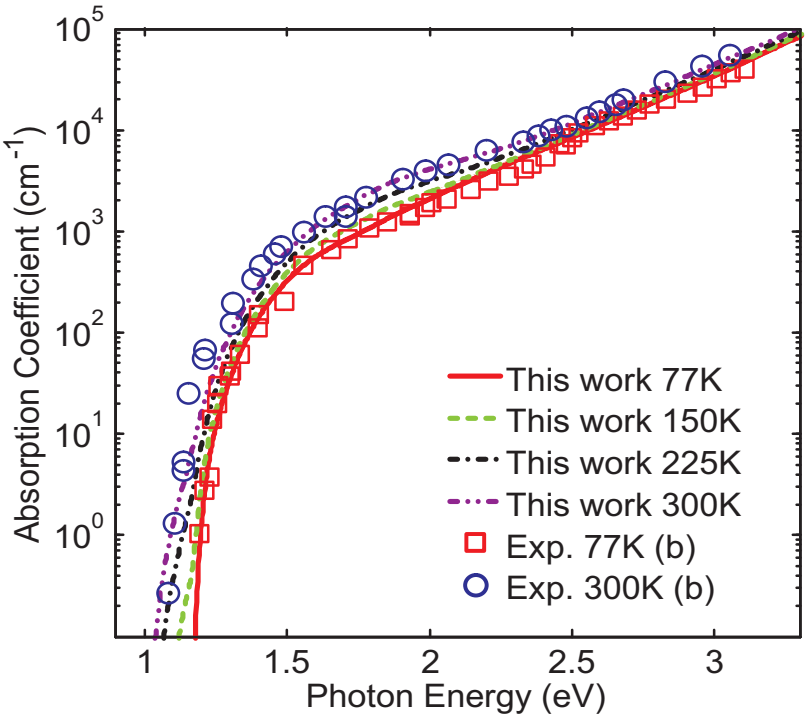


Figure 4·7: Absorption coefficient spectra for relaxed intrinsic Si at thermal equilibrium obtained from Green's function model are plotted with different lattice temperatures (colored lines). Experimental data (Dash and Newman, 1955) measured at the same condition are labeled by squares (77 K) and circles (300 K) as a comparison.

In this section we present the calculated optical properties of relaxed and strained Si using the same numerical model as we employed for Ge. Fig. 4·7 demonstrated the absorption coefficient for unstrained intrinsic Si at different lattice temperatures.

In order to compare the numerical results with the experimental data, the data of Dash and coworkers (Dash and Newman, 1955) measured at 77 K (open squares) and 300 K (open circles) is also plotted in Fig. 4·7. One can see that the results from the numerical model agree very well with the measured data. When compared to the same results obtained for Ge, we notice that for Si the PA indirect absorption process dominates the photon absorption below 3.4 eV, whereas in Ge both direct and PA indirect transition contribute to the total absorption. This makes the absorption spectrum for Si increase slowly with the photon energy in contrast to the abrupt rise observed in the Ge spectrum. The absorption spectra for different values of tensile strain have also been evaluated and presented in Fig. 4·8 for $\epsilon = 0\%$, 0.7%, 1.4% and 2.1%. As is shown, when a tensile strain is applied, the absorption spectrum shifts to lower photon energy while its shape remains similar since the strain-induced shrinkage of the direct band-gap has no impact on the absorption in the interested range of photon energy.

As was explained in Chapter 2, Section 2.4, the spectral density function employed in the model describes the energy broadening of the states due to different physical phenomena, for example phonon-electron interaction in this work. Therefore, by substituting a specific spectral density function in Eqs. (2.18) and (2.21) with a delta function, one can eliminate the PA process that pertains to a selected band transition. This enables the study of radiative processes involving multiple phonons. In particular, it is possible to study one-phonon and two-phonon assisted absorption processes in the current numerical model since two spectral functions are involved in Eqs. (2.18) and (2.21). Using this approach we have computed the temperature dependence of the radiative recombination rate in relaxed Si which is plotted in Fig. 4·9 with each curve obtained at a different lattice temperature. The Si is assumed to have an n-type doping of 10^{19}cm^{-3} and injection of 10^{18}cm^{-3}. We can notice that besides

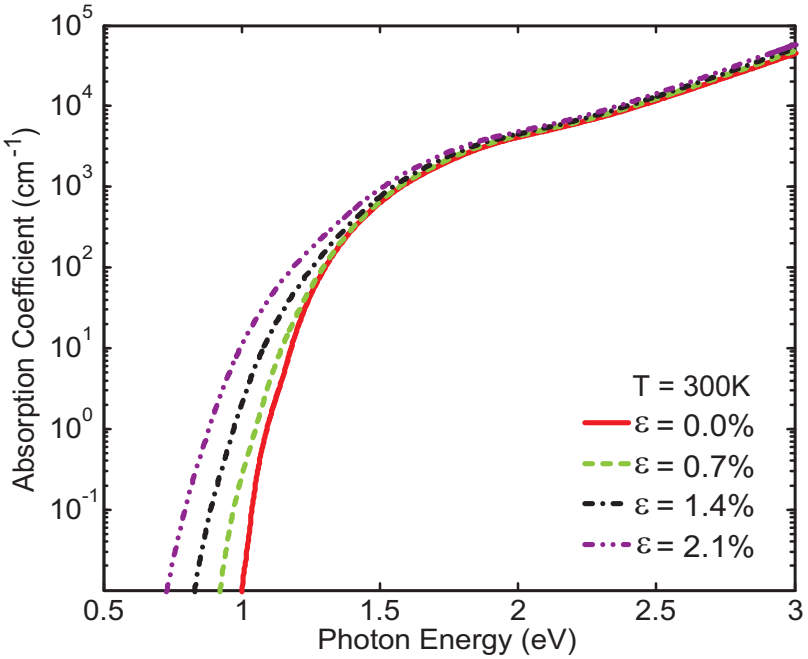


Figure 4·8: Absorption coefficient spectra for intrinsic Si at thermal equilibrium are calculated under the tensile strain of $\epsilon = 0\%$ (solid line), $\epsilon = 0.7\%$ (dashed line), $\epsilon = 1.4\%$ (dash-dot line) and $\epsilon = 2.1\%$ (dash-two-dot line). Lattice temperature is set at 300K.

the major peak, a secondary peak to the left of the curves also appears at all the temperatures considered. To investigate the physical origin of the smaller peak, we have calculated the contribution of the two-phonon process and the two one-phonon processes by replacing one of the electron spectral density functions with a delta function. The results show that the two one-phonon processes separately contribute to the main peak, while the two-phonon process is responsible for the smaller peak only. In fact, the separation of the two peaks is about 60 meV, approximately equal to the energy of one optical phonon in Si, which underscores the fact that the lower peak actually represents the two-phonon event in PA process. Further validation of this result is obtained from the experimental work of Green and coworkers (Green et al., 2001; Trupke et al., 2003). Fig. 4·9 reports the calculated data together with the experimental data (in arbitrary units in Fig.2(b) of Ref. (Green et al., 2001))

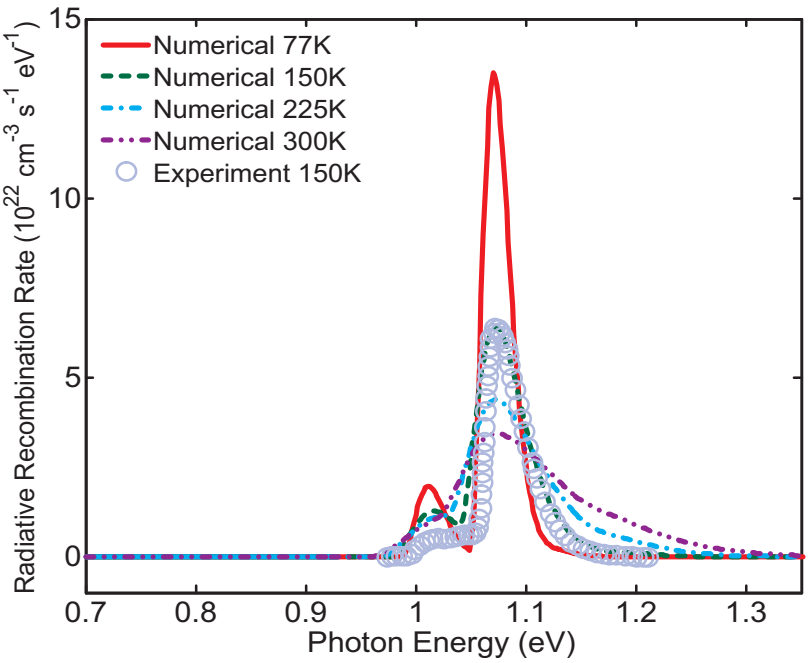


Figure 4·9: The spectra of radiative recombination rate for relaxed Si are computed with the lattice temperatures ranging from 77 K to 300 K. The Si is assumed to have an n-type doping of 10^{19}cm^{-3} and injection of 10^{18}cm^{-3}. Experimental data (Green et al., 2001) from the electroluminescence (EL) spectrum of a Si diode measured at 150 K are excerpted and converted to the same scale as the numerical result. The converted EL data (a.u.), shown in circles, has been multiplied by a constant number to fit the peak of the numerical result which is calculated at the same temperature, 150 K.

which is scaled to match the maximum value of the numerical result at 150 K. While a direct quantitative comparison between our numerical data and the experimental measurements may not be possible due to the lack of information about the bias level and and the device geometry, we can see that the shape of the calculated radiative recombination curve (dashed line) and the position of the peaks are very similar to the data measured at the same temperature (150 K, open circles).

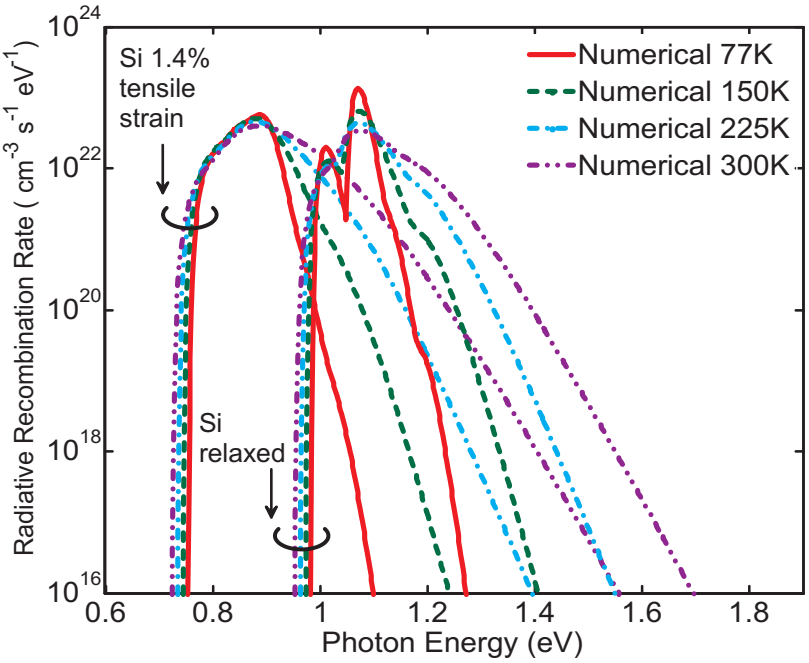


Figure 4·10: The spectra of the radiative recombination rate for relaxed (right group) and 1.4% tensile strained (left group) Si are computed with the lattice temperatures ranging from 77 K to 300 K.

4.2.2 Radiative Recombination Rates of Si and Two-Phonon Process

In addition to the calculated radiative recombination rates for relaxed Si, Fig. 4·10 presents the corresponding values when a tensile strain of $\epsilon = 1.4\%$ is applied to the system. It is noteworthy that the same low temperature carrier concentration effect, that we have mentioned in the discussion of Fig. 4·4, also appears in the case of relaxed Si, leading to a higher and narrower recombination peak at lower temperature. As to the results of 1.4% tensile strained Si, it is shown that the one phonon peak has broadened so much that it almost covered the two-phonon peak. This can be understood by considering the change in the density of states associated to the six valleys at the conduction band minima, Δ-point. When the strain is applied, the 6-fold Δ valleys split into a 4-fold lower valleys and a 2-fold upper valleys, reducing the actual density of state at the conduction band edge. Consequently, in order to

maintain the same carrier concentration, electrons will relocate to the higher energy states, leading to a broader radiative recombination spectrum. We want to conclude this section by comparing the spectrum for Ge presented in Fig. 4·4 with the one computed for Si, where it can be noticed that the two-phonon peak never appears in the plots of Ge, even in the case of relaxed material. This difference is due to several reasons. First, the conduction band density of states of relaxed Ge is similar to the one of $\epsilon = 1.4\%$ tensile strained Si. At the same doping/injection conditions, the broadening of the one-phonon peak in Ge is comparable to the strained Si, and same as the case in Fig. 4·10, the two-phonon peak is engulfed. Second, the small optical phonon energy in Ge makes the two-phonon peak even closer to one-phonon peak than in the case of Si. As an additional effect, this leads to the former one being even harder to distinguish from the total spectrum.

4.3 Conclusions of the Chapter

As the optical integrated circuit becomes increasingly important in building the next generation communication systems, a proper laser source that is compatible with the standard silicon process is highly desirable. In this chapter, we have investigated the optical properties of strained and relaxed Ge and Si for their potential applications in lasing medium.

Using the proposed numerical model described in Chapter 2, Section 2.2, and the full electronic band structures constructed in Chapter 3, we have calculated the radiative recombination rate of strained Ge to investigate the effect of different doping levels, injection and temperatures on the PA indirect and direct processes. These results provide additional insight into the possible ways to increase the photon emission in the material. The calculated optical gain for Ge indicates that while high carrier injection δn in tensile strained material could effectively enhance the gain through

the transition of CB-HH and CB-LH, extremely high δn, regardless of the tensile strain, would inevitably suppress the total optical gain by introducing strong SO-HH absorption. Specifically, at δn in excess of 10^{20}cm^{-3}, the attainable gain from CB-HH and CB-LH is totally canceled by the SO-HH transition, leaving a net absorption in the material. We have also evaluated the gain for TE and TM polarized light. A total gain of $4749\,\text{cm}^{-1}$ is predicted by our calculation for TM light in 2.1% tensilely strained Ge with n-type doping of 10^{20}cm^{-3} and injection of 10^{19}cm^{-3}. However, by adding other types of absorption such as free carrier absorption, the achievable optical gain will be reduced.

The radiative recombination properties of relaxed and strained Si were studied as well. We have demonstrated that the proposed model can indeed elucidate and reproduce the experimental data for the two-phonon events in the PA recombination process. The numerical model presented in this work is proved to be an effective tool to investigate the optical processes in indirect-gap semiconductors. In addition, this model can be further extended to inspect the non-radiative recombination mechanisms, such as Auger process, which compete with the optical gain at high injection conditions.

Chapter 5

Optical Absorption and Intrinsic Carrier Lifetime in Compound Materials for IR Detection

Aiming to operate the next-generation IR detectors at higher temperature with improved quantum efficiency over a wide spectral range, especially in the MWIR and SWIR band, the detector community has proposed many novel detector architectures. Among them are barrier detectors, normally referred to as XBn (Klipstein et al., 2010; D'Souza et al., 2011; Schuster et al., 2012) and photon-trapping structures (Keasler and Bellotti, 2011; Schuster and Bellotti, 2012) which exhibit high quantum efficiency, reduced crosstalk and low dark current compared to the conventional planar detectors (Schuster and Bellotti, 2013). A prerequisite to realize these novel detector structures with potentially superior performance is to understand the properties of the materials to be employed and how they influence the device operation. Of all the established and emerging narrow band-gap semiconductors, special consideration has been given to the lattice matched (LM) III-V ternary alloys. Specifically Indium Arsenide Antimonide (InAsSb), which is LM to Gallium Antimonide (GaSb), and Indium Gallium Arsenide (InGaAs) LM to Indium Phosphide (InP) have been employed alongside Mercury Cadmium Telluride (HgCdTe) for detector applications in the MWIR and SWIR spectral regions, respectively. Although, it is important to design and optimize the structure of a specific detector to exploit the unique features of a given material, the mechanisms that determine the intrinsic minority carrier lifetime impose a fundamental

limit to the device performance.(Edwall et al., 1998)

Unlike extrinsic carrier recombination mechanisms, such as Shockley-Read-Hall (SRH) recombination, that can be reduced by improving the material quality, Auger and radiative recombination processes are closely related to the details of the material's electronic structure and therefore are more difficult to engineer. As a result, comparing the intrinsic minority carrier lifetime of these narrow-gap III-V alloys with the widely used HgCdTe is important to understand the potential advantages of one specific material over the other for a given detector application.

A number of theoretical and experimental investigations have been performed to study Auger and radiative recombination processes and the minority carrier lifetime in HgCdTe, InAsSb and InGaAs. In particular, Krishnamurthy *et al.*(Krishnamurthy et al., 2006) and Grein *et al.*(Grein et al., 2008) respectively employed the hybrid pseudopotential tight binding and a 14-band $\mathbf{k}\cdot\mathbf{p}$ band structure and calculated the direct Auger and radiative lifetime in HgCdTe using Fermi's golden rule. With a similar formalism, Bertazzi *et al.*(Bertazzi et al., 2011) used band structures from the empirical pseudopotential method (EPM) and obtained comparable results for direct Auger-1 and Auger-7 lifetime in $Hg_{0.78}Cd_{0.22}Te$.

In the case of InAsSb alloys, the study of Auger recombination processes has been performed for superlattice structures (Lindle et al., 1995; Meyer et al., 1998; Steenbergen et al., 2011). Specifically, Grein *et al.*(Grein et al., 1995) theoretically compared the performances of type-II InAs/InGaSb superlattices to the conventional HgCdTe detectors operating at long wavelength IR (LWIR) using the $\mathbf{k}\cdot\mathbf{p}$ method and found that InAsSb-based detectors outperformed their HgCdTe-based counterparts due to their smaller Auger recombination rates. Vinter (Vinter, 2002) calculated the direct Auger coefficients in InAs, InSb and $InAs_{1-x}Sb_x$ alloys in both bulk materials and SL with an extended $\mathbf{k}\cdot\mathbf{p}$ band structures, where the results showed a good

agreement with the experimental data at 300 K. This indicated that the direct Auger process is the dominant Auger recombination mechanism at this temperature, but similar information at other temperatures in these materials is still missing.

As to InGaAs alloys, the majority of the studies of their Auger recombination have been performed using the parabolic band approximation (Bardyszewski and Yevick, 1985; Takeshima, 1982b; Takeshima, 1984a) to investigate their applications in lasers and high-speed photodetectors. However, except for some early works based on analytical models (Bardyszewski and Yevick, 1985; Takeshima, 1982b), few studies have focused on the phonon-assisted (PA) indirect Auger process which is believed to be important in InGaAs alloys (Bardyszewski and Yevick, 1985). In fact, it is still unclear whether PA indirect processes are important to understand the Auger and radiative properties in InGaAs especially with full-band structure based models, where the total density of states is different than that obtained from analytical bands.

In this chapter, a comparative study of the Auger and radiative recombination properties of HgCdTe, InAsSb and InGaAs will be performed to highlight potential advantages of one material system with respect to the others at given operating conditions. Specifically, both direct and PA indirect Auger as well as the radiative recombination processes are investigated using the Green's function formalism.(Wen and Bellotti, 2014; Wen and Bellotti, 2015b) Full band structures of each alloys, that are obtained using EPM, are also employed to overcome the limitations of analytical non-parabolic bands, leading to a more accurate evaluation of the interaction matrix element (Laks and Neumark, 1990; Harrison et al., 1999). By using the Green's function theory, the treatment of the divergences that appear in the second-order perturbation theory can be naturally avoided and the full band structure, as a result, can be incorporated into the calculation without *ad hoc* approximations. (Wen and Bellotti, 2014; Wen and Bellotti, 2015b).

In the first section, intrinsic minority carrier lifetime in $Hg_{0.78}Cd_{0.22}Te$ for LWIR application is demonstrated in terms of doping concentrations and lattice temperatures. Since few compound materials can compete with HgCdTe in LWIR spectral range, the obtained results represented an upper bound of minority carrier lifetime in the LWIR detectors at this moment. Section 5.2 and section 5.3 in this chapter presented a detailed comparison of both Auger and radiative lifetime in two sets of material compostions: $InAs_{0.91}Sb_{0.09}$ and $Hg_{0.67}Cd_{0.33}Te$ intended for detector operation in the MWIR and $In_{0.53}Ga_{0.47}As$ and $Hg_{0.38}Cd_{0.62}Te$ for SWIR, which have cutoff wavelengths of 4 μm and 1.7 μm at 200 K and 300 K respectively. In section 5.4, the overlap integral factor $|F_1F_2|$ for MWIR and SWIR materials that are needed to evaluate the Auger rates using the Beattie-Landsberg-Blakemore (BLB) expression will be derived by fitting the theoretical formula to the numerical Auger coefficients. This data will facilitate the researchers to set up a fast and accurate Auger recombination rate evaluation for device simulations without performing the full band calculation.

5.1 Intrinsic Minority Carrier Lifetime in HgCdTe for LWIR application

Since the $Hg_{1-x}Cd_xTe$ was first synthesized in 1958 by Lawson's research group (Lawson et al., 1959), it has been widely applied to the field of infrared (IR) detectors due to its adjustable energy gap with cut-off wavelength from 0.7 to 30 μm and nearly lattice matched for all the composition range (Norton, 2002). Alongside with the sustained effort to develop new device concepts, such as barrier detectors (Klipstein et al., 2011; Schuster et al., 2012) and photon-trapping (PT) structures (Keasler and Bellotti, 2011; Schuster and Bellotti, 2012), to improve the quantum efficiency, lower the dark current and raise the operating temperature, the study of the physical phenomena that determine the excess carrier lifetime is always an active topic since it imposes a

fundamental limit on the performance of all kinds of IR detectors (Edwall et al., 1998). Three recombination mechanisms, radiative, Auger and Shockley-Read-Hall (SRH), are generally considered as the dominant factors in determining the excess carrier lifetime, while the former two mechanisms are governed by the intrinsic properties of the material, e.g. composition and doping concentration, SRH recombination is typically affected by the impurities and defects formed in HgCdTe. Due to the significant improvement in the HgCdTe material quality, the SRH recombination rate has been greatly reduced as a result of the small number of defects present in the crystal. Consequently, Auger and radiative recombination have become the limiting factors in determining the minority carrier lifetime. Many theoretical studies have been performed to calculate the carrier lifetime due to Auger processes. Among them, Jiang *et al.* (Jiang et al., 1991) and Lopes *et al.* (Lopes et al., 1993) adopted the parabolic band approximation to simplify the evaluation of the recombination rate, while Bertazzi *et al.* and Krishnamurthy *et al.* (Krishnamurthy et al., 2005) employed the full band structure. In particular, Bertazzi and his collaborators (Bertazzi et al., 2011) have shown that the Auger-1 and Auger-7 recombination rate in n-type and p-type HgCdTe can be computed accurately using this approach.

Due to the ever pressing demand to increase the operating temperature of the state-of-the-art HgCdTe IR detectors, Auger and radiative recombination may become the dominant recombination mechanisms that limits the performance of devices (Rogalski, 2005). Specifically, Auger recombination (AR) is usually strong in the LWIR range, whereas for devices operating in the short-wavelength infrared (SWIR) and mid-wavelength infrared (MWIR) spectral regions, radiative recombination (RR) is dominant. The intrinsic limit of RR, however, could be significantly mitigated by the photon-recycling effect, which was first proposed by Humphreys (Humphreys, 1983) in 1983. Unfortunately, due to the strong coupling between photon and the device

boundary as well as the complexity of photon-carrier interaction processes, only a handful of studies, which include a number of approximations, have been performed to elucidate this phenomena (Jozwikowski et al., 2011; Jozwikowski et al., 2012). In any case, a prerequisite to investigate the photon-recycling effect is to develop an accurate model for the radiative recombination process. Furthermore, this model should be valid for all the relevant alloy compositions and doping levels. The majority models of the radiative recombination rate and lifetime developed up to now are based on the classical theory of van Roosbroeck and Shockley (VRS) (van Roosbroeck and Shockley, 1954) and an improved version developed by Finkman (Finkman and Nemirovsky, 1979), both of which were derived under the parabolic band approximation. An exception to them was the study performed by Krishnamurthy, (Krishnamurthy et al., 2005) who used Fermi's Golden rule and incorporated the full band structures of the material in the calculation. Despite the complexity of introducing the full band structures in the calculation of the radiative recombination lifetime, the advantages are obvious. A model based on the realistic electronic structure can be used to study both direct and indirect gap semiconductors and the anisotropic nature of valence band of HgCdTe will be automatically included. This avoids the ambiguities when using the VRS or Finkman's theory, where fixed values of m^*_{HH} and m^*_{LH} need to be chosen from various values reported from the experimental data (Capper, 1994).

5.1.1 Radiative Recombination in HgCdTe

Previous studies of the radiative lifetime have predominantly used the measured value of the absorption coefficient or have computed it assuming a parabolic band approximation (Kinch et al., 1973; Pratt and Hewett, 1983). The assumption of parabolic bands, though simple and accurate to some degree, requires that one has accurate values of both conduction and valence band effective masses, which is not always the case. Furthermore, due to the anisotropic nature of the heavy hole band,

researchers have reported many different values of the effective mass for this band ranging from $0.4m_0$ to $0.7m_0$ (Capper, 1994). Additionally, as a result of the fact that the shape of conduction band is more hyperbolic-like than parabolic-like, the electron effective mass can only be defined accurately very close to the band edge. Consequently, it is not always obvious how to choose proper effective masses in order to get a reliable value of $\alpha(E)$. Krishnamurthy (Krishnamurthy, 1996) for the first time applied the full band structures of HgCdTe to Fermi's Golden rule and computed the radiative recombination rate for the direct inter-band transition. Subsequently, Grein *et al.* (Grein et al., 2008; Chang et al., 2006) later developed an analytical model based on $\mathbf{k} \cdot \mathbf{p}$ method and the hyperbolic band structure, which gave a good fit to the experimental results. In this work, we go beyond the current state of the art and implement a more sophisticated numerical model based on Green's function formalism which is applicable to both direct and indirect, as well as narrow and wide gap semiconductors.

We used the Green's function theory outlined in Chapter 2, Section 2.2 which incorporates the full band structure (obtained from a nonlocal EPM) to evaluate both direct and phonon-assisted indirect optical (radiative) transitions. Traditional second-order perturbation theory which is widely used in studying the indirect transitions, is known as having the divergence difficulty when a virtual state becomes a real one (Bude et al., 1992). Depending on the way one chooses to deal with it, the final result would inevitably change thus introducing additional uncertainties to the theory (Takeshima, 1981). This situation would become even worse when a full band structure is employed as the denominator in the matrix element going to zero more frequently and more unpredictably. However, by using the Green's function theory, this difficulty can be avoided naturally and even higher order processes (e.g. two phonon-assisted transition) could be incorporated into the same theoretical framework. With this consideration,

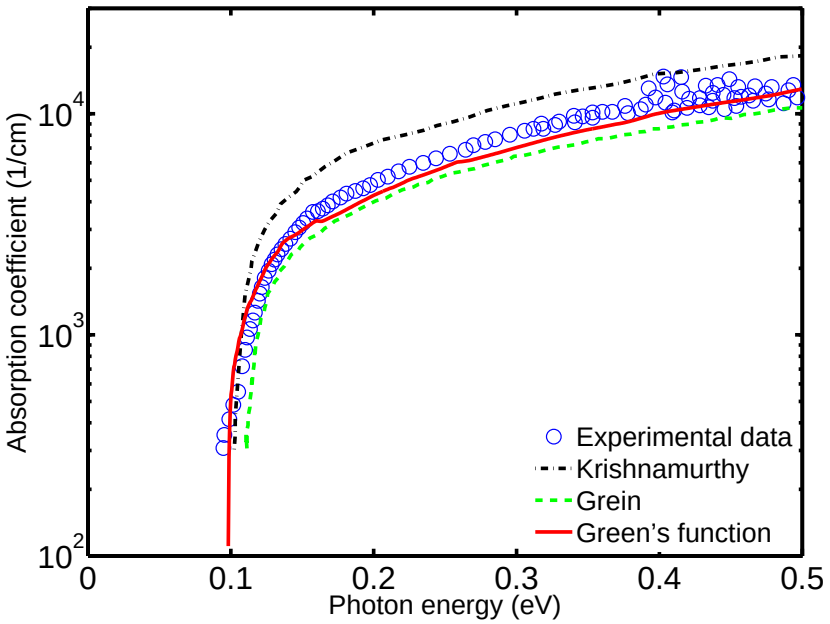


Figure 5·1: Absorption coefficient for $Hg_{0.79}Cd_{0.21}Te$ at 77 K from three different numerical models. Solid line is the result from Green's function formalism, while the dashed line and the dash-dot line represent Grein's (Grein et al., 2008) and Kirshnamurthy's (Krishnamurthy, 1996) models respectively. Circles are the experimental data for the same material measured at 80 K (Chang et al., 2006).

the numerical model we have developed is suitable for studying radiative recombination processes for both direct and indirect material with high accuracy. Using this model, we first computed the absorption coefficients and radiative recombination lifetimes of HgCdTe for three different cadmium compositions ($x = 0.23$, 0.35 and 0.4) and compared them with the theoretical expression that is widely used in the device simulation community (Lopes et al., 1993). After that, the temperature dependence of the radiative lifetime for the same compositions at different carrier concentrations is investigated and presented.

While it has been pointed out in previous works (Krishnamurthy et al., 2005) that the parabolic band approximation does not work very well at the absorption edge due to the fact that the shape of conduction band is actually hyperbolic and the heavy

hole band is anisotropic, it is worth comparing the results from Green's function with other numerical models which use non-parabolic band structure. Krishnamurthy *et al.* (Krishnamurthy, 1996) and Grein *et al.* (Grein et al., 2008) independently investigated this problem with full band structures and 14-band $\mathbf{k} \cdot \mathbf{p}$ models respectively. Their results as well as the experimental data (Chang et al., 2006) together are compared with our numerical model in Fig. 5·1 for $Hg_{0.79}Cd_{0.21}Te$ at 77 K. Despite the Urbach tail, which none of these models considered, it can be seen that the Green's function generally gives a better match in the whole range of photon energy, while for the higher photon energy, both Grein's result and Green's function give smaller values than the experimental measurement. Krishnamurthy's result, on the contrary, is consistently larger. Since the HgCdTe is a direct band gap material, we noticed that the contribution from the direct transition process is predominant in our result as the the phonon-assisted indirect transition only contributes less than 1% of the total absorption. The low operating temperature further reduced the phonon-assisted process by greatly suppressing the number of phonons in the crystal thus making it negligible in low temperature. This result agrees with a similar calculation done by F. Bertazzi on the Auger recombination in HgCdTe (unpublished), where the phonon-assisted process was also found minor compared to the direct process.

In addition to the absorption coefficients we have also evaluated the radiative recombination rates and the associated lifetimes. The calculated lifetime (symbols) as a function of carrier concentration for the three compositions at 90K is shown in Fig. 5·2. In the same figure we also report the values (lines) computed using the known analytical expression from Lopes *et al.*(Lopes et al., 1993). As pointed out earlier, in using this formula it is necessary to select the values of the effective masses, which is in general not possible for holes since valence bands are usually highly anisotropic. We found that by choosing the effective mass of holes as the geometric mean of the effective

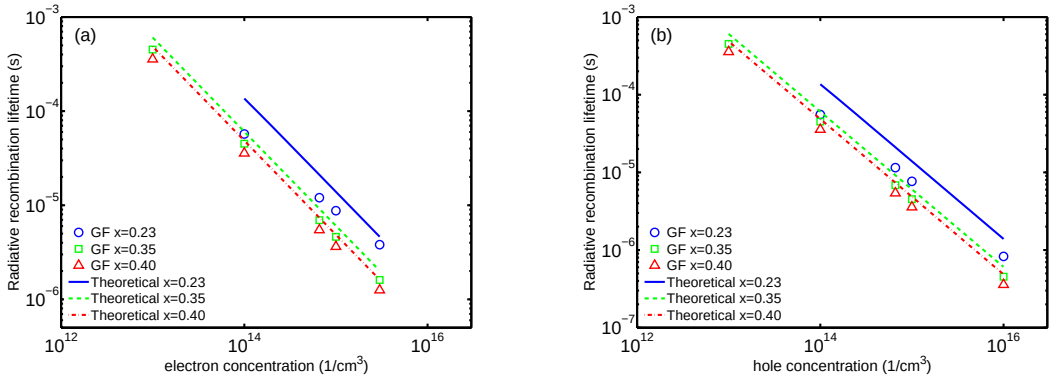


Figure 5·2: Radiative recombination lifetime as a function of carrier concentration for three compositions of $Hg_{1-x}Cd_xTe$: $x = 0.23$, $x = 0.35$ and $x = 0.40$ at 90K with (a): n-type doping and (b): p-type doping. Symbols in the figure represent results from Green's function formalism and lines are the theoretical formula mentioned in Lopes *et al.* (Lopes et al., 1993).

masses obtained from the EPM heavy hole band in [100], [110] and [111] direction, the two results (numerical and analytical) almost match to within a consistent difference of 30%. One exception here is the composition of $x = 0.23$ where the intrinsic carrier concentration is comparable to the extrinsic concentration making the concept of lifetime itself not well-defined, as one can see the results from Green's function have deviated from a straight line especially in the lower carrier density region.

For the theoretical values represented by the lines in Fig. 5·2, an effective mass of $0.49m_0$ for holes is used for all three compositions while the effective mass for electrons is calculated according to Weiler (Weiler, 1981). In fact, from our non-local EPM band structure, the actual effective mass for the heavy hole band is $0.58m_0$ in the $\Gamma - X$ direction, $0.49m_0$ in $\Gamma - K$ direction, and $0.41m_0$ in $\Gamma - L$ direction, whose geometric mean is $0.488m_0$. As is stated above, even though the two methods used the same effective mass of holes, their results differed. Further reducing the effective mass of holes to $0.4m_0$ in the theoretical formula would yield a perfect match of radiative lifetime between the two models. These facts indicate that though the analytical

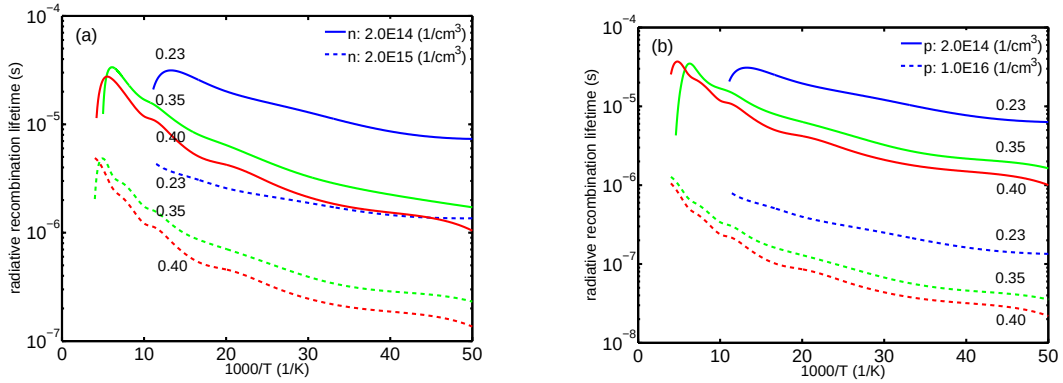


Figure 5·3: Temperature dependence of radiative recombination lifetime for (a): n-type and (b): p-type $Hg_{1-x}Cd_xTe$. Solid lines represent carrier concentration of $2.0 \times 10^{14}\text{cm}^{-3}$ while the dashed lines are for carrier concentrations of (a): $2.0 \times 10^{15}\text{cm}^{-3}$ and (b): $1.0 \times 10^{16}\text{cm}^{-3}$. The number around each curve gives the corresponding composition ratio of cadmium x.

formula does not agree with Green's function model, by choosing the effective mass of holes as $0.4m_0$ at 90 K, the simple theoretical formula will give a good estimation of the radiative lifetime which is important for the device simulation where a fast and accurate lifetime model is desirable.

We have also evaluated the temperature dependent radiative recombination lifetime for different carrier concentrations. Results are presented in Fig. 5·3 for n-type (a) and p-type (b) HgCdTe respectively. In Fig. 5·3(a) the calculated lifetime for two electron concentrations of $2.0 \times 10^{14}\text{cm}^{-3}$ (solid lines) and $2.0 \times 10^{15}\text{cm}^{-3}$ (dashed lines) are shown. The parameter on each curve represents the composition (cadmium molar fraction x) of the material. As the temperature decreases, the radiative recombination lifetime decreases slowly to several microseconds which is usually orders of magnitude higher than the lifetime of Auger recombination. On the other end of the curves, when the temperature approaches the intrinsic temperature of the material, there is a sharp decrease in the lifetime due to the presence of large amounts of thermal excited electron-hole pairs. This could potentially become the limiting factor of the

detector performance under conditions where SRH recombination is negligible and the photon-recycling effect is not significant. Similar results for p-type HgCdTe are presented in Fig. 5·3(b) with the hole concentrations of $2.0 \times 10^{14}\text{cm}^{-3}$ (solid lines) and $1.0 \times 10^{16}\text{cm}^{-3}$ (dashed lines). The trend for holes is similar to the one for electrons, with the only difference that the corresponding lifetime for holes is slightly shorter. The radiative lifetime also decreases as the hole concentration increases and the cadmium composition becomes larger, which is also shown in Fig. 5·2. Since we did not compare the dependence of lifetime on dopant types, the freeze-out effect, which will generally make the lifetime increase exponentially as the temperature decreases, was not included in the results above.

5.1.2 Auger Recombination in HgCdTe

Following the theory outlined in Chapter 2, Section 2.3, the direct Auger recombination lifetime in $Hg_{0.78}Cd_{0.22}Te$ is computed and compared to the values obtained in other theoretical works, including the BLB formula. Fig. 5·4 shows the temperature dependence of the minority carrier lifetime due to direct Auger process. In the figure, solid lines represent the results from this work, circle symbols show the 14-band $k \cdot p$ results from Grein *et al.*(Grein et al., 2008), dashed lines give the calculation results of Bertazzi *et al.*(Bertazzi et al., 2011) and dash-dot lines are from BLB formula. An overlap integral factor $|F_1F_2| = 0.16$ is used in the BLB formula to best fit to the numerical results.

As is shown, the three full band numerical results agrees well with each other in the high temperature range while at low temperatures some differences appeared between Grein's result and the one from this work. We argue that the differences may come from the details of the band structures used in these works. At high temperatures, since more electrons and holes present in the bands, the available Auger recombination final states are abundant making the contribution from each transition smaller. As a

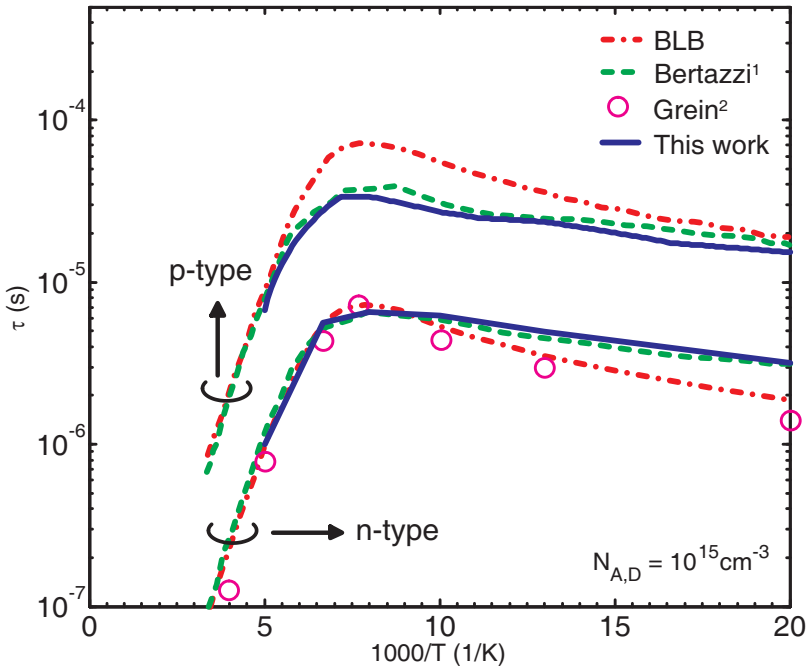


Figure 5·4: Calculated temperature dependence of direct Auger recombination lifetime in $Hg_{0.78}Cd_{0.22}Te$ from Green's function theory (solid lines), 1: Bertazzi's work (Bertazzi et al., 2011) (dashed lines), 2: Grein's work (Grein et al., 2008) (circles) and BLB formula with $|F_1F_2| = 0.16$ (dash-dot lines). The upper group of curves represent the results for p-type material (CHLH, Auger-7 process) while the lower group is for n-type material. The doping concentration in both cases is $10^{15}cm^{-3}$.

result, the impact of the band structure details, especially the states close to the band edge, on the total recombination rate reduced. On the contrary, at low temperatures, due to the fact that most of the electrons and holes are concentrated at the band edge, very few direct recombination paths can fulfill both momentum and energy conservation in the Auger process. This results in a total direct recombination rate that is sensitive to the nature of the involving energy values and wavefunctions. Since both this work and Bertazzi's calculation used the same EPM band, but a different formulation for the self-energy, the calculated carrier lifetime agrees better with each other, comparing to the one from the $\mathbf{k} \cdot \mathbf{p}$ bands. Similar behavior can also be

observed in the calculated hole lifetime which is shown in the upper group of curves. Furthermore, though not shown in the figure, the PA indirect contribution to the Auger recombination is also calculated along with the direct part and is found to be small compared to the direct process. In fact at 200 K, the difference between direct and total (direct + PA indirect) Auger lifetime is less than 7% which is within the error of the Monte-Carlo integration techniques. Comparing the numerical results to the BLB formula, it is obvious that the difference between numerical results and BLB values is larger at low temperature, which suggests that it may be necessary to use a temperature-dependent $|F_1F_2|$ in the BLB formula to fully reproduce the numerical results. Indeed, the overlap integral factor actually represents the coupling of the four Bloch states in the Auger process which will definitely change with the temperature through the change of the band gap energy. Consequently, it is of great importance to understand the temperature dependence of $|F_1F_2|$ in order to use the BLB formula effectively and accurately. Further information on the overlap integral factor $|F_1F_2|$ for HgCdTe can be found in Chapter 5, Section 5.2 and Section 5.3.

In summary, we have used the Green's function formalism and the full band structures and calculated both direct and indirect RR and AR processes in the HgCdTe for LWIR spectral range. The calculation of radiative lifetime for HgCdTe versus carrier concentration confirmed the accuracy of the widely used theoretical formula under the condition that the effective mass of holes in the formula must be carefully chosen. For example, at 90 K with $m_h = 0.4m_0$, the formula agrees well with our numerical calculation. Furthermore, the temperature dependence of the radiative recombination lifetime in HgCdTe was investigated for different Cadmium molar fractions and carrier concentrations. The results showed that when the temperature approached to the intrinsic temperature, a sharp decrease of lifetime appeared which would potentially limit the detector performance if the photon-recycling effect were

not significant. As to the calculation of Auger recombination lifetime in HgCdTe, the obtained results further confirmed the accuracy of the Green's function theory and showed the importance of using a temperature-dependent overlap integral factor in the BLB formula.

5.2 Intrinsic Minority Carrier Lifetime in Strained/Relaxed InAsSb and its Comparison with HgCdTe for MWIR application

Antimony based compounds have become key building blocks in a number of optoelectronic devices (Rogalski, 1994). Among them, $InAs_{1-x}Sb_x$ alloys grown on InAs or GaSb substrates which are employed in barrier (Klipstein et al., 2010; D'Souza et al., 2011) and strained layer superlattices (SLS) MWIR photodetectors (Haddadi et al., 2014), have become technologically important due to their high optical quality and material epitaxy layer uniformity (Ivanov et al., 2005). In spite of their increasing use, the understanding of the optical and electronic properties of relaxed, and more importantly biaxial strained, $InAs_{1-x}Sb_x$ alloys has not been developed apace yet. In fact, only a small number of studies have been devoted to investigate the radiative and non-radiative recombination processes in these materials (Vinter, 2002; Olson et al., 2013) and significant uncertainties about the theoretical values regarding to the measured data still present. As to the minority carrier lifetime in strained $InAs_{1-x}Sb_x$ alloys, which is a critical ingredient for the design and optimization of photon detectors' performances, even fewer theoretical works have been published.

This section intends to contribute to the understanding of the optical and electronic properties of relaxed and strained $InAs_{1-x}Sb_x$ alloys by employing the full band model described in Chapter 2, Section 2.2 and 2.3. Using the Green's function theory, both direct and phonon-assisted (PA) indirect Auger recombination as well as the radiative

recombination processes are studied in strained and lattice-matched $InAs_{1-x}Sb_x$ alloys. In particular, the intrinsic minority carrier lifetimes in n-type and p-type materials are computed as functions of lattice temperatures and antimony molar fractions. The Auger coefficients computed for the direct and PA indirect processes are also compared to determine their relative importance for different alloys composition. The results show that strained $InAs_{0.91}Sb_{0.09}$ grown on an InAs substrate has a smaller Auger recombination coefficient when compared to the lattice-matched $InAs_{0.91}Sb_{0.09}$ grown on a GaSb substrate. This difference could lead to further improvement of the InAsSb photon detectors.

5.2.1 n and p-type Relaxed InAsSb and HgCdTe

One of the III-V materials emerging as a possible competitor to HgCdTe and InSb in the MWIR is $InAs_{1-x}Sb_x$. Among all the compositions in the $0 < x < 1$ range, the lattice-matched $InAs_{0.91}Sb_{0.09}$/GaSb system is of great importance due to the possibility of growing high quality barrier detector structures and type-2 superlattices (T2SL). In this section, we investigated the radiative and Auger recombination lifetime in n-type $InAs_{0.91}Sb_{0.09}$ and compared it to the corresponding values for $Hg_{0.67}Cd_{0.33}Te$, since both of them have a cutoff wavelength of 4μm at 200 K.

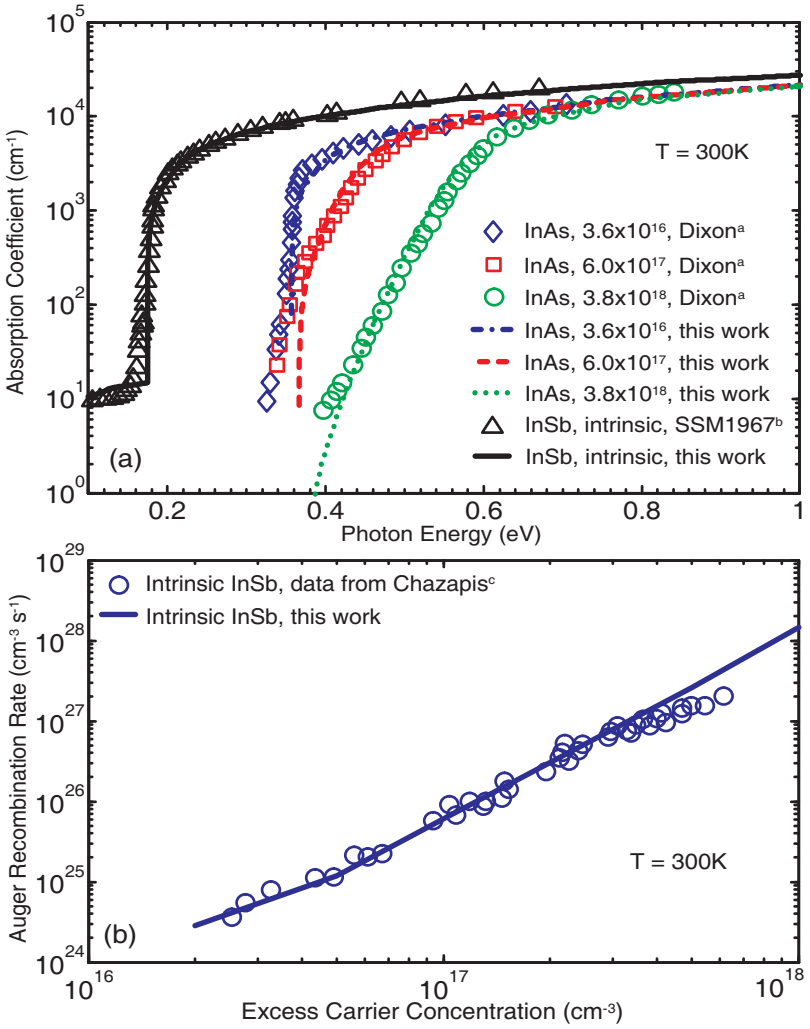


Figure 5·5: (a) Absorption coefficients and (b) Auger recombination rates in n-type InAs and InSb at 300 K are computed using Green's function theory and compared with the experimental results (symbols are from a: Ref. (Dixon and Ellis, 1961), b: Ref. (Willardson and Beer, 1967) and c: Ref. (Chazapis et al., 1995)). In (b), the Auger recombination rates for InSb are converted from the measured photoluminescence decay rates.

To assess the accuracy of the computational model and the fitted EPM band structure described in Chapter 2 and 3, we have performed a number of tests in which the theoretical results are compared with the experimental data obtained for the relevant semiconductor materials. (Wen and Bellotti, 2015b; Wen et al., 2015)

Fig. 5·5(a) presents the computed absorption coefficients for the binary materials evaluated at different doping concentrations and plotted with the corresponding experimental results. As shown, the theoretical values of the absorption coefficients computed for InAs match very well with the measured data at all the doping levels considered. In fact, since the Fermi level employed in our calculation is evaluated from the given doping concentration using the full-band structures, the Moss-Burstein effect, which appears in heavily doped materials, is properly accounted for in the calculation. As the data represented by the open circle shows, at an n-type doping of $3.8 \times 10^{18}\text{cm}^{-3}$, the absorption edge moves considerably toward higher photon energy when compared to the case of $3.6 \times 10^{16}\text{cm}^{-3}$ doping. It is also interesting to see that for InSb, the small tail at the photon energy below the band gap is properly reproduced, which, according to our calculation, is due to the free carrier absorption between the spin-orbit band and the heavy hole band. Finally, we need to point out that the Urbach tail below the band gap energy which comes from the disorder effects in the samples is not included in our model.

Fig. 5·5(b) presents the total Auger rates computed for intrinsic InSb as a function of injection level. It can be seen that the calculated values correctly predicted the measured Auger rates from an injection of $2.0 \times 10^{16}\text{cm}^{-3}$ to $5.0 \times 10^{17}\text{cm}^{-3}$ obtained from photoluminescence decay data reported in Ref. (Chazapis et al., 1995) at 300 K. In a recent work (Wen et al., 2015), we also showed that the calculated minority carrier lifetime in relaxed bulk $\text{InAs}_{0.91}\text{Sb}_{0.09}$ alloy is in good agreement with the measured lifetime from Ref. (Olson et al., 2013).

Before investigating the strained $\text{InAs}_{1-x}\text{Sb}_x$ alloys, it is beneficial to study the relaxed structures first, to establish baseline values of the material properties. Fig. 5·6 shows the calculated Auger-1 (CHCC for n-type), Auger-7 (CHLH for p-type) and radiative recombination coefficients for the relaxed alloys as functions of antimony

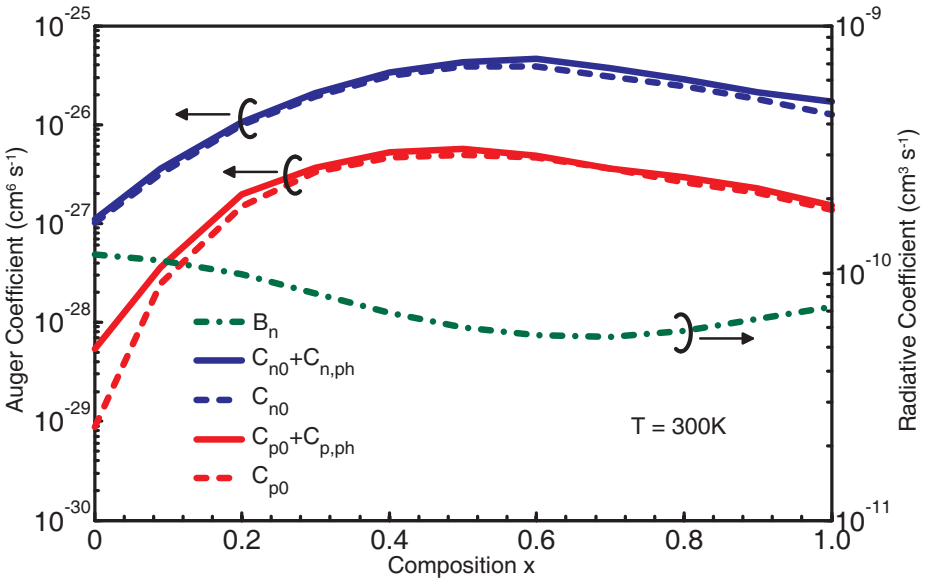


Figure 5·6: Auger coefficients (left y-axis) for CHCC (blue lines, for n-type) and CHLH (red lines, for p-type) processes and radiative coefficients (right y-axis) are plotted as functions of antimony composition in $InAs_{1-x}Sb_x$ alloys. The contributions from the direct Auger process to the total Auger coefficients (solid lines) are shown in dashed lines. The lattice temperature assumed in the calculation is 300 K.

compositions at room temperature. As can be seen, both direct and total (direct + PA indirect) Auger coefficients have a maximum around $x = 0.6$, which corresponds to the minimum of the band gap energy in InAsSb alloy. This result follows the known trends of the Auger process, where the recombination rate increases as the energy gap decreases. Similar theoretical results were also obtained by Vinter (Vinter, 2002) whose values of the direct Auger coefficients closely matched with ours. However, it should be noted that even including the PA indirect Auger processes, the calculated Auger coefficient for InAs is still smaller than some of the experimental data (Vodopyanov et al., 1992; Marchetti et al., 2002; Lindle et al., 1995), while for other measurements such as the one from the photoconductivity (Galkin et al., 1971), the obtained value is very close to what we have computed. Further calculation of the Auger S-process (CHSH, in which one of the holes transitions from the heavy-hole to the split-off

band) in intrinsic InAs also showed a negligible contribution (3.2%) to the total Auger coefficient. At this time, we found this discrepancy between the theoretical and experimental results remains an open question. Nevertheless, the calculated Auger coefficient for InSb still lies within the lower side of the measurement variability (Ciesla et al., 1996; Marchetti et al., 2001). In the case of p-type material, where the CHLH process is dominant, the corresponding Auger coefficient C_p is one order of magnitude lower than that in the n-type material as a result of the low number of final states in LH band. Although we can appreciate that in both n and p-type materials, when $0.2 < x < 0.8$, the contribution of the PA process to the total Auger recombination is not very significant, for the InAs-like ($x < 0.2$) and InSb-like ($x > 0.8$) alloys, the PA process is strong in the p-type and n-type materials respectively.

Fig. 5·6 also presents the calculated radiative recombination coefficient B_n (values on the right y-axis) for the relaxed alloys obtained from $R_r = B_n(np - n_i^2)$. In this expression, R_r, n, p and n_i are the radiative recombination rates, electron density, hole density and intrinsic carrier density respectively. The figure shows that the B_n slightly decreases from about $1.2 \times 10^{-10}\text{cm}^3/\text{s}$ for InAs to $5.5 \times 10^{-11}\text{cm}^3/\text{s}$ at $x = 0.7$ as the energy gap shrinks. This can be traced back to the reduced density of states of photons at lower photon energy. Nevertheless, the change of B_n as the antimony composition in the relaxed alloys is not very significant.

For the n-type InAsSb alloy, the corresponding Auger-1 lifetime as a function of lattice temperature are shown in Fig. 5·7. Depending on the recombination mechanisms used in the calculation, the total Auger coefficients, which consist of both phonon-assisted ($C_{n,ph}$) and direct (C_{n0}) Auger processes, are plotted alongside the direct Auger coefficients as a function of temperature. As can be seen, at high temperature range ($T > 200\,K$) when both materials become nearly intrinsic, the total Auger coefficient almost remains constant at $4.8 \times 10^{-27}\text{cm}^6/\text{s}$ and $3.7 \times 10^{-27}\text{cm}^6/\text{s}$ for

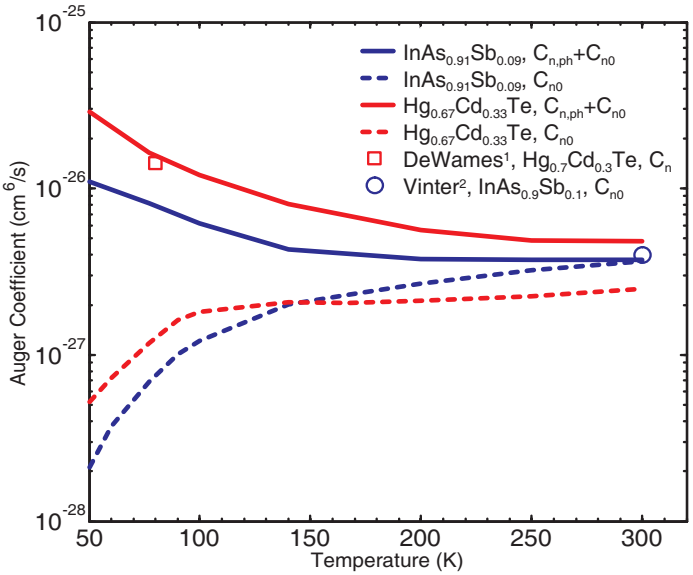


Figure 5·7: Temperature dependence of Auger-1 coefficients for $InAs_{0.91}Sb_{0.09}$ (blue lines) and $Hg_{0.67}Cd_{0.33}Te$ (red lines) are compared with 1 (square): fitted C_n of $Hg_{0.7}Cd_{0.3}Te$ at 80 K from DeWames's work (DeWames et al., 2011), and 2 (circle): calculated direct Auger coefficient C_{n0} in $InAs_{0.9}Sb_{0.1}$ at 300 K (Vinter, 2002). The solid lines here represent the total Auger coefficients (PA indirect+direct) and the dashed lines are computed from the direct Auger process only.

$Hg_{0.67}Cd_{0.33}Te$ and $InAs_{0.91}Sb_{0.09}$ respectively. The PA indirect Auger process in this temperature range does not provide a significant contribution to the total Auger rate as opposed to what happens at lower temperatures. For example, in the case of $InAs_{0.91}Sb_{0.09}$ the direct Auger coefficient almost equals to the total one at 300 K which also agrees with Vinter's calculation (circle) (Vinter, 2002). However, as the temperature decreases, the PA indirect Auger process becomes increasingly important as the direct Auger coefficient drops while the total Auger coefficient, on the contrary, increases. The square symbol in the figure shows the value of the fitted Auger coefficient for $Hg_{0.7}Cd_{0.3}Te$ from the experimental data with a donor concentration of $10^{15}cm^{-3}$ obtained by DeWames *et al.*(DeWames et al., 2011). We found that to reproduce the experimental value it is necessary to include the contribution of the indirect Auger process since the direct Auger coefficient at 80 K is one order of magnitude lower than

the measured data. Adding the indirect Auger process to the calculation leads to a much better agreement between the calculated and measured total Auger coefficient and we believe that the small remaining differences between them comes from small alloy composition differences and numerical errors.

The calculated radiative and Auger recombination rates are presented as a function of lattice temperature and doping concentration in Fig. 5·8(a) and (b) respectively. As is shown, though diminishing at higher doping concentration, the PA indirect Auger process contributes a large portion to the total Auger recombination lifetime in $Hg_{0.67}Cd_{0.33}Te$ over the doping range from $10^{14}cm^{-3}$ to $5 \times 10^{16}cm^{-3}$. In fact, at $N_D = 10^{14}cm^{-3}$, the indirect Auger rate is twice as large as the direct Auger rate whereas at $N_D = 10^{16}cm^{-3}$ the two processes almost contribute equally to the total recombination rate. In the case of $InAs_{0.91}Sb_{0.09}$, the relative strength of indirect Auger to the direct Auger is smaller and it can be seen that at high doping concentration, the contribution from the indirect process almost vanished. This variation of the relative strength between direct and indirect Auger rate with the doping concentration can be qualitatively explained by considering the available recombination paths. As was stated above, the low probability of indirect Auger process can actually be compensated by the large availability of final states in k-space. When the doping concentration is low, carriers tend to occupy the states at the band edge where the available direct recombination path is limited. As to indirect processes, with the assistance of phonon scattering carriers can firstly be scattered to a state far from the band edge and recombines with the final states that fulfills both energy and momentum conservation. Since the number of final states in this case is usually abundant, the total recombination rate therefore increases. Consequently, as the direct process is limited by the available final states, the indirect process becomes more important at low doping level. A similar analysis can also be applied to the case of high doping

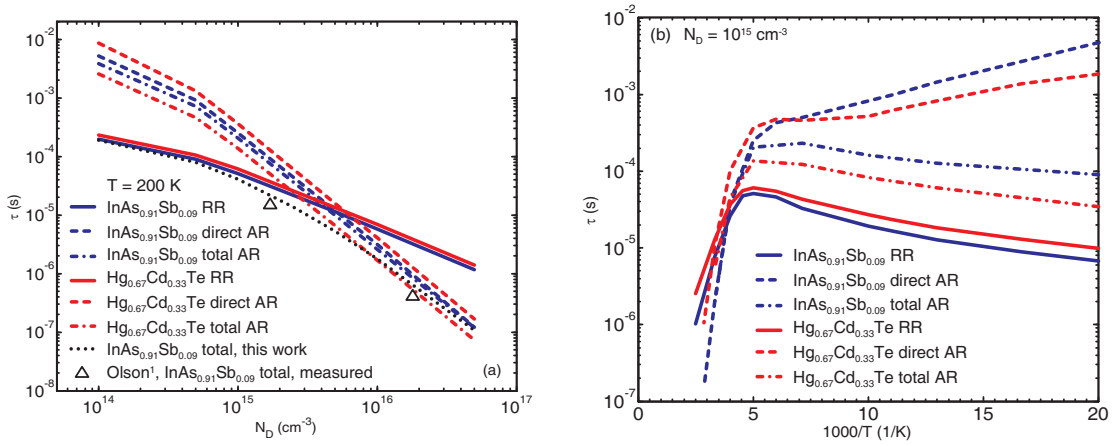


Figure 5·8: Calculated radiative (solid lines), direct (dashed lines) and PA indirect (dash-dot lines) Auger recombination lifetime as a function of (a) n-type doping concentration and (b) lattice temperature for $InAs_{0.91}Sb_{0.09}$ (blue lines) and $Hg_{0.67}Cd_{0.33}Te$ (red lines). A lattice temperature of 200 K is assumed in (a) while a doping of $N_D = 10^{15}cm^{-3}$ is used in (b). The dotted line in (a) is the total lifetime (Auger and radiative lifetime) for $InAs_{0.91}Sb_{0.09}$ at 200 K. For comparison, the open triangles show the experimental values measured by 1: Olson and coworkers(Olson et al., 2013).

concentrations, where more carriers will present in higher energy states, effectively increasing the available direct recombination paths. Due to its larger interaction strength, the direct process, as a result, dominates over the indirect process.

We further compared the Auger recombination lifetime with the radiative lifetime. As is expected, when the doping concentration is higher than $3 \times 10^{15}cm^{-3}$ for $Hg_{0.67}Cd_{0.33}Te$ or $5 \times 10^{15}cm^{-3}$ for $InAs_{0.91}Sb_{0.09}$, the minority carrier lifetime is dominated by the Auger process whereas at lower doping concentration, the radiative recombination is found to be the limiting factor. From Fig. 5·8(a) it can be seen that although the lifetime difference is not very significant, $InAs_{0.91}Sb_{0.09}$ has better intrinsic minority carrier lifetime at high doping concentrations thanks to its smaller total Auger recombination rate. In Fig. 5·8(a), the total lifetime (combined Auger and radiative) calculated for $InAs_{0.91}Sb_{0.09}$ at 200 K is also compared to the experimental

data from Olson and coworkers(Olson et al., 2013) that measured the carrier lifetime in bulk $InAs_{0.91}Sb_{0.09}$ at different temperatures and doping concentrations. It can be seen that the calculated values are in good agreement with the measured data which indicates that the intrinsic recombination mechanisms are indeed the limiting factors in determining the carrier lifetime in the material. Indeed, from Olson's reported data pertinent to unintentionally doped and doped bulk $InAs_{0.91}Sb_{0.09}$, one can evince that SRH recombination, although present, is not dominant. On the contrary, in the case of T2SL SRH recombination is the determining factor of the carrier lifetime below 200 K.

The calculated temperature-dependent hole lifetimes in the two alloys are presented in Fig 5·8(b) where radiative, direct and total Auger processes are compared at a constant n-type doping of 10^{15}cm^{-3}. Although it may seem counter-intuitive, the results show that the contribution from the PA indirect Auger process to the total Auger rate actually becomes larger at lower temperature when the average number of phonon decreased. This phenomenon, however, can be understood by considering the two competing factors in determining the PA process: the strength of the recombination and the final states available for the recombination process. The strength of the PA Auger recombination, which is related to the matrix element for Auger process and band broadening, decreases at lower temperature due to the diminishing phonon population. Indeed, further calculation of the self-energy confirmed that the broadening of the band under the phonon perturbation decreases with the temperature. However, it also should be noted that from Eq. (2.35) the broadening of the bands will converge to a finite value as the temperature approaches zero. In fact, if one sets the phonon population factor $P(\hbar\omega)$ to zero and the electron population factor to 1 or 0 depending on the value of the energy E, the band broadening $\text{Im}\Sigma(\mathbf{k}, E)$ always exists through the emission of optical phonons which, unlike phonon absorption, does not require the

presence of a phonon population. As a result, the PA process will always contribute to the total Auger recombination even at extremely low temperature (Kioupakis et al., 2011).

The other factor that will affect the total PA Auger rate is the number of available final states for Auger recombination. As the temperature decreases, carriers will tend to concentrate at the band edge. This in turn reduces the available direct recombination final states that fulfill both energy and momentum conservation. On the other hand, through the emission/absorption of a phonon, the momentum conservation among the four electron states can be relaxed with the phonon providing the additional momentum. Therefore, the number of available final states for PA Auger process increases. Moreover, in order to have a meaningful comparison between different materials that may have different dopant activation energies, the carrier concentration is kept fixed for all the temperatures investigated (no freeze-out effect is included). This means that, in order to maintain the same amount of electrons and holes in the conduction and valence band, the Fermi level moves closer to the band edges at lower temperature. As a result, the statistical factor for electron $\Theta(E)$ is found to be higher at lower temperature when evaluating at the conduction band edge. Same effect is also true for the hole distribution. Consequently, even though the strength of the PA process diminishes at low temperature, the total PA Auger rate slightly increases when the temperature changes from 200 K to 50 K for MWIR materials. Further calculation at 10 K gives a much lower total Auger rate indicating that the rate of PA process will eventually decrease at temperature below 50 K. It should be noted that even though the indirect Auger dominates over the direct Auger process at low temperature, the minority carrier lifetime in this range is limited by the radiative recombination. Indeed, when the temperature becomes lower than 250 K the radiative recombination becomes stronger than the total Auger recombination in both materials with the

$Hg_{0.67}Cd_{0.33}Te$ having larger overall lifetime. In the high temperature range, due to the fact that more carriers now have enough energy to occupy high energy states, the direct process gradually contributes to all of the Auger recombination events, and eventually exceeds the radiative process and becomes the limiting factor of the hole lifetime in the two materials. It is also worth noting that for all the temperatures considered here $Hg_{0.67}Cd_{0.33}Te$ generally has slightly better intrinsic minority carrier lifetime than $InAs_{0.91}Sb_{0.09}$ because of its larger radiative lifetime at low temperature as well as weaker direct Auger recombination rate at high temperature.

As a final remark, we want to emphasize that the calculations presented in this work do not include the effect of photon recycling (PR), which tends to increase the radiative recombination lifetime extracted from devices level measurements. In fact, earlier work on PR in HgCdTe (Humphreys, 1983; Grein et al., 1997), indicated that the enhancement of the radiative lifetime due to PR can be as high as 25 times which may explain why in some cases the calculated lifetime are lower than the reported values for MWIR HgCdTe (Kinch, 2007). However, as the derivation in Ref. (Humphreys, 1983) suggests, the effect of PR largely depends on the geometry of the devices such as the thickness of the layers and the shape of the device boundaries which can vary significantly from device to device. For these reasons, the aim of the present work is to provide a baseline evaluation of the radiative recombination lifetime in materials under the assumption that PR effect can be ignored. This is obviously the worst case scenario in which the lifetime is not enhanced by the PR effect.

5.2.2 n-type Strained InAsSb

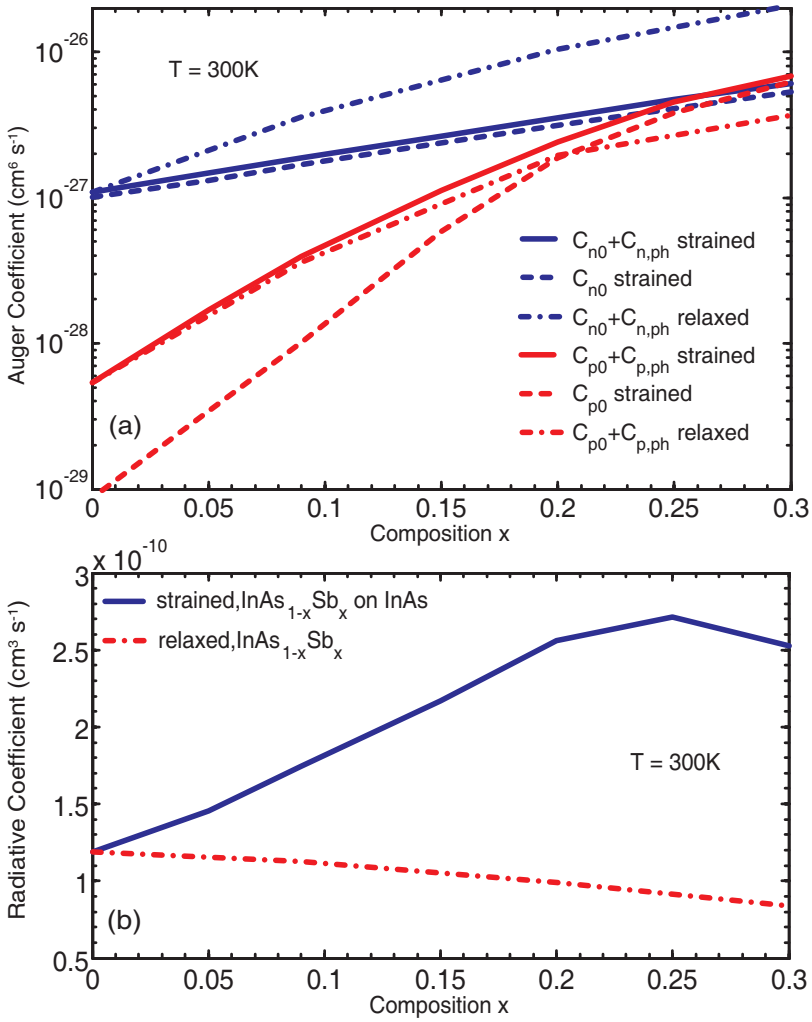


Figure 5·9: (a): Auger coefficients computed in n-type (blue lines) and p-type (red lines) strained materials are shown in solid and dashed lines for total and direct Auger processes respectively. The results from the relaxed material are plotted in dash-dot lines as comparisons. In (b), the radiative coefficients in strained (solid line) and relaxed (dash-dot line) materials are calculated under the same condition as that in (a). For the strain configuration, the $InAs_{1-x}Sb_x$ alloy is assumed to be grown on an InAs substrate. All the calculations assumed lattice temperature of 300 K.

At this point we are ready to discuss the effect of strain. Specifically, we consider the Auger and radiative recombination rates assuming that InAsSb alloys are grown

on an InAs substrate in which the applied compressive strain in the alloys becomes larger as the antimony fraction increases. In Fig. 5·9(a) the direct and the total Auger coefficients in strained InAsSb alloys are plotted as a function of the composition value x for both n and p-type materials. The results obtained for the relaxed alloys, which has been presented in Fig. 5·6 is also included as comparisons.

It can be seen that for the CHCC process, the calculated total Auger coefficients in strained alloys are significantly lower than that in the case of relaxed material. For example, at $x = 0.2$, $C_{n0} + C_{n,ph}$ in strained alloy is only one third of the corresponding value obtained for the relaxed InAsSb. Furthermore, as x increases, this difference becomes even larger. The decrease of the total Auger rate as a function of the increasing strain can be understood by considering the change of band gap energy under the strain. Indeed, the compressive strain usually leads to an increase of the energy gap in zincblend phase materials (Kuo et al., 1985). Consequently, both direct and PA indirect Auger rates decrease (Wen et al., 2015). For the the CHLH process, the compressive strain results in a total Auger coefficient that is slightly larger than the one for the relaxed alloy, which can be explained by the reduced effective hole mass. In fact, under the compressive strain, the LH band moves above the HH band making the effective hole mass near Γ point smaller. It should be noted that the contribution from the PA indirect Auger to the total Auger rates in the strained alloy follows the same trend as what was described for the relaxed alloys, and a large PA indirect Auger rate in p-type InAs-like alloys is also obtained.

It is worthy of note that when comparing the Auger coefficients in the strained and relaxed alloys at the same band gap energy, the obtained coefficients in strained material are still about 40% smaller than that in the relaxed material. Since with the same band gap energy, the reduction of Auger rates due to the gap shrinking has been eliminated, we conclude that other effects such as the change of effective masses

under strain also contribute to the suppression of Auger process when a compressive strain is applied.

Fig. 5·9(b) compares the radiative coefficients in both strained and relaxed InAsSb alloys. As shown, since the band gap energy increases under the compressive strain for each antimony composition, B_n for the strained material becomes larger than the value for the the relaxed material as a consequence of the increased photon density of states. We need to point out that in the present calculation, only the fundamental microscopic radiative recombination is evaluated and summed to get its final rates. No photon recycling effect has been included due to its dependence on the geometry of the specific devices. However, the photon recycling effect has indeed been observed in some experiments (Borrego et al., 2001) which can lead to reduction of B_n values as much as 17% (Connelly et al., 2010). Therefore, we could reasonably expect that in the real devices, the increase of radiative recombination rate due to the compressive strain should be mitigated by the photon recycling effects.

We want to point out that due to the presence of strain in the material, the thickness of the epitaxial layers should be kept below a critical value that depends on the alloy composition. When the InAsSb layer becomes sufficiently thin, other effects like surface/interface states and quantum confinement will change the electronic structures of the material and the corresponding matrix elements. Consequently, to compute the rates of the Auger and radiative processes the electron wavefunctions need to be modified accordingly. (Vinter, 2002)

Finally we consider the effect of temperature and strain together. To investigate the dependence of the two recombination mechanisms on strain at different temperatures, we analyze the intrinsic carrier lifetime of the commonly used $InAs_{0.91}Sb_{0.09}$ alloy assuming an InAs (GaSb) substrate for the strained (relaxed) situation. As shown in Fig. 5·10, the minority carrier lifetime due to Auger and radiative processes is

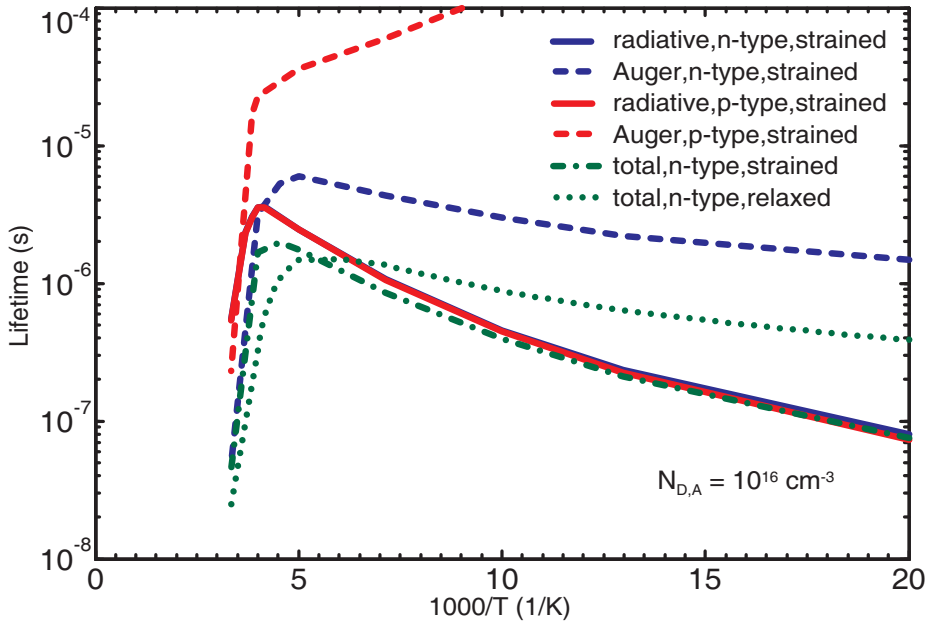


Figure 5·10: Auger (dashed lines) and radiative (solid lines) lifetime in n-type (blue lines) and p-type (red lines) $InAs_{0.91}Sb_{0.09}$ is plotted as functions of inverse lattice temperatures. Despite the dotted line which represents the total intrinsic carrier lifetime in the relaxed material, all other curves are for the compressive strained $InAs_{0.91}Sb_{0.09}$ assuming an InAs substrate. A doping concentration of 10^{16}cm^{-3} is used in the calculation.

plotted as a function of the inverse lattice temperature at the doping concentration of 10^{16}cm^{-3} for both n-type and p-type materials.

Similar to what was found for relaxed $InAs_{0.91}Sb_{0.09}$ (Wen et al., 2015), the dominant recombination process in the strained n-doped material at low temperature ($T < 240\,\text{K}$) is the radiative recombination. In fact, if one compares the total intrinsic minority carrier lifetime in the relaxed layer with the one in the strained layer, the increase of the radiative recombination rate due to strain can be clearly observed. On the contrary, at high temperature, the dominant recombination mechanism becomes the Auger process. As a result of the reduced Auger rate under compressive strain, the strained layer shows a much longer minority carrier lifetime compared to that for the lattice matched situation. In the case of p-type material, the calculated radiative

recombination lifetime is very close to the values computed for the n-doped strained layer. For Auger process, the corresponding recombination rate in this case is found to be very weak. As a final note, we want to point out that the Auger lifetime shown in Fig. 5·10 is the combination of both direct and PA indirect processes.

In conclusion, we have numerically investigated the minority carrier lifetime in $InAs_{1-x}Sb_x$ alloy due to the fundamental radiative and non-radiative processes using a full-band Green's function formalism. Both direct and PA indirect Auger processes as well as the radiative process were studied as functions of antimony molar fractions and lattice temperatures. The calculated results indicated a reduced Auger recombination rate and enhanced radiative rate in the alloys when a compressive strain is applied. Further comparison between the calculated compressively strained and lattice-matched $InAs_{0.91}Sb_{0.09}$ material pointed to a large increase of the minority carrier lifetime under the strained condition when the temperature is above 200 K. This could be potentially useful for the fabrication of improved MWIR detectors if the SRH recombination rate due to the applied strain can be controlled.

5.3 Intrinsic Minority Carrier Lifetime in Strained/Relaxed InGaAs and its Comparison with HgCdTe for SWIR and ESWIR Application

Due to the possibility of exploiting the nightglow phenomena (Battaglia et al., 2007) in the spectral range between 1 - 2 μm, which is of special importance in the field of night vision, surveillance and remote sensing (Onat et al., 2007; Tidhar and Segal, 2011), photon detectors operating in the SWIR band have been the subject of a significant development effort. Over the past decades, many material systems such as Germanium, InGaAs, InSb and HgCdTe have been intensively studied in terms of material properties, growth techniques and device architectures which enable SWIR detectors to be

fabricated and widely used in both civilian and defense applications. Among them, the $In_{0.53}Ga_{0.47}As$ lattice matched to InP substrate with cutoff wavelength of 1.7 μm has emerged as the most successful material due to its high quantum efficiency, low dark current at room temperature and relatively low growth and processing cost (Hansen and Malchow, 2008). However, many other applications, for example the hyperspectral imaging which is essential for object identification, biomedical imaging and geosciences, require photon detectors with cutoff wavelength beyond 1.7 μm. As a result, the study of ESWIR material systems and their photon detector architectures has attracted a significant interest. To extend the cutoff wavelength of InGaAs alloy further into 2 - 3 μm spectral range, an Indium molar fraction around 80% is usually required. The corresponding material lattice constant in this case will be 1.9% larger than that for InP substrate and compressive strains will be introduced into the material (Arslan et al., 2015).

To avoid the resulting high interfacial defects and the associated SRH recombination in the absorption layer due to the excessive lattice mismatch, a graded buffer layer is usually grown between the InGaAs absorption layer and the substrate(Sasaki et al., 2013). The problem of lattice defects in InGaAs, therefore, can be partially mitigated. However, even with a graded buffer layer, the InGaAs absorption layer is still under strain which will change its optical and recombination properties. In contrast, due to the small lattice constant difference in HgTe and CdTe, the HgCdTe alloy with cutoff wavelength from 1.5 μm up to 12 μm is almost free of strain. Consequently it is worth comparing their optical and carrier recombination properties in order to understand the potential advantages of each alloy.

Few works on the intrinsic carrier lifetime in strained/relaxed InGaAs alloys have been published in the literature. Among those limited number of publications, Fuchs *et al.* (Fuchs et al., 1993) measured and fitted the Auger coefficients in strained and relaxed

InGaAs/InGaAsP quantum well and obtained a value of 1.0×10^{-28}cm^6/s independent of the applied strain. Hausser *et al.* (Hausser et al., 1990) measured an Auger coefficient of 3.2×10^{-28}cm^6/s in the intrinsic bulk InGaAs material (lattice-matched to InP), which is higher than a later measurement with a value of 8.1×10^{-29}cm^6/s in n-type material done by Ahrenkiel *et al.* (Ahrenkiel et al., 1998). By measuring the decay of photoluminescence in the bulk material, Sermage *et al.* (Sermage et al., 1985) extracted the Auger coefficient in relaxed InGaAsP alloy as 2.6×10^{-29}cm^6/s, which is believed to be underestimated due to the overestimation of the carrier density in the material (Hausser et al., 1990). The theoretical calculations of the intrinsic carrier lifetime in this material system also see a large variance of the computed coefficient (Bardyszewski and Yevick, 1985; Takeshima, 1982b; Takeshima, 1984a), which indicates the difficulties of obtaining accurate Auger and radiative coefficients in this alloy. Furthermore, all of these theoretical works cited above only considered the relaxed material and the strain effect on the intrinsic carrier lifetime in this alloy remains unclear.

In this section, a numerical study comparing the intrinsic carrier recombination lifetime in n-type ESWIR InGaAs and HgCdTe is performed using Green's function theory incorporating the full band structure. Radiative recombination together with direct and phonon-assisted indirect Auger recombination are investigated at different cutoff wavelengths under room temperature, where the strain effect is taken into account by assuming an InP substrate for the InGaAs alloy. We first present the obtained results for relaxed materials to establish references of the material properties. After that strained materials will be studied and compared to the relaxed situation.

5.3.1 n-type Relaxed $In_{0.53}Ga_{0.47}As$ and $Hg_{0.38}Cd_{0.62}Te$

Radiative and Auger recombination lifetimes in $In_{0.53}Ga_{0.47}As$ and $Hg_{0.38}Cd_{0.62}Te$ which have a cutoff wavelength of 1.6μm at 300 K are presented in Fig. 5·11(a) and

(b). Unlike $Hg_{0.38}Cd_{0.62}Te$ in which the contribution from the PA indirect Auger recombination accounts for about 83% of the total Auger rate at 300 K, the indirect Auger process in $In_{0.53}Ga_{0.47}As$ shown in Fig. 5·11(a) is very weak and the two lines representing its total Auger lifetime and the lifetime due to direct Auger process almost overlap. As opposed to MWIR materials, the dominant recombination process for the two SWIR materials in the doping range from $10^{14}cm^{-3}$ to $5 \times 10^{16}cm^{-3}$ is always the radiative recombination and we found that the calculated radiative recombination lifetime indeed matches very well with the measured lifetime in the case of $In_{0.53}Ga_{0.47}As$ (square symbol) (Ahrenkiel et al., 1998). Furthermore, by extrapolating the calculated curve of $In_{0.53}Ga_{0.47}As$ to higher doping concentrations we find that an intersection between the total Auger lifetime and radiative lifetime occurs around a doping concentration of $5 \times 10^{18}cm^{-3}$, which also agrees with the measurement done by Ahrenkiel *et al.*(Ahrenkiel et al., 1998). We notice that even though the Auger lifetime in the two SWIR materials are similar, due to the large radiative lifetime in $Hg_{0.38}Cd_{0.62}Te$ and the high intersection point between Auger and radiative process, the overall intrinsic minority carrier lifetime in $Hg_{0.38}Cd_{0.62}Te$ outperforms that in the $In_{0.53}Ga_{0.47}As$ by large amount within the investigated doping range.

The temperature dependence of the two recombination processes in SWIR materials is shown in Fig. 5·11(b). As can be seen, the Auger lifetime, including both direct Auger and PA indirect Auger processes, increases dramatically as the temperature decreases which is different from what was found in MWIR and LWIR materials. In fact, as the band gap of the material increases, the available final states in the band structures decrease, leading to a weaker Auger recombination regardless of the assistance from the phonons. It is obvious that from 300 K to 50 K, the dominant recombination process for the two materials is still the radiative process with the

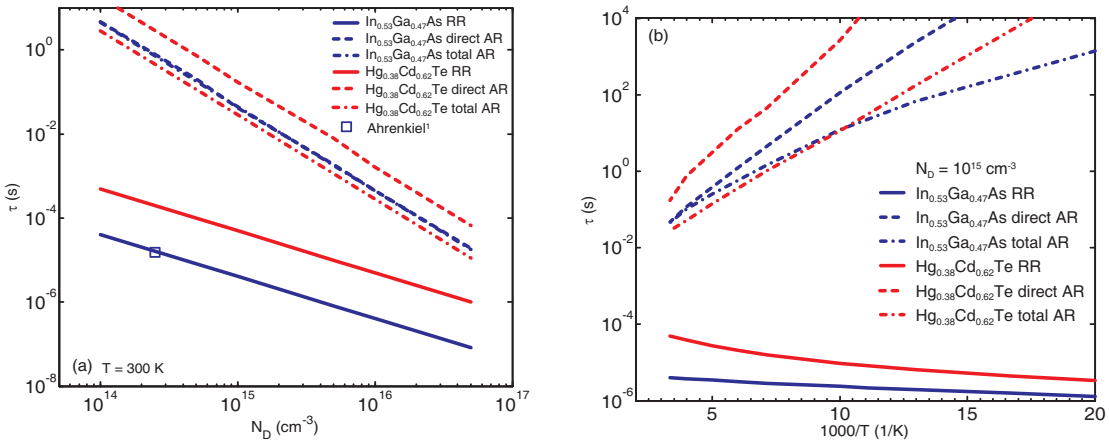


Figure 5·11: Radiative (solid lines), direct (dashed lines) and PA indirect (dash-dot lines) Auger recombination lifetime is computed as a function of (a) n-type doping concentration and (b) lattice temperature for $In_{0.53}Ga_{0.47}As$ (blue lines) and $Hg_{0.38}Cd_{0.62}Te$ (red lines). 300 K is assumed in the calculation of (a) while an n-type doping of $N_D = 10^{15}cm^{-3}$ is used in (b). In (a), the lifetime measured in $In_{0.53}Ga_{0.47}As$ at low-injection from 1: Ref. (Ahrenkiel et al., 1998) is labeled as square.

$Hg_{0.38}Cd_{0.62}Te$ having longer carrier lifetime.

The Auger coefficients calculated from the total Auger-1 lifetime of the two SWIR materials are presented in Fig. 5·12. Although the PA indirect Auger process here also contributes substantially to the total Auger coefficient at low temperatures, the calculated total Auger coefficient $C_{n,ph} + C_{n0}$, in contrast to the MWIR material, drops fast as the temperature decreases. In this figure, the measurement done by Ahrenkiel *et al.*(Ahrenkiel et al., 1998) for $In_{0.53}Ga_{0.47}As$ at 300 K, averaged over n-type and p-type materials, is reported in the form of triangle, whereas an early theoretical calculation using density functional theory (DFT) for direct Auger process is shown as circle symbol(Picozzi, 2002). The square symbol is the calculated Auger coefficient C_p due to CHSH process (Auger-S) from our numerical model. It can be seen that, although the measured $C_{n,p}$ is about three times larger than our numerical results C_n for n-type material, by taking the C_p into consideration our calculated

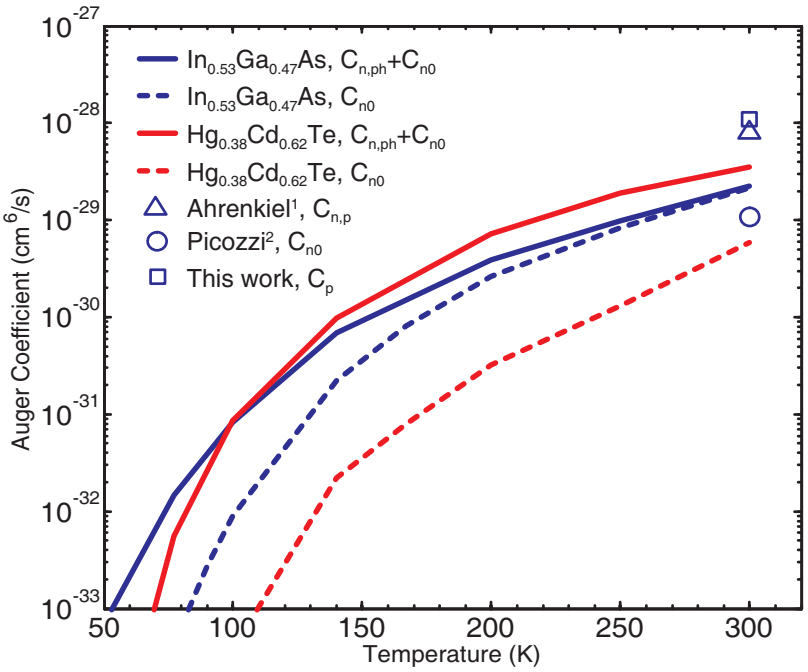


Figure 5·12: Temperature dependence of Auger-1 coefficient for $In_{0.53}Ga_{0.47}As$ (blue lines) and $Hg_{0.38}Cd_{0.62}Te$ (red lines) are compared with 1: fitted $C_{n,p}$ of $In_{0.53}Ga_{0.47}As$ at 300 K from the measured data (Ahrenkiel et al., 1998), and 2: calculated direct Auger coefficient C_{n0} of $In_{0.53}Ga_{0.47}As$ (Picozzi, 2002). The solid lines here represent the total Auger coefficient (PA indirect+direct) and the dashed lines are computed from the direct Auger process only. The Auger coefficient C_p due to CHSH process from this work is labeled as square.

value agrees with Ahrenkiel's measurement within a reasonable error. In fact, using the measured lifetime in the sample with $N_A = 2 \times 10^{19}\text{cm}^{-3}$ from Ref. (Ahrenkiel et al., 1998), we estimate a C_p of $1 \times 10^{-28}\text{cm}^6/\text{s}$ which is very close to the C_p from our calculation, that is $1.08 \times 10^{-28}\text{cm}^6/\text{s}$. A relatively good agreement is also found between the numerical direct Auger coefficient C_{n0} from our model and Picozzi's theoretical calculation. Finally, we want to point out that although the PA indirect Auger process diminished at high temperature in $In_{0.53}Ga_{0.47}As$, it is still important in $Hg_{0.38}Cd_{0.62}Te$. In fact, the direct Auger coefficient at this temperature is obviously lower than the total coefficient which has been reported in both Fig. 5·11(b) and

Fig. 5·12.

5.3.2 n-type Strained InGaAs and Relaxed HgCdTe

In the literature, n-type doping between 10^{15}cm^{-3} (for unintentionally doped sample (Arslan et al., 2015)) and $3 \times 10^{16}\text{cm}^{-3}$ is usually used in the fabricated InGaAs photon detector (Cao et al., 2014; Onat et al., 2007). As a result, a doping concentration of 10^{16}cm^{-3} is assumed for the InGaAs alloy investigated in this section to accommodate the most relevant doping condition in real detectors. Both direct and phonon-assisted indirect Auger-1 (CHCC) recombination and direct radiative recombination are included in the calculation. The carrier lifetime is then computed as a function of cutoff wavelength in the ESWIR spectral region from 1.7 μm to 2.7 μm. Since the energy gap and lattice constant for InAs is 0.352 eV and 6.058 Å respectively compared to 1.418 eV and 5.642 Å for GaAs, when increasing the Indium molar fraction of the ternary material, the cutoff wavelength for InGaAs increases, so does the lattice constant. We consider two cases. The first in is fully relaxed InGaAs material. In the second case, we consider a special situation in which the film is relaxed when grown on a lattice-matched InP substrate so that no strain is induced. Therefore, by varying the Indium composition, with values higher than 53%, we can modify the gap energy and strain configuration. It should be noted that since compressive strain leads to an increase of the material's gap energy, its film cutoff wavelength will be shorter than the relaxed alloy which has the same x. Consequently, in order to compare the strained and relaxed alloys at the same cutoff wavelength, the Indium composition x in the strained case needs to be adjusted accordingly.

Fig. 5·13(a) presents the calculated minority carrier lifetime in both strained and relaxed n-type InGaAs ternary alloys at room temperature under low injection condition. As is shown, for the fully relaxed film, the carrier lifetime for the Auger process decreases from 2×10^{-4} s to about 10^{-5} s when the cutoff wavelength changes

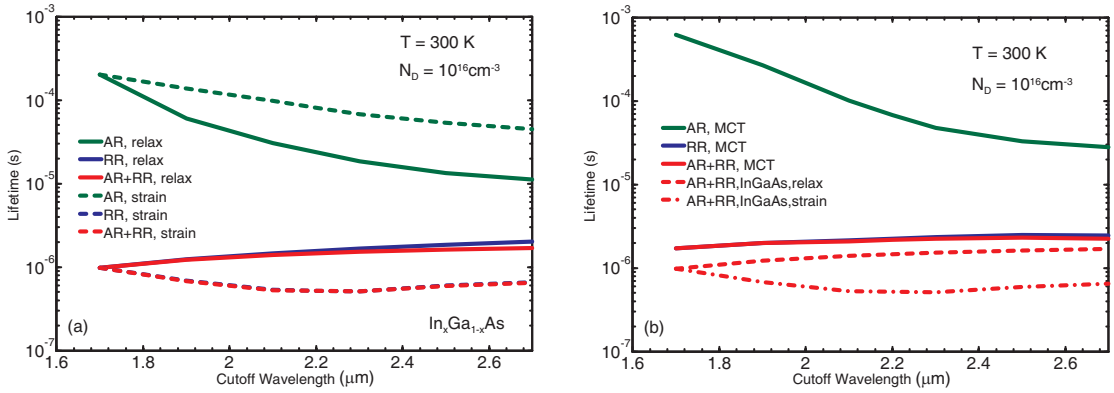


Figure 5·13: (a): Minority carrier lifetime in relaxed (solid lines) and strained (dashed lines) InGaAs is computed in terms of the material cutoff wavelength. The calculation assumed n-type doping concentration of 10^{16}cm^{-3} at 300 K in which both Auger (green lines) and radiative (blue lines) lifetime has been investigated. The total carrier lifetime in (a) is shown in red lines. In (b), same calculation is done for relaxed HgCdTe (solid lines) and the results are compared to the total carrier lifetime in relaxed (red dashed line) and strained (red dashed-dot line) InGaAs.

from 1.7 μm to 2.7 μm. The decrease of Auger lifetime follows the usual trend, in which the Auger rate for low band gap material becomes higher due to stronger coupling between conduction and valence band (Wen and Bellotti, 2015a). In the case of strained InGaAs on InP substrate, the corresponding Auger lifetime (green dashed line), also decreases with the increasing cutoff wavelength. However, when compared to the fully relaxed InGaAs, the strained material shows longer Auger lifetime at longer cutoff wavelengths and the difference between the values for relaxed and strained films grows as larger strain is applied. On the other hand, the radiative recombination, which are represented by blue solid and blue dashed lines for relaxed and strained material respectively, shows the opposite trend. When larger strain is applied, the radiative recombination lifetime for InGaAs first decreases from 10^{-6} s to 5.16×10^{-7} s and then slightly increases to 6.6×10^{-7} s as the cutoff wavelength approaches 2.7 μm, whereas in the case of fully relaxed InGaAs, the radiative lifetime

always increases with the decreasing band gap energy.

Since for both Auger and radiative recombination processes, the carrier lifetime in relaxed and strained InGaAs is compared at the same cutoff wavelength, the difference shown in Fig. 5·13(a) is entirely due to the compressive strain effect. The first consequence of the presence of the strain is the increase the effective mass of electrons and decrease of the effective mass of holes (Kim and Fischetti, 2010). Moreover, the strain-induced band structure change will also affect the overlap integral among the initial and final electronic states making it difficult to predict the aforementioned carrier lifetime change according to the analytical model (Haug, 1983). As a result, a numerical analysis in this case is not only desirable but also necessary. The total intrinsic carrier lifetime combining both Auger and radiative processes for relaxed and strained InGaAs is shown by red solid and red dashed lines in Fig. 5·13(a). It is easy to observe that the dominating recombination process in this situation is the radiative one, and the relaxed InGaAs shows longer minority carrier lifetime compared to the strained cases. We also note that both direct and phonon-assisted indirect Auger processes have been included in the calculation above in which only the total Auger lifetime is plotted for simplicity and brevity.

Figure 5·13(b) presents the same quantities calculated for relaxed HgCdTe with the same doping concentration as the InGaAs at different cutoff wavelengths. As is shown, both Auger and radiative recombination lifetime for HgCdTe follows the same trend as that for InGaAs, with the radiative process still dominating the total intrinsic recombination. Consequently, when compared to the total minority carrier lifetime for relaxed and strained InGaAs, which are reported in Figure 5·13(b) by red dashed and red dashed-dot lines, the carrier lifetime in HgCdTe is consistently longer than the two InGaAs cases at all the investigated cutoff wavelengths due to its longer radiative lifetime. Furthermore, we expect that for a real material, straining

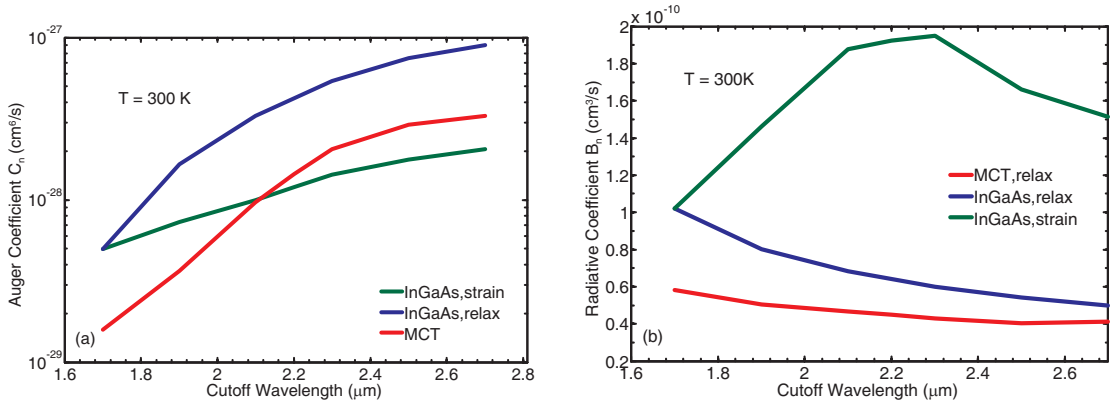


Figure 5·14: (a): Auger and (b): radiative recombination coefficients are calculated for relaxed (blue lines), strained (green lines) InGaAs and relaxed HgCdTe (red lines) at different material cutoff wavelengths. The lattice temperature is assumed to be 300 K. Here, the Auger recombination coefficient is the combination of both direct and phonon-assisted indirect processes.

InGaAs will inevitably introduce lattice defects which would increase the Shockley-Read-Hall recombination rate dramatically and further degraded the carrier lifetime. We speculate that the resulting material performances for strained InGaAs therefore, in addition to its shorter intrinsic carrier lifetime, would be worse than that in HgCdTe alloy which enjoys almost strain-free condition from ESWIR to MWIR spectral range.

The Auger and radiative recombination coefficients for InGaAs and HgCdTe materials at different cutoff wavelengths are presented separately in Fig. 5·14(a) and (b). In Fig. 5·14(a), the calculated Auger coefficients increase with the cutoff wavelength of the material in all three cases (relaxed and strained InGaAs, and relaxed HgCdTe) and it is easy to see that the compressive strain in InGaAs alloy greatly reduced its Auger recombination rate compared to that in the relaxed material. When considering the HgCdTe alloy, at the cutoff wavelength below 2.1 μm the HgCdTe alloy has the lowest Auger coefficient giving it longest Auger lifetime in this spectral range. For the alloy compositions with cutoff wavelength longer than 2.1 μm, carriers in the strained InGaAs showed weaker Auger recombination than that in relaxed HgCdTe

due to its slow increase of Auger coefficient in this situation. However, even though at longer cutoff wavelength the Auger coefficient in HgCdTe is higher than that in the strained InGaAs/InP structure, since in the ESWIR range the dominating intrinsic recombination is radiative process (Fig. 5·13(b)), the overall material performance of HgCdTe is still better than the strained InGaAs. In fact, as it is shown in Fig. 5·14(b), for all the cutoff wavelengths we considered in this work, HgCdTe always has the lowest radiative recombination coefficient B_n among the three materials studied. It is also worthy of note that compressive strain in InGaAs significantly increases the radiative recombination coefficient. For example, at the cutoff wavelength of 2.2 μm, the corresponding B_n for strained InGaAs is more than 3 times larger than that in the relaxed InGaAs. Though B_n for compressively strained InGaAs drops when the material cutoff wavelength increases above 2.2 μm, at 2.7 μm its corresponding B_n is still the largest among the three material configuration, making its minority carrier lifetime shorter compared to the latter two situations.

As a final remark, we notice that our calculated radiative recombination coefficient $B_n = 1.02 \times 10^{-10}\text{cm}^{-3}\text{s}^{-1}$ for relaxed $In_{0.53}Ga_{0.47}As$ is comparable with the measured value by Ahrenkiel *et al.* (Ahrenkiel et al., 1998) who obtained $1.43 \times 10^{-10}\text{cm}^{-3}\text{s}^{-1}$ from the fitting of radio-frequency photoconductive decay data.

5.4 Overlap Integral Factors for MWIR and SWIR Materials

In order to facilitate the use of BLB formula in calculating the Auger coefficients, we put the calculated overall Auger-1 coefficient back into Eq. (2.32) and evaluated the corresponding overlap integral factor $|F_1F_2|$ which is plotted in Fig. 5·15. As can be noticed, for the MWIR materials the values of $|F_1F_2|$ for $InAs_{0.91}Sb_{0.09}$ and $Hg_{0.67}Cd_{0.33}Te$ do not change significantly in the temperature range between 100 K to 300 K in which their values are 0.07-0.085 and 0.1-0.125, respectively. As to the SWIR

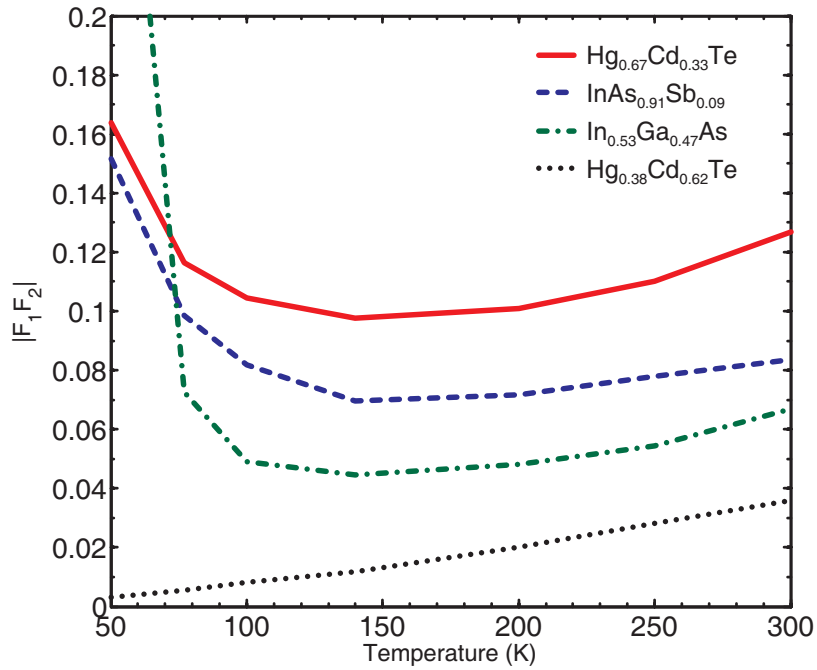


Figure 5·15: Overlap integral factors $|F_1F_2|$ for $Hg_{0.67}Cd_{0.33}Te$ (solid line), $InAs_{0.91}Sb_{0.09}$ (dashed line), $In_{0.53}Ga_{0.47}As$ (dash-dot line) and $Hg_{0.38}Cd_{0.62}Te$ (dotted line) are calculated as a function of temperature using the BLB formula in Eq. (2.32) and the numerical Auger-1 coefficient $C_{n,ph} + C_{n0}$ from this work.

materials, due to the much weaker Auger recombination at low temperature, the calculated factors $|F_1F_2|$ varies significantly between 100 K to 300 K, with their values spanning the range 0.04-0.07 for $In_{0.53}Ga_{0.47}As$ and 0.005-0.035 for $Hg_{0.38}Cd_{0.62}Te$.

In the low temperature regime, when $T < 100$ K, $|F_1F_2|$ for all the four materials changes significantly with the temperature and we speculated that it is probably caused by the assumption of non-degenerate material approximation used in the BLB formula. In fact, the Boltzmann distribution can only approximate the Fermi-Dirac distribution in the limit of non-degenerate situation which is reasonable at high temperature. In the low temperature range, the distribution of carriers is governed by the Pauli exclusion principle and the use of Boltzmann statistic would lead to the an un-physical increase of the overlap factor as is the case in Fig. 5·15. Although the BLB

formula becomes inaccurate at low temperature, it is still a quick way to estimate the Auger coefficient and lifetime in the context of device simulation as long as the $|F_1F_2|$ factor is selected carefully. However, instead of choosing the overlap factor between 0.1-0.3 as a common practice, for example in $Hg_{0.78}Cd_{0.22}Te$ a $|F_1F_2| = 0.16$ gives good results from BLB formula, we found that for MWIR material a proper $|F_1F_2|$ usually lies around 0.1 whereas for SWIR material, $|F_1F_2|$ is always below 0.1 and varies drastically with temperature. Indeed, as the band gap of material increases, the interaction between the electronic states that are involved in the Auger recombination process becomes weaker, which in turn leads to a smaller $|F_1F_2|$.

To conclude, we want to point out that in the original derivation of BLB formula (Beattie and Landsberg, 1959), only the pure collision mechanism is taken into account. However, in our calculation both pure collision and PA Auger process are considered. As a result, the overlap factor values that we derive are larger than the ones computed by Krishnamurthy *et al.*(Krishnamurthy et al., 2006) who followed the same calculation procedure as that proposed by Beattie and Landsberg.

5.5 Conclusions of the Chapter

Infrared detectors have received great attention from the research community in recent years due to its important application in the night vision, security surveillance, object tracking and identification. As the material growth and device fabrication technologies become mature, the degradation of carrier lifetime due to the defects in the crystal has diminished to a point where the intrinsic recombination mechanisms starts to become the limiting factor. In this chapter, the minority carrier lifetime due to both Auger and radiative recombination is studied using Green's function theory and the full band structure presented in Chapter 3.

We have calculated and compared the doping-dependent and temperature-

dependent radiative and Auger recombination lifetime, as well as the corresponding Auger coefficients, in the relaxed and strained HgCdTe, InAsSb and InGaAs intended for operation in the MWIR, ESWIR and SWIR spectral bands. Using the Green's function theory and EPM full band structures, both direct and PA indirect Auger processes are included in the numerical calculation which is implemented through the Monte-Carlo integration method. Our results indicate that for $InAs_{0.91}Sb_{0.09}$ ($Hg_{0.67}Cd_{0.33}Te$), when the n-doping concentration is greater than $5 \times 10^{15}\text{cm}^{-3}$ ($3 \times 10^{15}\text{cm}^{-3}$) the Auger process dominates over the radiative process whereas, at $T < 250\,\text{K}$, the radiative process becomes the limiting factor for the intrinsic minority carrier lifetime in both materials. Comparing the two materials for the MWIR band, we find that $Hg_{0.67}Cd_{0.33}Te$ generally has slightly better intrinsic minority carrier lifetime over $InAs_{0.91}Sb_{0.09}$ in the temperature range of 50 K to 300 K due to its larger radiative lifetime at low temperature as well as weaker direct Auger recombination rate at high temperature.

For the ESWIR material, the computed minority carrier lifetime values indicate that for $In_xGa_{1-x}As$ ($x > 0.53$) grown on InP substrate, the presence of compressive strain leads to a decrease of the Auger recombination rate and increase of radiative recombination rate. Since the dominant intrinsic recombination mechanism in this spectral range is radiative recombination, the overall intrinsic carrier lifetime in strained InGaAs alloys is shorter than that in the case of fully relaxed InGaAs. When compared to the relaxed HgCdTe, both relaxed and compressively strained InGaAs alloys showed shorter intrinsic carrier lifetime at the same cutoff wavelength at room temperature which confirms the advantage of HgCdTe as wide-band infrared detector material. However, considering the difficulties of growing HgCdTe alloys with high Mercury molar fraction, and the fact that strained InGaAs alloy may be significantly defective it is not clear which one of the two alloys can be effectively used for ESWIR

applications. While HgCdTe is very well established and mature, InGaAs may still be an alternative solution for the ESWIR detector as long as the material defects due to lattice mismatch can be controlled to avoid further degradation of the minority carrier lifetime.

For SWIR materials, namely $In_{0.53}Ga_{0.47}As$ and $Hg_{0.38}Cd_{0.62}Te$, our calculations indicate that over the n-type doping from $10^{14}cm^{-3}$ to $5 \times 10^{16}cm^{-3}$ and temperature between 50 K to 300 K, the radiative recombination process always dominates over the Auger process which is more than three orders of magnitude weaker than that in MWIR materials. Moreover, the results also suggest that $Hg_{0.38}Cd_{0.62}Te$ has a better intrinsic minority carrier lifetime compared with $In_{0.53}Ga_{0.47}As$ due to its lower radiative recombination rate.

Furthermore, the estimation of the overlap integral factor $|F_1F_2|$, which is evaluated by matching the BLB formula to our numerical results, shows that for MWIR materials, $|F_1F_2|$ usually lies around 0.1 and does not change much within the temperature between 100 K to 300 K. Specifically, for $InAs_{0.91}Sb_{0.09}$ and $Hg_{0.67}Cd_{0.33}Te$ $|F_1F_2|$ is found to be 0.07-0.085 and 0.1-0.125 respectively. In the case of SWIR materials, $|F_1F_2|$ is always below 0.1 and varies significantly with temperature.

Overall, our calculations unveil the importance of Auger and radiative processes in determining the total intrinsic minority carrier lifetime in MWIR, ESWIR and SWIR materials, which will further benefit the simulation of dark current, internal quantum efficiency and detectivity in the design of IR detectors.

Chapter 6

Simulation of Electrical Characteristics of Bulk-like and Wire-like Silicon Light Emitting Diodes

CMOS based light sources have been shown to exhibit illumination levels that can be utilized in micro-displays (Chen et al., 2005; du Plessis et al., 2002). In these micro-display applications, CMOS based avalanche electroluminescent micro-display devices has distinct advantage compared to the well-established LCD and OLED technologies due to its wide operational temperature range (Venter and du Plessis, 2014). However, since all these CMOS micro-display arrays work in high reverse bias condition (reach-through or punch-through) where avalanche breakdown contributes to the electroluminescence, carrier distribution and electric field in the device especially at the *pn* junction need to be understood and optimized in order to increase the luminescence efficiency and prevent device failure.

This chapter presents the results obtained in the study of the silicon-based nanowire structures, which are currently used as light emitting devices in the micro-display arrays (du Plessis et al., 2015). Both carrier concentrations and electric field distribution have been simulated in prototype and realistic silicon nanowire LEDs. The ionization rates in the nanowire are also computed at different temperatures in order to investigate the effect of temperature on the spectral characteristics. Both classical drift-diffusion model and multiple quantum well model have been applied to the silicon nanowires by using the simulation software Synopsys Sentaurus Device and Crosslight

APSYS respectively. The results did not show much difference between classical model and quantum model, indicating that the carrier confinement in the cross sections of nanowires is minimal compared to the avalanche generation along the nanowire direction.

In this chapter, we first consider the bulk-type devices that are fabricated in CMOS compatible processes and subsequently a smaller wire-like devices of much smaller dimensions. For the bulk-type devices, a one-dimensional model is employed in the simulation in which the device is assumed to be operating at the nominal current density. As to the realistic silicon nanowire, device I-V characteristics and electrical field are calculated in order to evaluate the non-equilibrium carrier distribution in the nanowire which further determines the impact ionization rate at the *pn* junction. We noted that in this work, the impact ionization is treated as a local function of the electric field which clearly limits our ability to gain further insight into the microscopic behavior of the device. Nevertheless, since the model is relatively simple, it is convenient to use it to analyze large devices and complex geometries.

6.1 Prototype $N^+P^-N^+$ Silicon Wire

6.1.1 Geometry and Doping Profile

We first investigated the case of a prototype $N^+P^-N^+$ silicon wire with cross section of 10 nm square and length of 580 nm. The I-V characteristics of the device is computed using both APSYS and Synopsys software packages. As is shown in Figure 6·1, the wire is uniformly doped with $10^{19}\,cm^{-3}$ donors, $10^{16}\,cm^{-3}$ acceptor and $10^{19}\,cm^{-3}$ donor in three regions N^+, P^-, N^+ along the z direction. Each region has length of 40 nm, 500 nm and 40 nm, respectively. In the simulation model, two perfect ohmic contacts are placed at the coordinates $z = 0$ and $z = 580$ nm planes. The current flows in the negative z direction. Unless explicitly noted, we assume a lattice temperature

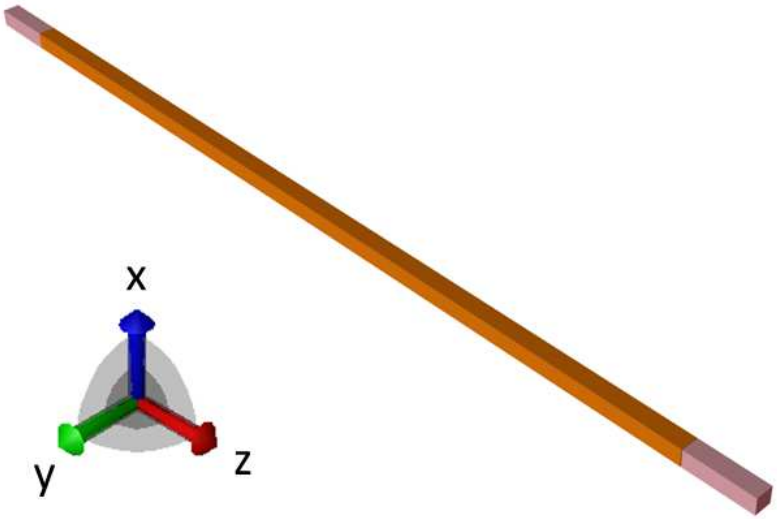


Figure 6·1: Geometrical model of the $N^+P^-N^+$ silicon nanowire.

of 300 K. The relevant information on the band structures and all the other material parameters used in the simulators are obtained from the values demonstrated in Chapter 3. The only exception is the model used for the impact ionization coefficients. In APSYS, Chynoweth (Chynoweth, 1958) impact ionization model is used, while in the Synopsys, the New University of Bologna model (UniBo2) (Gnani et al., 2002; Reggiani et al., 2004b; Reggiani et al., 2004a) is employed instead.

6.1.2 Electrical Properties

Figure 6·2 shows the calculated I-V characteristic of the silicon wire using (a) APSYS and (b) Synopsys. Current control is applied when the voltage gets above 3 V for both simulators. It can be seen that the results from both simulators show a region of negative differential resistance when the applied voltage reaches the critical point, defined as voltage and current values of the first peak in the I-V characteristic, which is about 5 V, 500 μA/μm^2 for APSYS and 6.7 V, 1000 μA/μm^2 for Synopsys. For all these simulations, the biasing is changed from constant voltage to constant current bias, before the critical point is reached.

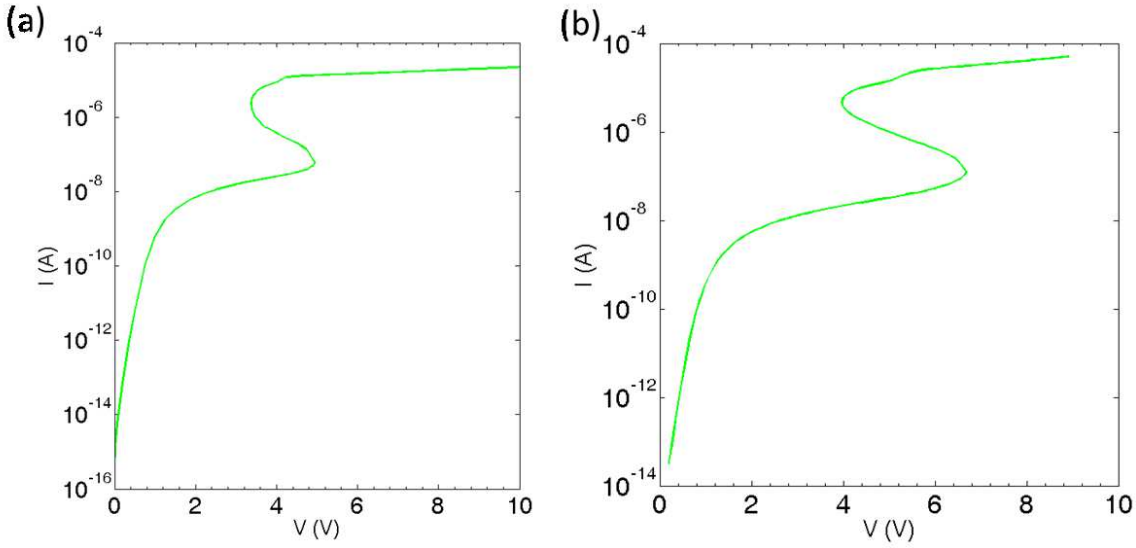


Figure 6·2: Calculated I-V characteristic of the $N^+P^-N^+$ silicon nanowire. (a) Result from APSYS with critical point of 5 V, 500 μA/μm^2 and (b) 6.7 V, 1000 μA/μm^2 for Synopsys.

This behavior can be understood by looking at the effect of driving the device at a constant current (current control) as opposed to the normal operation in which the device is subjected to a constant voltage at its contacts. At moderate voltages the avalanche impact ionization occur in the depleted region of the nanowire which acts as a usual reverse-biased diode. For this bias condition, the generated carriers only give a minimal contribution to the total current which is set by the external current generator. At high enough applied voltage, impact ionization will eventually occur in the P^- depleted region. When the bias reaches the critical point, avalanche impact ionization process leads to super-exponential increase of the device current. By applying the current control, the magnitude of the current itself is fixed and cannot change. Consequently, in order to maintain such a value of current, the ionization rate must drop requiring a reduction of the electric filed in the device. This results in a corresponding drop of the voltage at the ohmic contacts. This process produces a negative differential resistance region, that can be observed for currents between

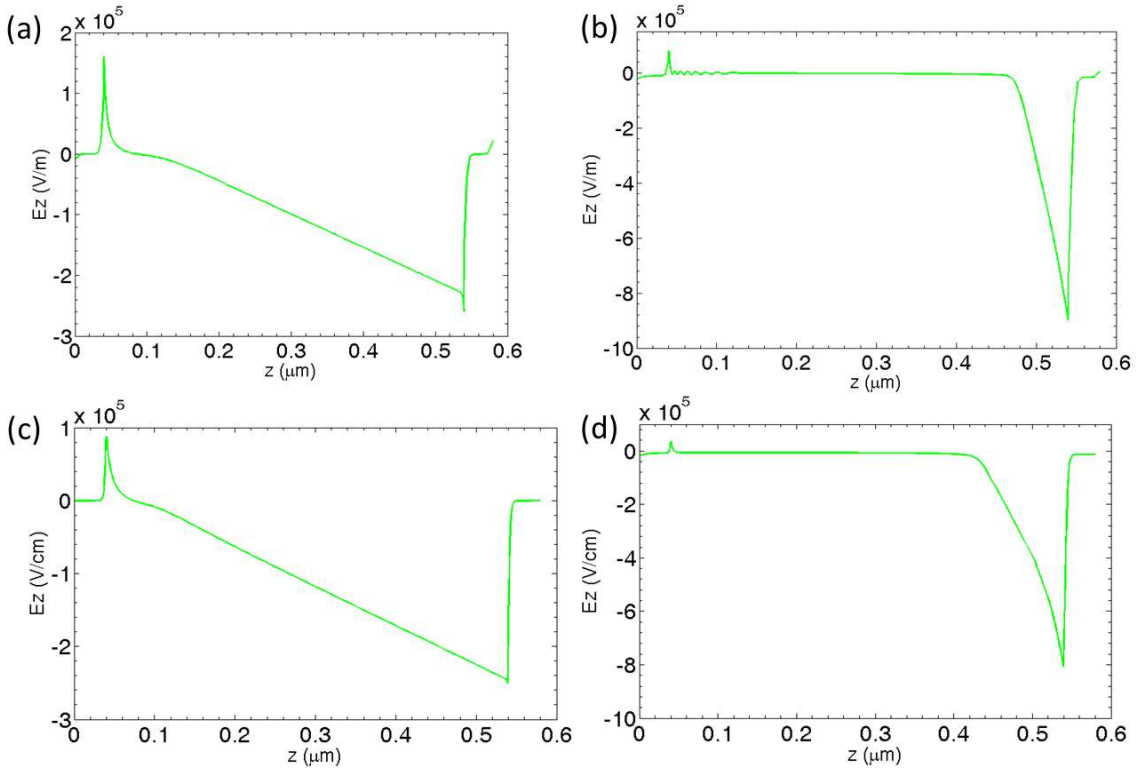


Figure 6·3: Calculated electric field along the z direction at the center of wire. (a) APSYS result at current density of 418 μA/μm^2; (b) APSYS results at current density of 0.0205 A/μm^2; (c) Synopsys result at current density of 418 μA/μm^2;(d) Synopsys result at current density of 0.0205 A/μm^2.

0.1 μA and 1 μA. If the value of the current injected in the device is further increased, the impact ionization process needs to provide the additional amount of electron-hole pairs. However, in the situation where the electric field is screened by the electrons and holes plasma, it is more and more difficult to further increase the field by increasing the contact voltages. As a result, from this point on, the the nanowire behaves as a normal breakdown reverse-biased diode.

In order to investigate the properties of silicon nanowire in detail, we calculated the component of electric field, the carrier distribution, the electron/hole ionization coefficients and its generation rate along the z direction. These results are shown in Figure 6·3, Figure 6·4, Figure 6·5 and Figure 6·6 respectively.

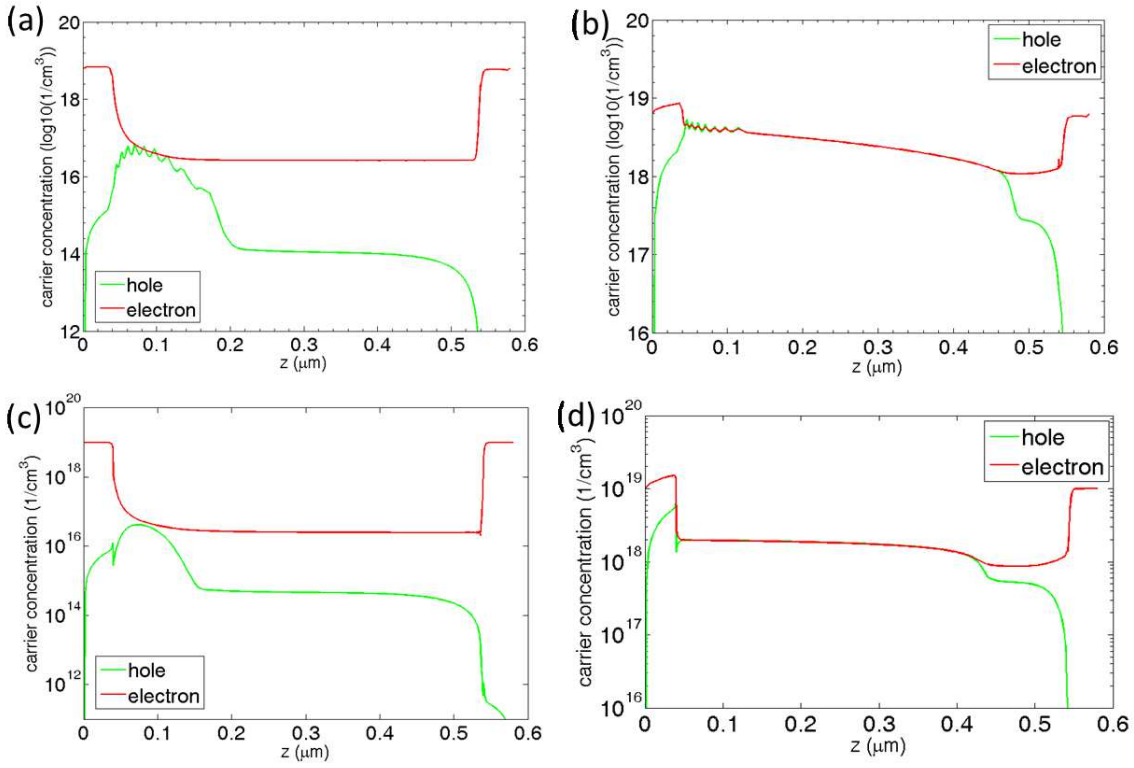


Figure 6·4: Calculated carrier distribution along the z direction at the center of wire. The red line is for electrons, green line is for holes. (a) APSYS result at current density of 418 μA/μm^2; (b) APSYS results at current density of 0.0205 A/μm^2; (c) Synopsys result at current density of 418 μA/μm^2;(d) Synopsys result at current density of 0.0205 A/μm^2.

Figure 6·3(a) and Figure 6·3(c) present the electric field profiles, when the current flowing in the device is 418 μA/μm^2 and below the critical value. Furthermore, Figure 6·3(a) and Figure 6·3(c) also show that the electric field decreases linearly in the P^- region, indicating that this region is depleted of holes and flooded with electrons injected from the forward-biased junction. The mobile charge, shown in Figure 6·4(a) and Figure 6·4(c), together with the fixed ionized acceptors result in electric field with linear profile. We can also notice that the hole concentration is two orders of magnitude lower than the original doping. A significantly different situation occurs when the wire is drive in current above the critical point. Figure 6·3(b) and Figure 6·3(d) present

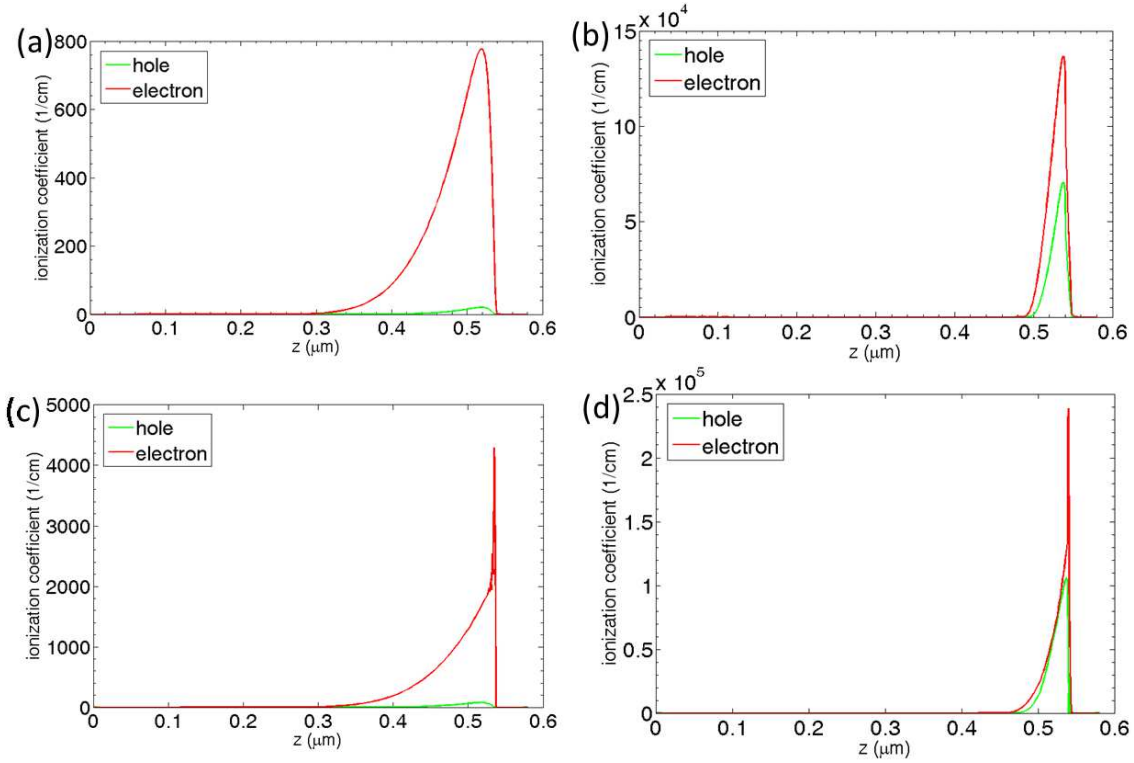


Figure 6·5: Calculated ionization coefficient values for electrons and holes along the z direction at the center of wire. The red line is for electrons, green line is for holes. (a) APSYS result at current density of $418\,\mu\mathrm{A}/\mu\mathrm{m}^2$; (b) APSYS results at current density of $0.0205\,\mathrm{A}/\mu\mathrm{m}^2$; (c) Synopsys result at current density of $418\,\mu\mathrm{A}/\mu\mathrm{m}^2$;(d) Synopsys result at current density of $0.0205\,\mathrm{A}/\mu\mathrm{m}^2$.

the electric filed profile when the current in the wire is $0.0205\,\mathrm{A}/\mu\mathrm{m}^2$, which is above the critical point. The electric field in the P^- region is almost zero except near the reverse-biased junction where we can observe the usual triangular profile. Due to the very large avalanche impact ionization occurring near this junction, a large number of avalanche carriers flood the P^- region, making the whole region almost neutral, as it can be seen in Figure 6·4(b) and Figure 6·4(d). It also important to point out that in this situation the density of injected and generated carriers is much larger than that in the equilibrium condition for both minority and majority carriers. It is clear that in the case of high-level injection the conventional depletion approximation can not

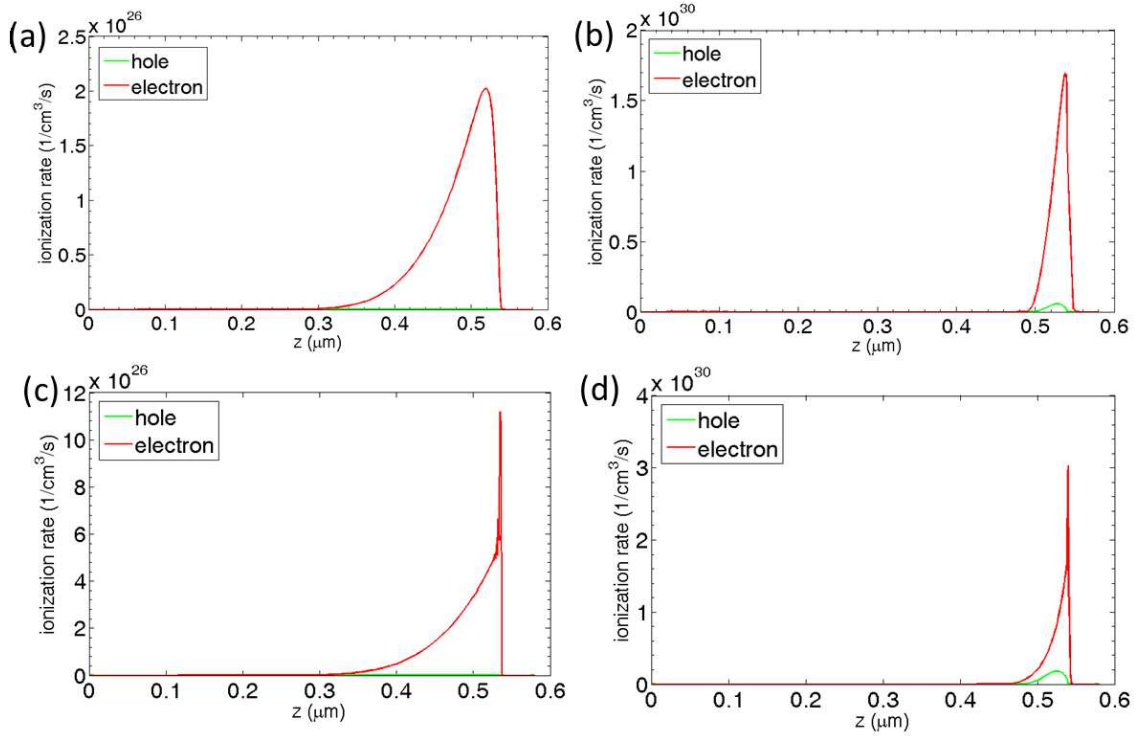


Figure 6·6: Calculated impact ionization generation rates for electrons and holes along the z direction at the center of wire. The red line is for electrons, green line is for holes. (a) APSYS result at current density of $418\,\mu\mathrm{A}/\mu\mathrm{m}^2$; (b) APSYS results at current density of $0.0205\,\mathrm{A}/\mu\mathrm{m}^2$; (c) Synopsys result at current density of $418\,\mu\mathrm{A}/\mu\mathrm{m}^2$;(d) Synopsys result at current density of $0.0205\,\mathrm{A}/\mu\mathrm{m}^2$.

be used to analyze device operation.

Figure 6·5(a) and Figure 6·5(c) shows the calculated ionization coefficients at the same condition under which Figure 6·3 and Figure 6·4 were obtained. As is expected, the impact ionization process occurs in a small portion of the P^- region closed to the reverse-biased pn junction. Same situation is also shown in Figure 6·6(a-d) which demonstrated the calculated impact ionization rates for electrons and holes below/above the critical point using APSYS and Synopsys. It is worthy of note that as for the computed impact ionization coefficients and generation rates, there exist significant differences between APSYS (panels (a) and (b)) and Synopsys (panels (c)

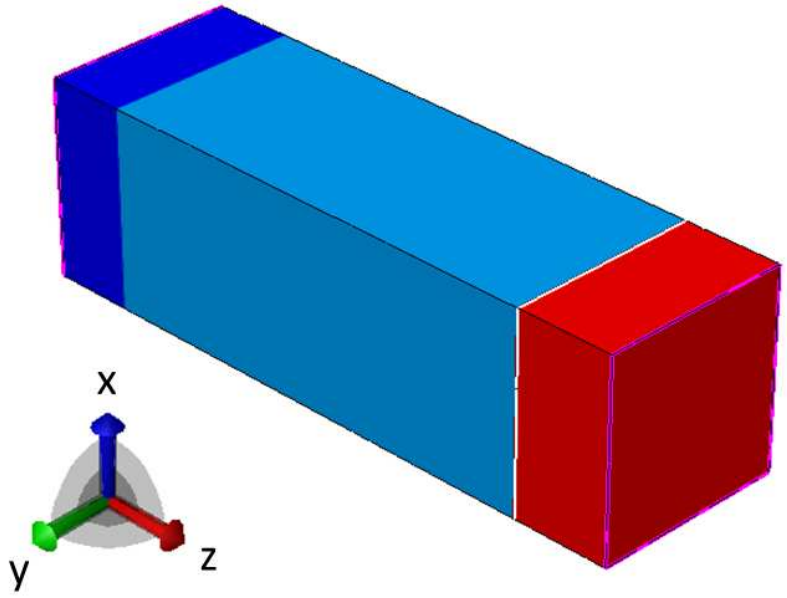


Figure 6·7: Geometrical model of the $P^+P^-N^+$ silicon wire.

and (d)) model. Compared to the results from APSYS, the relatively high impact ionization coefficient for holes from Synopsys is due to the model employed to describe the ionization process. In general, the UniBo2 model is more precise under high electric field in a wide range of temperature while the Chynoweth model is accurate only in the low field and room temperature. Consequently, in this situation results from Synopsys is more reliable.

6.2 Prototype $P^+P^-N^+$ Silicon Wire

6.2.1 Geometry and Doping Profile

The second type of device we consider is a $P^+P^-N^+$ silicon wire with a cross section of $1\,\mu$m square and length of $3.3\,\mu$m. Similar to the previous structure, the doping profile is $10^{19}\,\mathrm{cm}^{-3}$ acceptor, $10^{17}\,\mathrm{cm}^{-3}$ acceptor and $10^{19}\,\mathrm{cm}^{-3}$ donor along z direction, as is shown in Figure 6·7. Ohmic contacts are placed at the two ends of wires ($z = 0$ and $z = 3.3\,\mu$m) so that the current flows is in the -z direction. Both of the simulators,

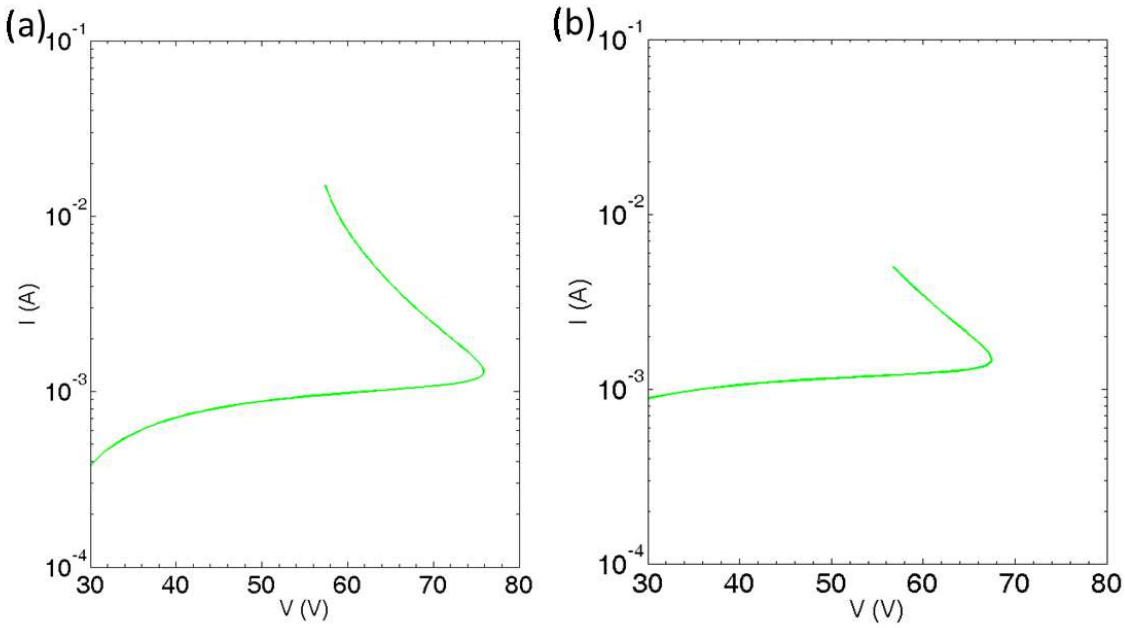


Figure 6·8: Calculated I-V characteristic of the wire. (a) result obtained using APSYS with critical point of 76 V and 1310 μA/μm^2; (b) result obtained using Synopsys with critical point 67.5 V, and 1450,μA/μm^2.

APSYS and Synpsys are used with the same set of material parameters as it was done in the previous section.

6.2.2 Electrical Properties under High Field

Beside the same simulation performed in the previous section, for this particular device, we further consider the temperature dependence of the calculated quantities. In particular, the carrier distribution in the multiplication region is calculated and plotted. We first evaluated the I-V characteristic using the current control bias approach to obtain the breakdown critical point for this device.

Figure 6·8 shows the results obtained using (a) APSYS and (b) Synopsys at 300 K lattice temperature. The corresponding curves also exhibit the negative differential resistance behavior after the critical point which is 76 V, 1310 μA/μm^2 for (a) APSYS

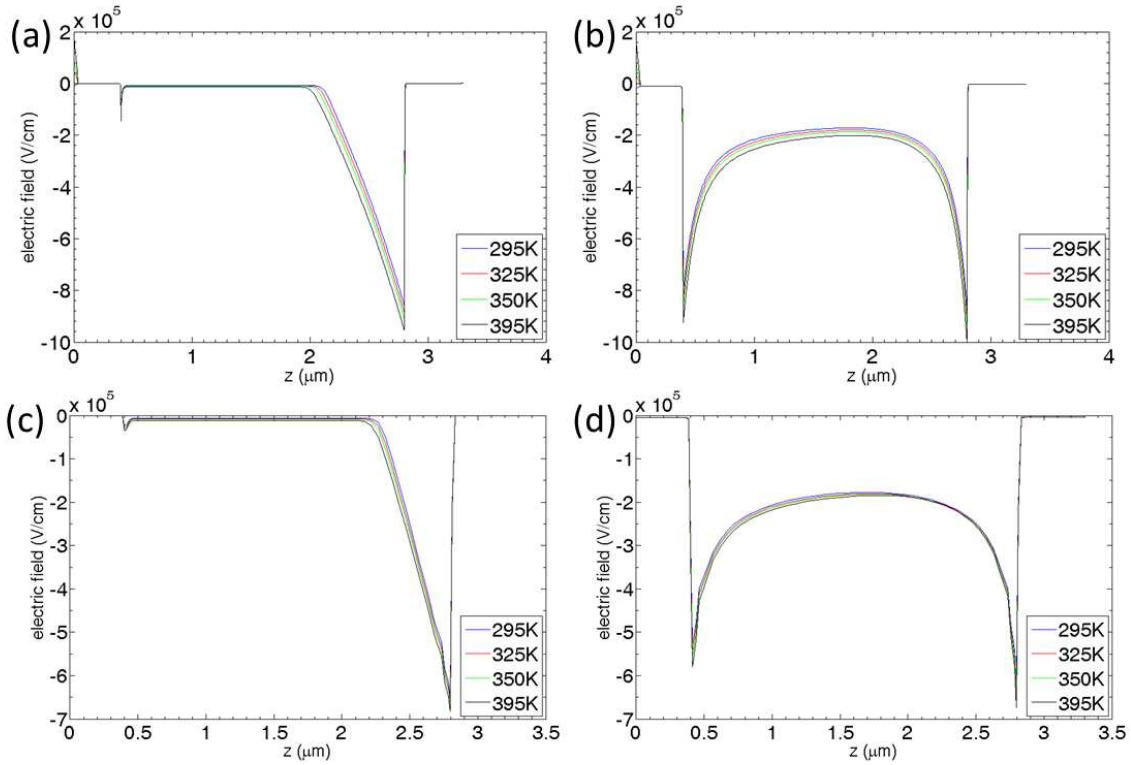


Figure 6·9: Calculated electric field profile along the z direction as a function of the temperature, sampled at the center of wire. (a) APSYS at current density of 300 μA/μm^2; (b) APSYS at current density of 5000 μA/μm^2; (c) Synopsys at current density of 300 μA/μm^2; (d) Synopsys at current density of 5000 μA/μm^2.

and 67.5 V, 1450,μA/μm^2 for (b) Synopsys. The much higher breakdown voltage for this wire, compared to the $N^+P^-N^+$ structure in the previous section, is due to the replacement of the N^+ layer with P^+ layer, and the fact that the wire is much longer (3.3 μm instead of 580 nm). Also, instead of having two back-to-back pn junctions in the previous case, the current device has only one.

Electric field, carrier distribution, impact ionization coefficient and their respective rates for a number of lattice temperatures equal to 295 K, 325 K, 350 K and 395 K have been simulated and plotted in Figure 6·9-6·12. Figure 6·9(a) and (c) present the calculated electric field profile for a current density of 300 μA/μm^2 which is below the critical point. It can be seen that the electric field increases slightly with temperature.

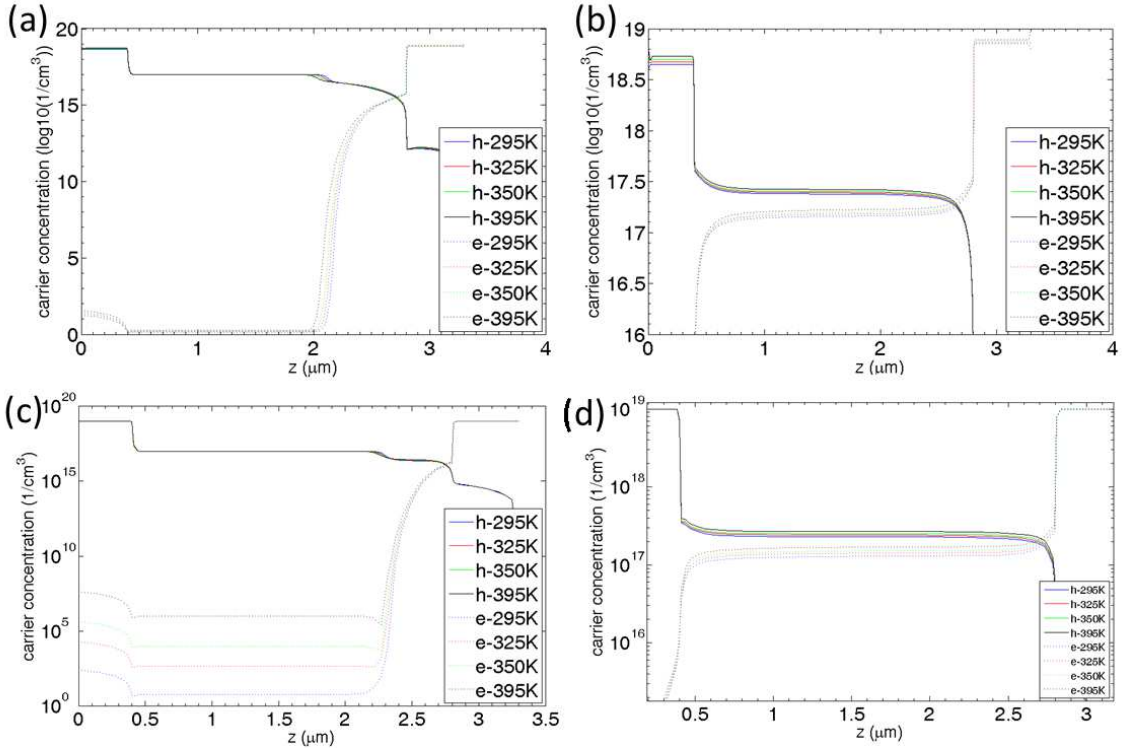


Figure 6·10: Calculated carrier distribution along the z direction as a function of the temperature, sampled at the center of wire. (a) APSYS at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (b) APSYS at current density of $5000\,\mu\text{A}/\mu\text{m}^2$; (c) Synopsys at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (d) Synopsys at current density of $5000\,\mu\text{A}/\mu\text{m}^2$.

In fact, since the ionization coefficient drops as temperature increases, in order to maintain a constant current set by the external generator, the electric filed needs to increase so does the voltage at the device contacts.

Similar to the previous device structure, the electric field changes substantially when the device is operating under high injection condition. Figure 6·9(b) and Figure 6·9(d) present the calculated electric field profile when the current is set above the critical point to $5000\,\mu\text{A}/\mu\text{m}^2$. The electric field in this situation is almost a constant, approximately $1.6 \times 10^5\,\text{V/cm}$, in P^- region, and is a direct consequence of the large impact ionization, which overwhelms the equilibrium carrier concentration. Figure 6·10(c) and Figure 6·10(d) underscore this situation. In fact, almost an equal

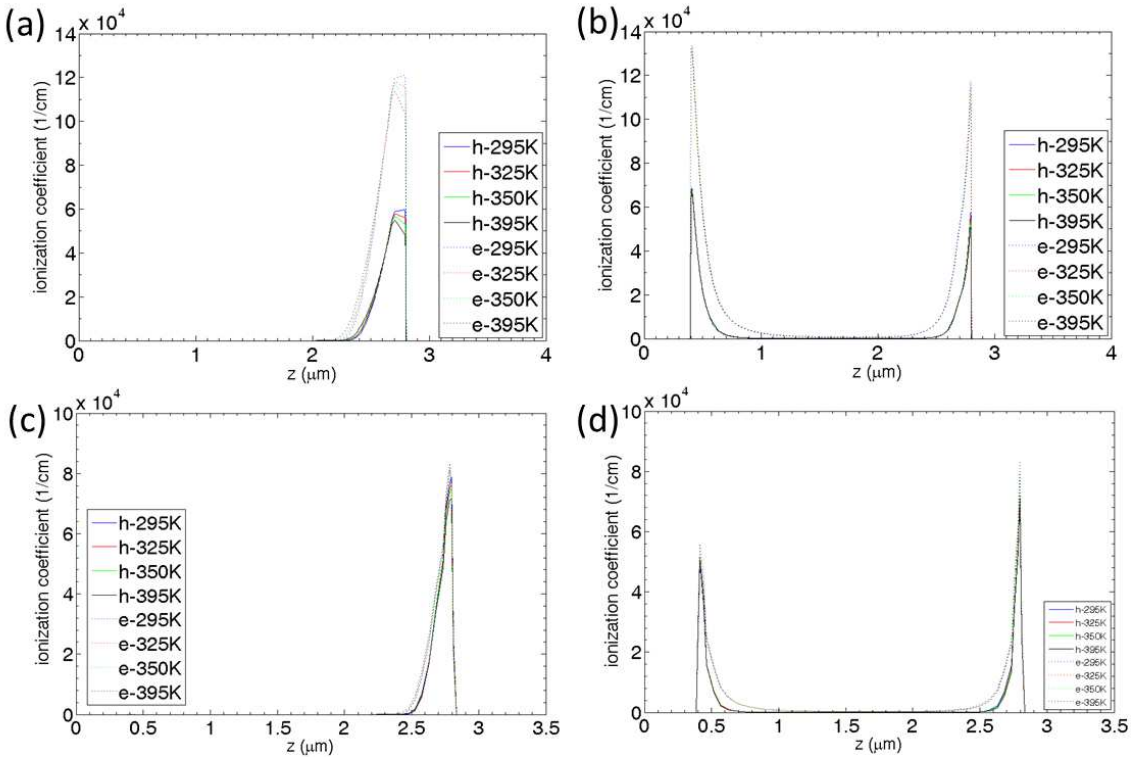


Figure 6·11: Calculated impact ionization coefficients for electrons and holes along the z direction as a function of the temperature, sampled at the center of wire. (a) APSYS at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (b) APSYS at current density of $5000\,\mu\text{A}/\mu\text{m}^2$; (c) Synopsys at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (d) Synopsys at current density of $5000\,\mu\text{A}/\mu\text{m}^2$.

number of electrons and holes, approximately $10^{17}\,\text{cm}^{-3}$, which are generated by impact ionization, flooded the P^- region leading to reduced electric field at the center of the region. We noted that this phenomenon is actually desirable for obtaining efficient luminescence devices. Indeed, from the electric field profile, one could expect a considerable radiative recombination rate in the region where the product np is maximum.

Figure 6·11 and Figure 6·12 present the corresponding impact ionization coefficient and rate. As is shown in Figure 6·11, Figure 6·12 panels (b) and (d), the rate and coefficient showed the same behavior as that in $N^+P^-N^+$ structure except for the presence of a second peak at the boundary of P^+P^- region for device current above

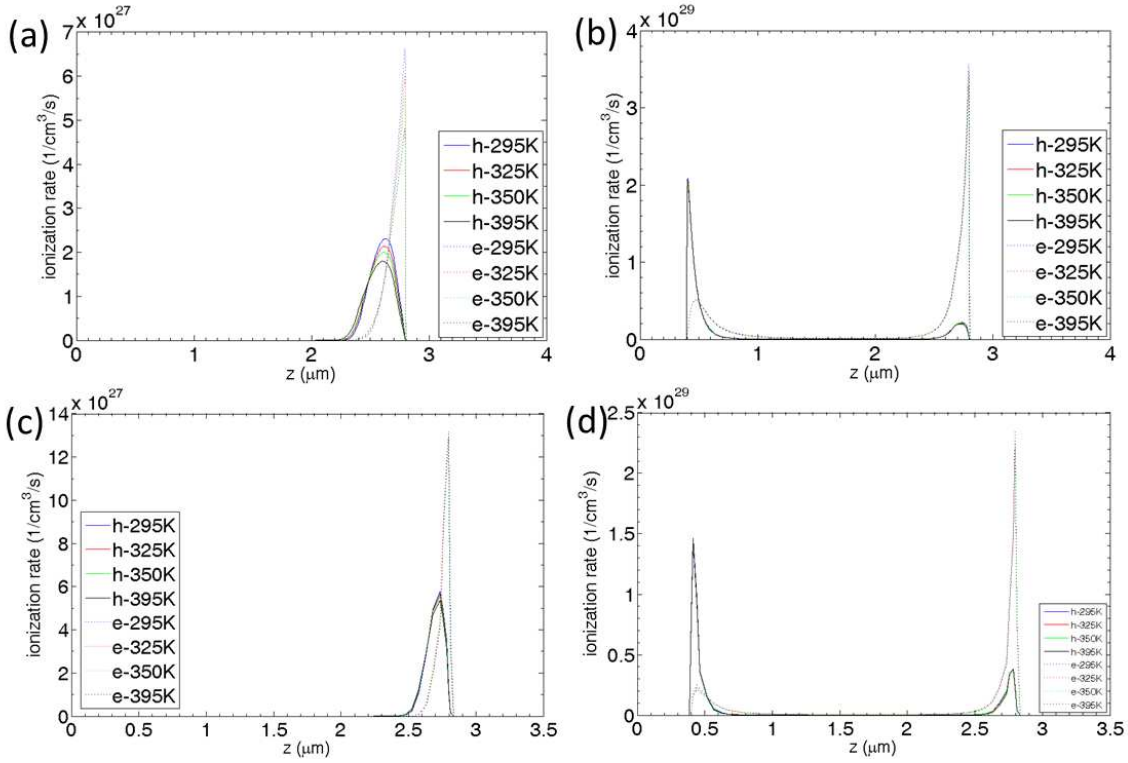


Figure 6·12: Calculated impact ionization generation rates for electrons and holes along the z direction as a function of the temperature, sampled at the center of wire. (a) APSYS at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (b) APSYS at current density of $5000\,\mu\text{A}/\mu\text{m}^2$; (c) Synopsys at current density of $300\,\mu\text{A}/\mu\text{m}^2$; (d) Synopsys at current density of $5000\,\mu\text{A}/\mu\text{m}^2$.

the critical point. This second peak of ionization coefficient is due to the large electric field at the boundary of P^+P^-. As expected, the impact ionization coefficient and rate show slightly negative temperature dependence. The discrepancies between results from APSYS and Synopsys are due to the different ionization model employed as is stated above.

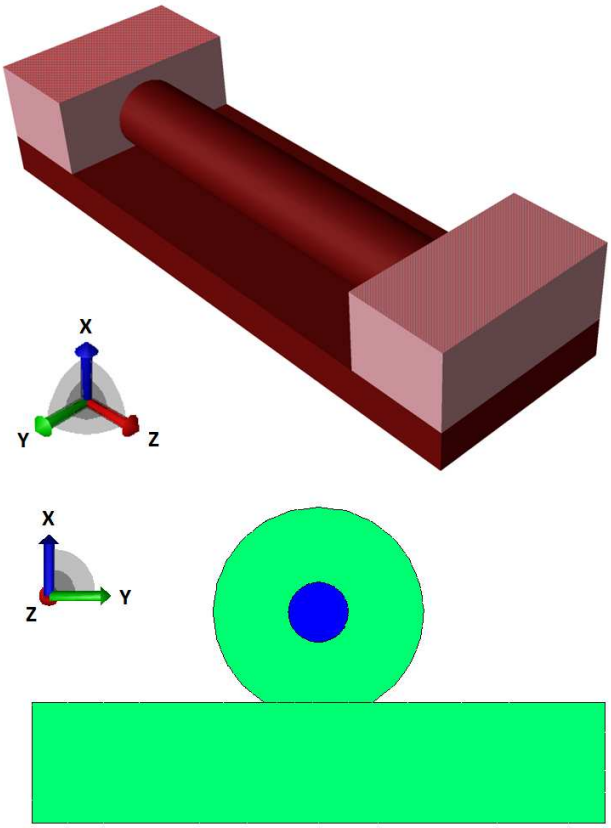


Figure 6·13: Geometrical model of the realistic $N^+P^-N^+$ silicon wire. Above: 3D geometric profile of complex N+P-N+ structure; Below: cross section of the wire region. The diameter of the wire and SiO_2 shell is 20 nm and 35 nm respectively. Dimension of the contacts regions and electrode are $40 \times 190 \times 600\,\text{nm}^3$ and $80 \times 190 \times 100\,\text{nm}^3$ respectively. Length of the wire as well as the shell is 400 nm.

6.3 Realistic $N^+P^-N^+$ Silicon Wire with Oxide

6.3.1 Geometry and Doping Profile

In this section we consider the electrical properties of a realistic silicon nanowire which can be fabricated through standard CMOS process. Figure 6·13 presents the geometrical characteristics of the device. The nanowire is cut out from a layer of silicon above a buried oxide (BOX) and is passivated by another layer of oxide around the silicon active region. Two pieces of bulk silicon located at the two ends of nanowire are

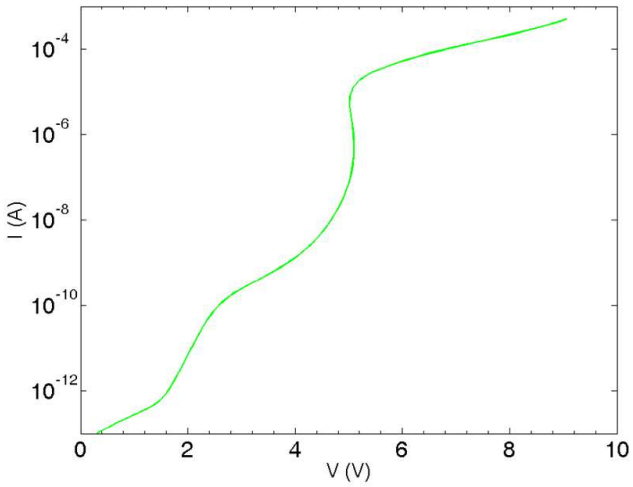


Figure 6·14: Calculated I-V characteristic of the wire obtained using Synopsys with critical point 5.1 V, and 1082 μA/μm^2.

used as the electrodes and contacts are made on the top surface of the two electrodes. In reality, many parallel wires, sharing the same bulk silicon electrodes, are etched out from the same silicon layer to form a nanowire array. Here we only simulated one unit cell as is shown in the upper part of Figure 6·13. The wire buried in the passivation layer is uniformly doped with $10^{16}\,\mathrm{cm}^{-3}$ acceptor and the two electrodes are doped with $10^{19}\,\mathrm{cm}^{-3}$ donor. The cross section of the silicon wire is plotted in the lower part of Figure 6·13, where the green region represents silicon dioxide and blue region the P^- silicon wire. The diameter of the wire is 20 nm and the SiO_2 shell has the diameter of 35 nm. The length of the nanowire is 400 nm. The two ends of wire locate at $z = 0$ and $z = 400$ nm respectively, and we further assumes that the device current flows along -z direction. For this nanowire structure, we have employed only the Synopsys simulator since APSYS cannot handle the large memory needed to represent the complex geometry.

6.3.2 Electrical Properties in Breakdown Region

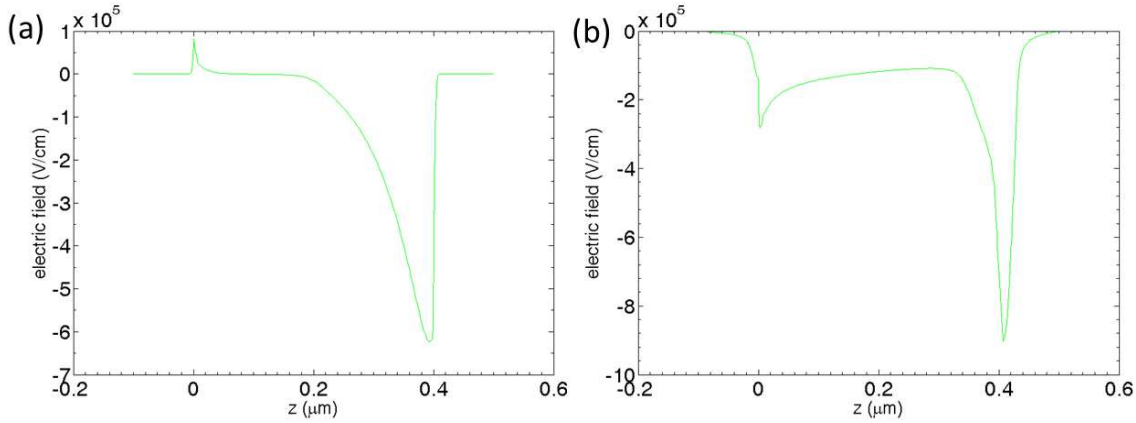


Figure 6·15: Calculated electric field at the center of finger along z direction. Breakdown happens at about 5.1 V with current density 1.59 A/μm^2. (a) is for a current density of 229 μA/μm^2; (b) is for current of 1.59 A/μm^2.

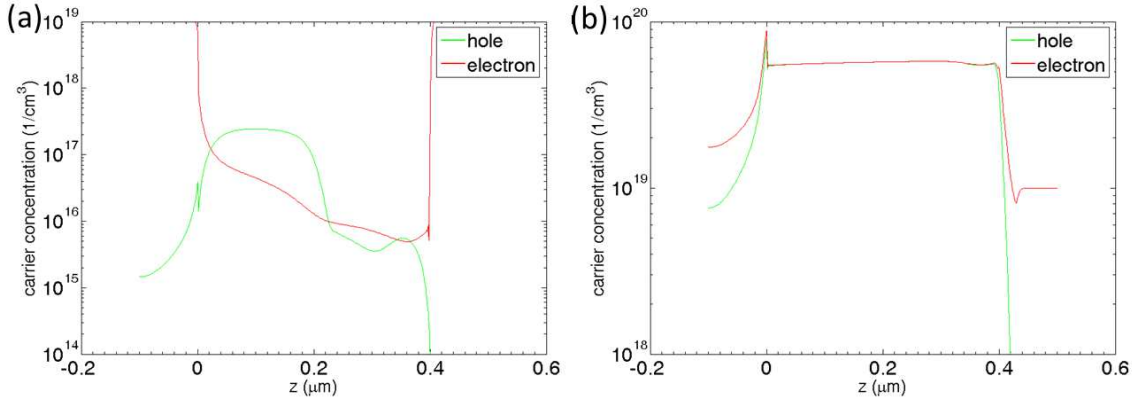


Figure 6·16: Calculated carrier distribution at the center of finger along z direction. (a) is for a current density of 229 μA/μm^2; (b) is for current of 1.59 A/μm^2.

As before, all the material parameters are set as default values of silicon and SiO_2, and lattice temperature is assumed to be 300 K. Figure 6·14 presents the calculated I-V characteristic of the device. The critical point is reached at 5.1 V, 1082 μA/μm^2, after which a small region of negative differential resistance appeared. This result is qualitatively similar to the result of the prototype $N^+P^-N^+$ wire, presented in

Chapter 6, Section 6.1 despite the fact that the two N+ regions are replaced by two pieces of bulk material here. The calculated electric field and carrier distributions at the center part of cylinder nanowire are shown in Figure 6·15 and Figure 6·16. It should be noted that when the current is above the critical point, for example at $1.59\,\mathrm{A}/\mu\mathrm{m}^2$ shown in Figure 6·16(b), a large number of electron-hole pairs are generated by impact ionization in the nanowire region that the whole region becomes neutral. This phenomenon leads to a constant electric field of $1.2 \times 10^5\,\mathrm{V/cm}$ throughout the region, as is shown in Figure 6·15(b). The calculated impact ionization coefficients and their respective rates are plotted in Figure 6·17 and Figure 6·18. These results once again are similar to the previous results for the prototype nanowire. Furthermore, from Figure 6·18, it is obvious that most of the impact ionization takes place in the P^- region with the electron process being dominant.

6.4 Conclusions of the Chapter

As an application of the device simulation using drift-diffusion model, prototypes of $N^+P^-N^+$ and $P^+P^-N^+$ nanowire and the realistic silicon wire structure have been studied in this chapter. The studied silicon nanowires can potentially be employed as light emitters in micro-display arrays which have the advantages of low cost and are fully compatible with the standard silicon process.

The results showed that, for all the structures we have considered, the calculated I-V characteristic presents a region with negative differential resistance. This is a direct consequence of driving the device at constant current. For the current density levels at which these structures are normally operated, we showed that a significant number of electron-hole pairs are generated by impact ionization in the depletion regions of *pn* junction. From the analysis of the carrier distribution in the device, we observed that at the nominal operating current density, the device functions in

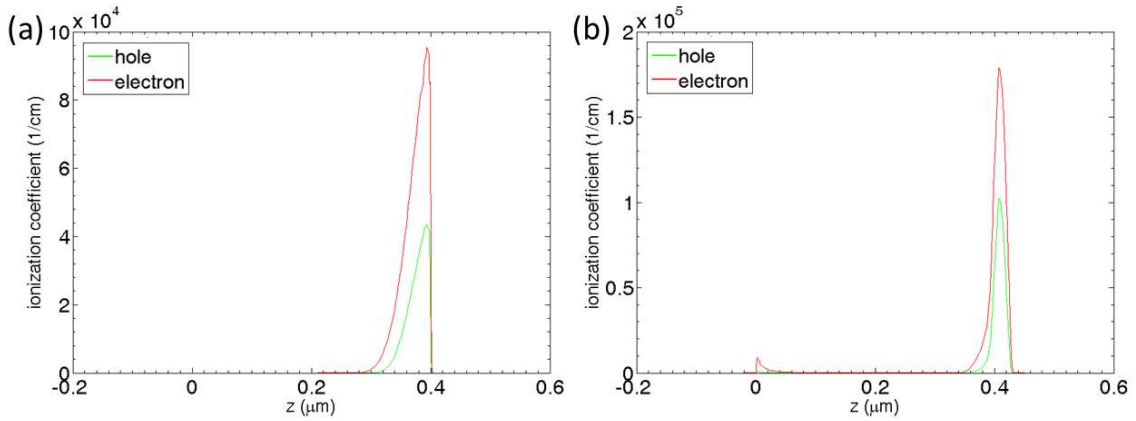


Figure 6·17: Calculated electrons and holes impact ionization coefficients at the center of finger along z direction. (a) is for a current density of 229 μA/μm^2; (b) is for current of 1.59 A/μm^2.

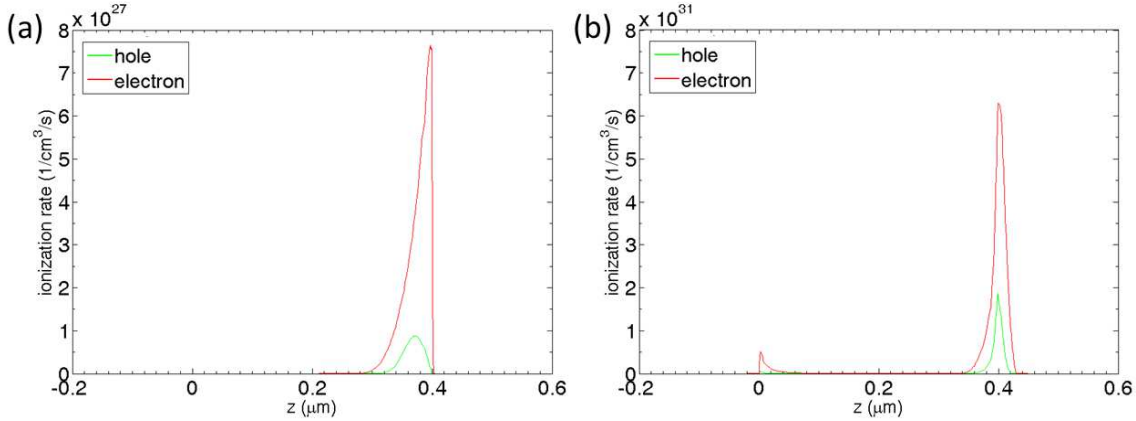


Figure 6·18: Calculated impact ionization rates at the center of finger along z direction. (a) is for a current density of 229 μA/μm^2; (b) is for current of 1.59 A/μm^2.

high injection condition. This results in free electron and hole densities in the low doping region being significantly greater than the equilibrium values. Consequently, in the active region of the nanowire, we have $n(\vec{r}) \approx p(\vec{r})$. From the point of view of optimizing the radiative recombination in these devices, this condition is closed to the optimal situation since the radiative recombination rate is maximized when both electron and hole have a similar carrier concentration in the same region of the device.

Chapter 7

3D Monte-Carlo Simulation Model

In Chapter 2, Section 2.5, we have reviewed the necessities and challenges of developing 3D Monte-Carlo simulation model based on unstructured mesh for space discretization. As is shown, a 3D model is crucial in simulating modern semiconductor devices with complex geometry and/or doping profiles. Several difficulties in developing such models have been mentioned but no solution yet has been provided. This chapter is intended to extend the discussion made in Section 2.6 and provides a study on the numerical implementation of the 3D Monte-Carlo model, which is of practical use. In particular, Section 7.1 will provide a quick way to search and locate one specific element in an unstructured tetrahedron mesh, which is of great importance in building the 3D MC model. Section 7.2 demonstrated a method to eliminate the particle-mesh forces in an tetrahedron mesh which is caused by the lack of symmetry of the hosting elements. Finally, Section 7.3 gives testing results for the proposed particle-mesh force elimination scheme.

7.1 Search in an Unstructured Mesh

Cubic or tensor meshes are traditionally used in the Monte-Carlo simulation due to the easy way of generating, storing and searching in them (Laux, 1995). Specifically, a 3D homogeneous tensor mesh can easily be stored in a 3D array based on the position of element center. The number of connection between the adjacent elements is fixed and the index of one element can be obtained by simply calculating its offset

to the reference element using the unit length of the mesh and its own position. This operation takes constant time and has $O(1)$ time complexity. However, as is presented in Section 2.6, tensor mesh does not work well with curved faces, and when the boundaries or internal interfaces of the device are not aligned with coordinate axes, excessive number of cubic elements will be need to recover the device structure, which reduces the accuracy of the numerical model and unnecessarily increases the computational load.

On the other hand, the unstructured mesh, especially the triangular mesh in 2D and tetrahedron mesh in 3D, have demonstrated excellent accuracy and sufficient efficiency in the numerical modeling of devices with irregular structures (Aldegunde et al., 2008; Agarwal et al., 2015). Furthermore, since the unstructured mesh can be refined locally, when the doping profile or the shape of junction is crucial in the simulation, it is more beneficial to employ the unstructured mesh than the tensor one. Despite the ability of reproducing complex faces and doping profiles, one difficulty of using such meshes in the calculation is implementing search operation in the mesh. A simple-minded brutal-force searching by iterating all the elements in the mesh is too expensive to use in the real device simulations. For example in the case of MC simulation, each particle need to be located every time after the drifting to update charge distribution on each node and the material/local electric field seen by the particle. Assuming there are m particles in a mesh with n elements, the brutal-force searching will have a time complexity of $O(mn)$ at every time step, which will significantly slow down the whole simulation.

In this section, a searching algorithm based on a local search by constantly approaching the destination point (e.g. the position of particles in MC simulation) is demonstrated as an efficient way of implementing search operation for tetrahedron mesh. Assuming that the destination point is P and the searching domain is convex,

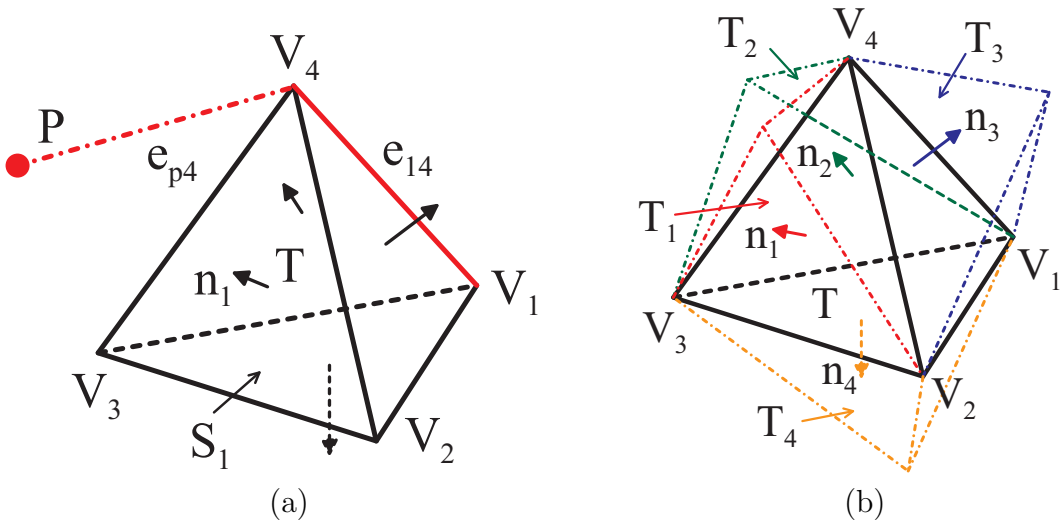


Figure 7·1: (a): Schematic of assessing the relative position of point P to the tetrahedron T. S_1 is the face opposite to vertex V_1 with normal vector n_1. (b): Schematic of adjacent tetrahedrons (T_1 to T_4, colored lines) of the center element (T, black lines). Arrows with solid head in the figure represent the normal vector of the corresponding faces.

the algorithm will try to find the element in the mesh that contains P. The procedure is the following:

- Choose an element T as the starting element of searching. In this step, we utilize a homogeneous cubic grid on top of the unstructured mesh, which contains one pointer to a tetrahedron whose geometrical center is inside the cube. The purpose of this grid is to find a starting element that is close enough to the destination to reduce the number of unnecessary searching in the area far from P. As a result, this cubic grid, which served as a coarse search, does not need to be accurate.

- Determine if P is outside of the current element T. This step will compare two singed distances to one of the face of the element. As is shown in Fig. 7·1(a), position of P relative to T is obtained by computing the signed distance of P and V_1 to S_1, which is $\mathbf{e}_{p4} \cdot \mathbf{n}_1$ and $\mathbf{e}_{14} \cdot \mathbf{n}_1$ respectively. If the two distance are

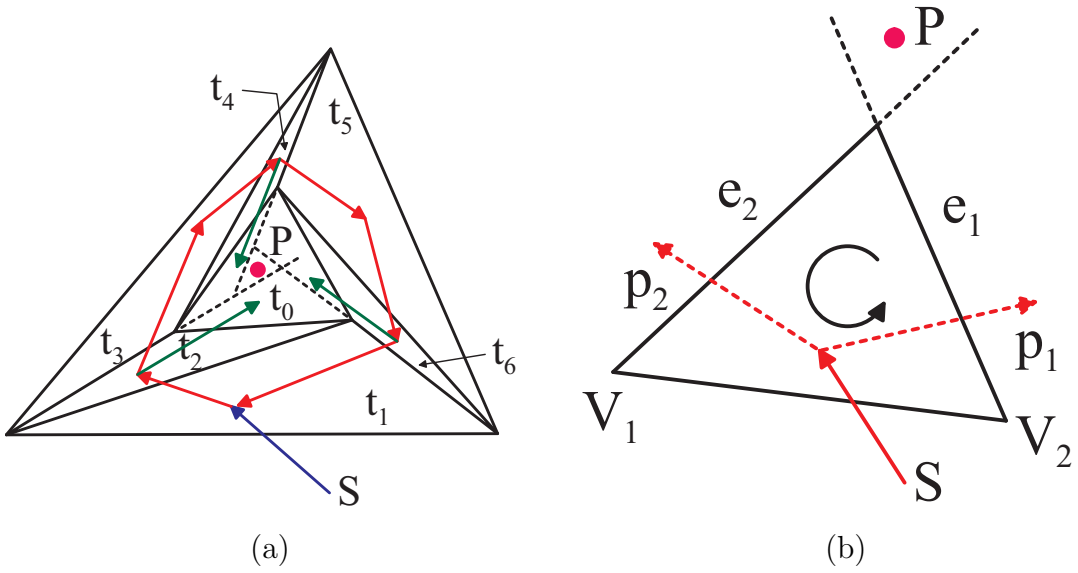


Figure 7·2: (a): Schematic of searching in the unstructured triangular mesh in 2D using walking algorithm. The starting point is S and the destination point is P. A successful searching will return the element t_0. Red arrows indicate a possible indefinite searching loop when the next element in the algorithm is chosen in a deterministic way. Green arrows show the break of searching loop when the next element in the searching is chosen randomly. (b): Schematic of randomly choosing the next element when the searching in a 2D triangular mesh.

of different signed, which is the case in Fig. 7·1(a), T will be change to one of its adjacent element T_1 that is opposite to V_1 as is shown in Fig. 7·1(b). If the two distances are of the same sign, which means P and V_1 are on the same side of S_1, similar check will be performed to node V_2 to V_4.

- The previous step will be repeated until all the four nodes of T has been checked and all pairs of distances have the same sign; or T becomes a null which means boundary of the device has been reached. In the former case, the search completes and T will the element that contains P. In the latter case, P is outside of the mesh boundary and cannot be reached from the current convex region.

Although the aforementioned algorithm has been proved to have time complexity

of $O(\log(n))$ (Devillers et al., 2001) and is guaranteed to terminate on a delaunay mesh (Shewchuk, 2001), it does not guaranteed to terminate if the mesh quality is bad (non-delaunay) and T jumps in a deterministic way. A 2D triangular mesh is used here to make the demonstration clear. As shown in Fig. 7·2(a), when starting from S, if every time the algorithm searches from the beginning of the stored vertex array of an element and jumps to a new T once it detects different signs of the two tested distances, the jumping path may end up in a loop shown by red arrows. In this case, the algorithm always try to reach to P by jumping across the first edge that separate P and the current element (from t_1 to t_6). To break the loop and make the algorithm work for any triangulation, a stochastic jump can be introduced. Shown in Fig. 7·2(b), when deciding which element to jump to as both edge e_1 and e_2 separating P with the current triangle, the algorithm can randomly choose either e_1 or e_2. Consequently, the loop in (a) can be broken at triangle t_2, t_4 and t_6 with the green arrows, and P can be reached in a finite steps.

As a final note, we want to point out that the searching operation is not necessary for drift-diffusion simulation since the matrix element and RHS have to be assembled and updated by iterating all the elements. However, in the case of MC simulation, positions of each particle need to be traced after the drift, which makes the searching operation of great importance in the numerical model.

7.2 Particle-Mesh Force Elimination

The particle-mesh force present in the MC simulations is due to the asymmetric charge distribution on the mesh nodes (Hockney and Eastwood, 1988). When the Poisson equation is discretized, charges carried by a particle need to be redistributed on its surround nodes. Unphysical electric field could be generated at the particle from its surrounding nodes if the charges are not assigned properly. Basically, the particle in

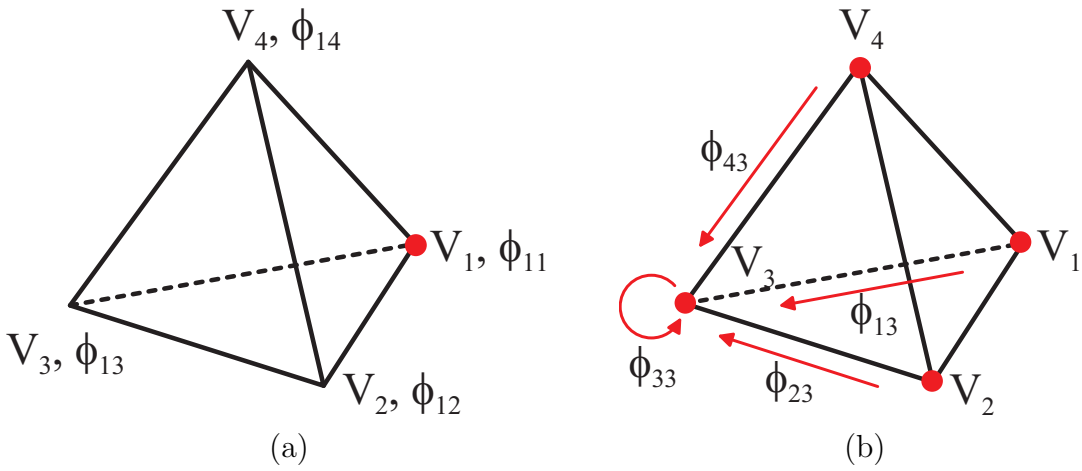


Figure 7·3: (a): Schematic of electrostatic potential at the four vertexes of a tetrahedron. A point charge has been put at V_1 and no other charge exist in the mesh. (b): Schematic of self-potential at V_3 corrected by adding the contributions from all the surrounding nodes.

this situation will see a electric field generated by the charges from itself, which is also called "self-force".

In the case of traditional tensor mesh, it has been proved that by using cloud-in-cell (CIC) scheme, one can eliminate such unphysical force if the same interpolation of charge assignment and electric field calculation have been used (Hockney and Eastwood, 1988). On the contrary, when a tetrahedron mesh that consists of non-regular elements is adopted, people have demonstrated that there is no charge assignment scheme exists that can eliminate self-force in the same way as CIC (Laux, 1995). The difficulty of eliminating the self-force in an tetrahedron mesh limited its use in the MC simulation in the past decades. In this section, we present a way to eliminate the self-force on a particle by directly canceling the force when interpolating the electric field (Aldegunde et al., 2010; Aldegunde and Kalna, 2015).

In order to correct the self-force during the electric field interpolation, we first calculated the potential generated by one unit charge placed at the location of a node V_1 in the mesh, which is described in Fig. 7·3(a). The Poisson equation in this case

becomes (assuming in a homogeneous material):

$$-\epsilon_0 \nabla^2 \phi^{v,\text{ref}} = \delta(\mathbf{r} - \mathbf{r}_v), \tag{7.1}$$

$$\phi^{v,\text{ref}} \,|_{\partial\Omega_D} = -\frac{1}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_v|}. \tag{7.2}$$

Here, $\mathbf{r}_v$ is the position of a mesh node where the unit charge is placed. Eq. (7.2) is the BC for all the boundaries of mesh. By setting the BC to the potential generated by a unit point charge at $\mathbf{r}_v$, and noticing that $\lim_{|\mathbf{r}-\mathbf{r}_v|\to\infty} = 0$, we implied that the whole space is vacuum and the mesh does not have any physical effect. Using the similar techniques of deriving Eq. (2.82), the weak form of Eq. (7.1) is:

$$\epsilon_0 \sum_{j=1}^{K} \phi_j^{v,\text{ref}} \int_\Omega \nabla N_j \cdot \nabla N_i \, \mathrm{d}\mathbf{r} = \int_\Omega \delta(\mathbf{r} - \mathbf{r}_v) \, N_i \, \mathrm{d}\mathbf{r} = N_i(\mathbf{r}_v) = \delta_{i,v}, \ \forall \ i = 1 \text{ to } K. \tag{7.3}$$

The last step used the properties of the linear interpolation function of tetrahedron element, which is shown in Eq. (2.80). It is worthy of note that although the Neumann BC is not applicable in this situation, the surface integral over the boundary Ω_D, which comes from the partial integration of the LHS of Eq. (7.1), can still be proven to be zero. As a result, Eq. (7.3) has a similar form as Eq. (2.82), which can be directly solved by FEM.

After obtaining the potential generated by one unit charge at $\mathbf{r}_v$, we record all the potentials at the direct surrounding nodes of $\mathbf{r}_v$. The potential at node i in an element T, $\phi_{i,T}$, therefore can be corrected using the $\phi_i^{j,\text{ref}}$ at each vertex i for a unit charge putting at j, where $i, j \in T$. Assuming particles locates at $\mathbf{r}_p$ which are inside T, the corrected potential at nodes of T is:

$$\phi_{i,T} = \phi_i - \epsilon_r \sum_p \sum_{j=1}^{4} q_p \, N_j(\mathbf{r}_p) \, \phi_i^{j,\text{ref}}, \quad i, j \in T. \tag{7.4}$$

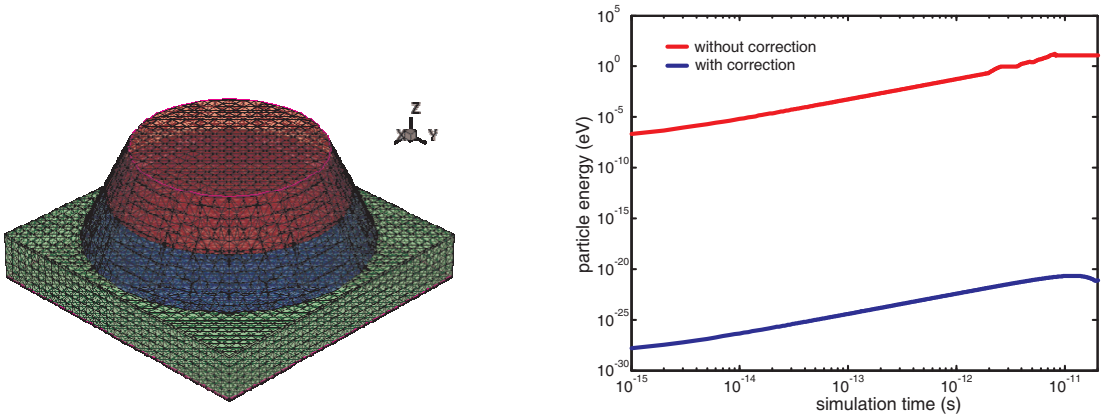
Here, ϕ_i is the potential calculated from Eq. (2.79) at node i, ϵ_r is the relative dielectric constant of T and q_p is the charge carried by the particle p. Since a unit charge is assumed when calculating the self-potential ϕ_i, when correcting the self-potential for a real particle, the actual amount of charge should be taken into consideration. Fig. 7·3(b) also provides a schematic for correcting the potential at V_3. Notice that the corrected potential $\phi_{i,T}$ is only valid inside element T, and the corresponding electric field in at position $\mathbf{r}, \mathbf{r} \in T$ is:

$$\mathbf{E}_T(\mathbf{r}) = -\sum_{i=1}^{4} \phi_{i,T} \nabla N_i(\mathbf{r}). \tag{7.5}$$

Since we are using linear interpolation function in this work, the obtained $\mathbf{E}_T$ is constant inside each tetrahedron T.

We noted that in order to cancel the self-force in all the elements in the mesh, the quantity $\phi_i^{j,\text{ref}}$ need to be calculated at all the nodes $j = 1, 2, 3, ..., K$, which means Eq. (7.3) need to be solved K times with K being the total number of nodes in the mesh. This procedure usually takes a long time. However, since the self-force is caused by the lack of symmetry of the mesh, it only relies on the geometry structure of the mesh. As a result, for the same mesh structure, $\phi_i^{j,\text{ref}},\ j = 1, 2, ..., K$ only need to be computed once. So the high computational cost for this self-force correction can be amortized.

As a final remark, one should be aware that the proposed self-force cancellation scheme in Eq. 7.4 will completely set the self-force sensed by a particle to zero inside tetrahedron T. If there are multiple particles in a same element, due to the superposition principle, the total field sensed by each particle will be very weak, which effectively underestimated the short-range Coulomb interaction. Consequently, if the electron-electron interaction is important in the MC simulation, other correction needs to be included to compensate the short-range Coulomb forces.



(a) (b)

Figure 7·4: (a): Mesh structure used in self-force calculation generated by Synopsys. (b): The calculated particle energy with(blue line)/without(red line) self-force correction when only one particle exists in a vacuum mesh.

7.3 Simulation Results

To test the effectiveness of the self-force correction scheme described in Section 7.2, we simulated the motion of one single particle in a generated tetrahedron mesh with irregular geometry boundary which is shown in Fig. 7·4(a). In this simulation, one particle with charge q is placed in the center region of the mesh without initial kinetic energy. The simulation space is assumed to be vacuum so no material properties (scattering) will be considered. Since there is no other charge in the whole space, by discretizing the Poisson equation in the same way as Eq. (2.82) with proper BCs, the particle will drift in the electric field generated by itself. One can imagine that if the self-force is not zero, the particle will begin to gain energy as the simulation running.

Fig. 7·4(b) gives the simulation results for a particle carrying 100 electron charge over a 20 ps flight. As is shown, without the self-force correction, the particle gains energy from the mesh at the very beginning of the simulation, and keep accelerating until its energy reaches several electron volts. With high energy, the particle travels

very fast towards the boundary of the mesh where it will get bounced back and forth. As a result, its energy will stay almost the same since there is no space for it to accelerate further. On the contrary, when the correction is applied, as is shown in blue line, the particle only gain a very tiny amount of energy 10^{-20} eV at the end of the simulation. We speculate that this small amount of energy comes from the floating point rounding error in the program, where the electric field cannot be canceled to exact zero. Nevertheless, the self-force correction gives a reduction of the unphysical energy by more than 10^{20} times which is sufficient for a MC simulation.

7.4 Conclusions of the Chapter

As another important device simulation model, 3D Monte-Carlo method has become a necessary tool of investigating the carrier transport properties in semiconductor devices in which the carrier scattering events are of interests. By using the unstructured mesh to discretize the simulation space, devices with complex geometry structures and doping profiles can be properly modeled with high accuracy and low computational cost. In this chapter, two important issues regarding to the use of unstructured real space mesh is discussed.

First is the tracking of simulated particles in the unstructured tetrahedron mesh. By introducing the walking algorithm (Devillers et al., 2001) in the mesh, the time complexity of searching can be greatly reduced from $O(n)$ to $O(\log(n))$ which mitigated the big disadvantage of finding a specific element in an unsorted array of mesh elements, and makes the usage of unstructured mesh applicable.

Secondly, the particle-mesh forces generated by the lack of symmetry of the unstructured mesh is studied. By using a set of pre-calculated electrostatic potentials defined at each mesh node, one can correct the obtained potential from Poisson equation in a certain way that the particle-mesh forces on the particles can be reduced

as much as 10^{20} times. This will eliminate the unphysical energy gain of the particles in the simulation and makes the whole Monte-Carlo method stable and reliable.

Chapter 8

Conclusions

In this dissertation, a suite of numerical models capable of predicting the semiconductor properties from atomic level to device level is presented. Starting from the atomic compositions and the structures of a material, full band structures including the electronic band energy, effective masses at different valleys and electron wavefunctions can be obtained using EPM. By fitting the pseudo-potentials for different atomic species in a material to the results from first-principle calculations and experimental measurements, full bands for the elemental or binary material can be obtained by solving the single electron Schrodinger equation. VCA is subsequently used to generate the electronic bands of compound alloys with disorder effects considered. After that a material model based on Green's function theory is developed and used to investigate the intrinsic recombination processes in different semiconductor materials. Both direct and phonon-assisted indirect radiative and Auger recombination have been studied and compared with the measured minority carrier lifetime which showed good agreements. Finally 3D device simulation models which can utilize the calculated material parameters are developed. Specifically, the key parts of implementing the 3D MC model are studied in terms of space discretization and particle-mesh interaction, which is crucial in modeling the modern semiconductor devices with increasingly complex geometry and doping profiles.

Electronic band structures for the semiconductor materials investigated in the dissertation have been calculated in the framework of EPM. By fitting the effective

atomic local potentials of an atom to the material parameters obtained from DFT or measured results, the realistic potential which changes rapidly near the atom core can be represented by a smoothed pseudo-potential that is easy to expand with a small number of plane waves. Therefore, both of the band structure calculation, where Schrodinger equation need to be solved using the plane wave basis, and the evaluation of overlap integral between two electronic states, which gives rise to the matrix elements for various microscopic processes, can be done quickly and accurately. Strain effect is also included in the numerical model by using a cubic spline to interpolate the atomic potentials. By changing both potential itself and its first derivative at the same time to reproduce key material properties, such as the band gap energy and deformation potentials under strain condition, we first fitted the realistic band structures of group IV materials: silicon, germanium and tin. Bands for their alloys are also investigated and constructed using VCA including both atomic potential disorder and spin-orbit coupling disorder effects. It is found that only by including the both disorder effects in the model, can we obtain band structures for GeSn that agree with the measured data. Band structures for group II-VI and III-V alloys are also studied using the same EPM model. In particular, bands for HgCdTe, InAsSb and InGaAs are presented with/without in-plane biaxial strain. It is shown that the obtained results matched to the DFT and experimental data very well, further confirming the accuracy of the EPM model.

Applying the fitted full band structures of strained silicon and germanium, we investigated the optical absorption radiative recombination rates as well as the optical gain in the Ge lasing medium using the proposed Green's function theory. Different doping levels, injection and temperatures on the PA indirect and direct processes are calculated, and the results provide additional insight into the possible ways of increasing the photon emission in the material. In fact, the calculated optical gain

for Ge indicates that while high carrier injection δn in the tensile strained material could effectively enhance the gain through the transition of CB-HH and CB-LH, extremely high δn, regardless of the tensile strain, would inevitably suppress the total optical gain by introducing strong SO-HH absorption. Specifically, at δn in excess of 10^{20}cm^{-3}, the attainable gain from CB-HH and CB-LH is totally canceled by the SO-HH transition, leaving a net absorption in the material. We have also evaluated the gain for TE and TM polarized light. A total gain of $4749\,\text{cm}^{-1}$ is predicted by our calculation for TM light in 2.1% tensilely strained Ge with n-type doping of 10^{20}cm^{-3} and injection of 10^{19}cm^{-3}. However, by adding other types of absorption such as free carrier absorption, the achievable optical gain will be reduced. Radiative recombination properties of relaxed and strained silicon were studied as well, in which, it has been demonstrated that the proposed model can indeed elucidate and reproduce the experimental data for the two-phonon events in the PA recombination process.

Doping-dependent and temperature-dependent radiative and Auger recombination lifetime and their corresponding coefficients in relaxed and strained HgCdTe, InAsSb and InGaAs, intended for operation in the LWIR, MWIR, ESWIR and SWIR spectral range, are further studied in the framework of Green's function theory. First, minority carrier lifetime in $Hg_{0.79}Cd_{0.21}Te$ for LWIR application is computed and compared with other theoretical and experimental results. A good agreement has been achieved indicating the validity of the numerical model for studying this group of materials. As to the two materials working in the MWIR band, namely, $InAs_{0.91}Sb_{0.09}$ and $Hg_{0.67}Cd_{0.33}Te$, the results indicate that when the n-doping concentration of InAsSb (HgCdTe) is greater than $5 \times 10^{15}\text{cm}^{-3}$ ($3 \times 10^{15}\text{cm}^{-3}$), the Auger process dominates over the radiative process whereas, at $T < 250\,\text{K}$, the radiative process becomes the limiting factor for the intrinsic minority carrier lifetime in both materials. It was further shown that $Hg_{0.67}Cd_{0.33}Te$ generally has slightly better intrinsic minority

carrier lifetime over $InAs_{0.91}Sb_{0.09}$ in the temperature range of 50 K to 300 K due to its larger radiative lifetime at low temperature as well as weaker direct Auger recombination rate at high temperature.

For the ESWIR material, the computed minority carrier lifetime values indicate that for $In_xGa_{1-x}As$ ($x > 0.53$) grown on InP substrate, the presence of compressive strain leads to a decrease of the Auger recombination rate and increase of radiative recombination rate. Since the dominant intrinsic recombination mechanism in this spectral range is radiative recombination, the overall intrinsic carrier lifetime in strained InGaAs alloys is shorter than that in the case of fully relaxed InGaAs. When compared to the relaxed HgCdTe, both relaxed and compressively strained InGaAs alloys showed shorter intrinsic carrier lifetime at the same cutoff wavelength at room temperature which confirms the advantage of HgCdTe as wide-band infrared detector material. However, considering the difficulties of growing HgCdTe alloys with high Mercury molar fraction, and the fact that strained InGaAs alloy may be significantly defective, it is not clear which one of the two alloys can be effectively used for ESWIR applications. While HgCdTe is very well established and mature, InGaAs may still be an alternative solution for the ESWIR detector as long as the material defects due to lattice mismatch can be controlled to avoid further degradation of the minority carrier lifetime.

For SWIR materials, namely $In_{0.53}Ga_{0.47}As$ and $Hg_{0.38}Cd_{0.62}Te$, our calculations indicate that over the n-type doping from $10^{14}cm^{-3}$ to $5 \times 10^{16}cm^{-3}$ and temperature between 50 K to 300 K, the radiative recombination process always dominates over the Auger process which is more than three orders of magnitude weaker than that in MWIR materials. Moreover, the results also suggest that $Hg_{0.38}Cd_{0.62}Te$ has a better intrinsic minority carrier lifetime compared with $In_{0.53}Ga_{0.47}As$ due to its lower radiative recombination rate.

Furthermore, this dissertation also estimated the overlap integral factor $|F_1F_2|$ using the results of Auger processes presented above. For MWIR materials, the obtained $|F_1F_2|$ usually lies around 0.1 and does not change much within the temperature between 100 K to 300 K. Specifically, for $InAs_{0.91}Sb_{0.09}$ and $Hg_{0.67}Cd_{0.33}Te$ $|F_1F_2|$ is found to be 0.07-0.085 and 0.1-0.125 respectively. In the case of SWIR materials, $|F_1F_2|$ is always below 0.1 and varies significantly with temperature. It should be emphasized that the material parameter calculations in this work unveiled the importance of Auger and radiative processes in determining the total intrinsic minority carrier lifetime in LWIR, MWIR, ESWIR and SWIR materials, which would further benefit the simulation of dark current, internal quantum efficiency and detectivity in IR detectors designs.

On the device simulation part, electrical properties of silicon LED using drift-diffusion model have been computed. In particular, prototypes of $N^+P^-N^+$ and $P^+P^-N^+$ nanowire and a realistic silicon wire structure that are currently employed as light emitters in micro-display arrays have been studied. The results showed that, for all the structures that are considered, the calculated I-V characteristic presents a region with negative differential resistance. This is a direct consequence of driving the device at constant current. For the current density levels at which these structures are normally operated, it is shown that a significant number of electron-hole pairs are generated by impact ionization in the depletion regions of *pn* junction. From the analysis of the carrier distribution in the device, we observed that at the nominal operating current density, the device functions in high injection condition. This results in free electron and hole densities in the low doping region being significantly greater than the equilibrium values. Consequently, in the active region of the nanowire, it can be observed that $n(\vec{r}) \approx p(\vec{r})$. From the point of view of optimizing the radiative recombination in these devices, this condition is closed to the optimal situation since

the radiative recombination rate is maximized when both electron and hole have a similar carrier concentration in the same region of the device.

Besides the classical device simulation using DD model, a 3D MC model is also proposed as an effort to accommodate the need for simulating the increasingly complicated modern device structures. Solutions to the two key problems in developing 3D MC models, namely, the space discretization scheme and the related searching algorithm as well as the elimination of particle-mesh forces in an unstructured mesh, have been demonstrated. It is found that by first calculating the electrostatic potentials generated by a unit charge at the mesh nodes, the artificial electric field caused by the unsymmetrical charge discretization of particles could be canceled by correcting the nodal potentials with the calculated single-charge potentials. The results showed a 10^{20} times suppression of particle energy when it drifted in the self-generated electric field for 20 ps.

Overall, this dissertation have developed a suit of numerical models that are able to calculate the semiconductor properties starting from the atomic level to the microscopic processes in the material and finally to the device level. Three groups of numerical models have been proposed which are EPM model for band structure calculation, Green's function theory for calculating intrinsic carrier recombination in the materials and 3D Monte-Carlo model for simulating complex device structures. The theoretical and experimental comparison of the numerical results from these models showed the efficacy and accuracy of numerical method on solving the real world problems. It can be expected that the computational models described in this work will benefit the ongoing effort of designing new materials and developing exotic device architectures at very low cost, which is indispensable in both academic study and industrial product development.

Appendix A

Proof and Formula Derivation of Phonon-Assisted Optical Absorption and Radiative Recombination using Green's Function Theory

In this appendix, a detailed formula derivation for absorption coefficient and radiative recombination rate using Green's function formalism is presented.

The starting point of the derivation for the absorption coefficient and radiative recombination rate in Green's function theory is the total Hamiltonian for the system of interest, which includes the Hamiltonian of electrons, photons and their interaction:

$$H = \frac{1}{2m_0}\left[-i\hbar\nabla - e\mathbf{A}(\mathbf{r},t)\right]^2 + V(\mathbf{r}) + V + \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}}. \tag{A.1}$$

where the vector potential of the radiation field is given by:

$$\begin{aligned}
\mathbf{A}(\mathbf{r},t) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \mathbf{A}(\mathbf{k},t) e^{i\mathbf{k}\cdot\mathbf{r}} \\
\mathbf{A}(\mathbf{k},t) &= \sqrt{\frac{\hbar}{2\omega_{\mathrm{ph}} v\epsilon}}\hat{\mathbf{e}}(a_{\mathbf{k}}\mathrm{e}^{-i\omega_{\mathrm{ph}}t} + a^{\dagger}_{-\mathbf{k}}\mathrm{e}^{i\omega_{\mathrm{ph}}t}).
\end{aligned} \tag{A.2}$$

By neglecting the non-linear interaction term $A^2(\mathbf{r},t)$, Eq. (A.1) can be expanded and

re-grouped as:

$$H = H_{\text{elec}} + H_{\text{photon}} + H_{\text{e-photon}} \tag{A.3}$$

in which,

$$\begin{aligned} H_{\text{elec}} &= -\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r}) + V \\ &= \sum_{l,\mathbf{k},\sigma} E_l(\mathbf{k}) C^{\dagger}_{l,\mathbf{k},\sigma} C_{l,\mathbf{k},\sigma} \end{aligned} \tag{A.4}$$

$$H_{\text{photon}} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} \tag{A.5}$$

$$H_{\text{e-photon}} = -\frac{e}{m_0}\mathbf{A}(\mathbf{r},t)\cdot\mathbf{P}. \tag{A.6}$$

We notice that, besides the electron kinetic energy and periodic potential operators for the host crystal $V(\mathbf{r})$, H_{elec} also includes the electron-phonon and the electron-electron interaction V. As a result, the carrier energies $E(\mathbf{k})$ and the electron creation and annihilation operators $C^{\dagger}_{l,\mathbf{k},\sigma}$, $C_{l,\mathbf{k},\sigma}$ then represent the realistic band structure of the bulk material. In our model, this requirement is fulfilled by employing the bands from a local empirical pseudopotential model in which the screened atomic potentials are optimized to reproduce a number of experimental band data, such as: effective masses, transition energies and deformation potentials. By representing the momentum operator $\mathbf{P}$ in the base of electron eigenvectors, Eq. (A.6) for a given photon frequency can be written as:

$$H_{\text{e-photon}} = \sum_{i,j} P_{ij} A(\mathbf{k},t) C^{\dagger}_i C_j, \tag{A.7}$$

with

$$P_{ij} = -\frac{e}{m_0}\left\langle i\left|\hat{\mathbf{e}}\cdot\mathbf{P}\right|j\right\rangle, \quad i,j \in \{(l,\mathbf{k},\sigma)\}. \tag{A.8}$$

where i, j represent two different particle states in the band structure. Assuming that only the dipole interaction contributes to the inter-band transition, the number operator for electrons (holes) in band l for all the wavevectors $\mathbf{k}$ and spin states σ, is: $\hat{N}_l = \sum_{\mathbf{k},\sigma} C^{\dagger}_{l,\mathbf{k},\sigma} C_{l,\mathbf{k},\sigma}$, and will commute with all the terms in Eq. (A.3) except for $H_{\mathrm{e-photon}}$ which is the starting point of our derivation. In the following equations, without affecting the final results, we shall drop the time dependence of $A(\mathbf{k}, t)$ and simply re-write it as A for convenience.

Defining the system density operator as $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$, the number of electrons in the conduction band can be expressed as $N_c(t) = \mathrm{Tr}[\rho(t)\hat{N}_c]$. From the Schrodinger equation, the time evolution of $\rho(t)$ can be obtained as follows:

$$\frac{\partial \rho(t)}{\partial t} = \frac{1}{i\hbar}[H, \rho(t)] \tag{A.9}$$

and the time evolution of $N_c(t)$ is:

$$\frac{\mathrm{d}}{\mathrm{d}t} N_c(t) = \frac{1}{i\hbar}\mathrm{Tr}(\rho(t)[\hat{N}_c, H]) \equiv \frac{1}{i\hbar}\langle[\hat{N}_c, H]\rangle. \tag{A.10}$$

We have introduced the common notation of trace operator in Eq. (A.10) and will keep using it thereafter. Applying the basic rules of commutator algebra, Eq. (A.10) can be simplified as:

$$\begin{aligned}
[\hat{N}_c, H] &= [\hat{N}_c, H_I] \\
&= \left[\sum_{\mathbf{k},\sigma} C^{\dagger}_{l_c,\mathbf{k},\sigma} C_{l,\mathbf{k},\sigma}, \sum_{i,j} P_{ij} C^{\dagger}_i C_j A\right] \\
&= \sum_{1,2} P_{1,2} C^{\dagger}_1 C_2 A \Delta^c_{1,2} \equiv H^c_I
\end{aligned} \tag{A.11}$$

with

$$\Delta^{c}_{1,2} = \delta_{l_c,l_1} - \delta_{l_c,l_2}. \tag{A.12}$$

where $\delta_{i,j}$ is the Kronecker delta function. Introducing a grand canonical ensemble, the density operator of the system at equilibrium can be written as:

$$\rho_0 = e^{\frac{\Omega - \bar{H}_0}{k_B T}}, \tag{A.13}$$

with

$$\bar{H}_0 = H_0 - \mu_c N_c - \mu_v N_v. \tag{A.14}$$

where $H_0 = H_{\text{elec}} + H_{\text{photon}}$ is the unperturbed system Hamiltonian and the perturbation from electron-photon interaction is $H_I = H_{\text{e-photon}}$ as it will be denoted below. Ω is the thermal dynamic potential of the system and is defined through $exp(-\Omega/(k_B T)) = \text{Tr}(exp(-\bar{H}_0/(k_B T)))$. We assume that the perturbation is adiabatically turned on at $t_0 \to -\infty$, after which the electron and hole ensembles in the conduction and valence bands remain at thermal equilibrium within themselves with the corresponding Fermi levels noted as μ_c and μ_v.

By using linear response theory, the density operator of the system can be expanded to first order as:

$$\rho(t) = \rho_0 + \rho_1(t). \tag{A.15}$$

Combining Eq. (A.15) with Eq. (A.9) and retaining the lowest order of non-vanishing terms, we have:

$$i\hbar \frac{\partial \rho_1(t)}{\partial t} = [H_0, \rho_1(t)] + [H_I, \rho_0]\,. \tag{A.16}$$

In order to find a closed form of $\rho_1(t)$ and evaluate $N_c(t)$, we consider the equivalent quantity of ρ_1 in the interaction picture $\tilde{\rho}_1(t)$, with $\rho_1(t) = e^{-iH_0 t/\hbar}\tilde{\rho}_1(t)e^{iH_0 t/\hbar}$. The

time derivative of $\rho_1(t)$ is given by:

$$i\hbar\frac{\partial\rho_1(t)}{\partial t} = [H_0, \rho_1(t)] + e^{-iH_0t/\hbar}i\hbar\ \frac{\partial\tilde{\rho}_1(t)}{\partial t}e^{iH_0t/\hbar} \tag{A.17}$$

and a direct comparison between Eq. (A.16) and Eq. (A.17) yields:

$$\tilde{\rho}_1(t) = \frac{1}{i\hbar}\int_{-\infty}^{t}\left[\tilde{H}_I(t'), \rho_0\right]\mathrm{d}t' \tag{A.18}$$

$$\tilde{H}_I(t) = e^{iH_0t/\hbar}H_Ie^{-iH_0t/\hbar}. \tag{A.19}$$

Combining Eqs. (A.10), (A.11) and (A.18) and recalling that at equilibrium $\mathrm{Tr}(\rho_0 H_I^c) = \langle H_I^c\rangle_0 = 0$, one obtains:

$$\frac{\mathrm{d}}{\mathrm{d}t}N_c(t) = \frac{1}{(i\hbar)^2}\int_{-\infty}^{t}\mathrm{d}t'\left\langle\left[\tilde{H}_I^c(t), \tilde{H}_I(t')\right]\right\rangle_0 \tag{A.20}$$

$$[H_I^c, H_I] = \sum_{1,2}\sum_{3,4}P_{1,2}P_{3,4}\Delta_{1,2}^c\left[C_1^\dagger C_2A, C_3^\dagger C_4A\right]. \tag{A.21}$$

Here $\tilde{H}_I^c(t)$ and $\tilde{H}_I(t')$ are the operators in the interaction picture of the corresponding quantities in the Schrodinger picture. To represent the commutator in Eq. (A.20) in the form of Green's function, we define two new operators: $D_1 = C_1^\dagger C_2A$ and $D_2 = C_3^\dagger C_4A$, so that the trace of the new commutator $\langle[\tilde{D}_1(t), \tilde{D}_2(t')]\rangle_0$ can be found by first evaluating the trace below:

$$\begin{aligned}\left\langle\tilde{D}_1(t)\tilde{D}_2(t')\right\rangle_0 &= \mathrm{Tr}\left(\rho_0e^{iH_0t/\hbar}D_1e^{-iH_0t/\hbar}e^{iH_0t'/\hbar}D_2e^{-iH_0t'/\hbar}\right)\\ &= e^{i(t-t')(\mu_c\Delta_{1,2}^c+\mu_v\Delta 1,2^c)/\hbar}\left\langle d_1(t-t')d_2(0)\right\rangle_0\end{aligned} \tag{A.22}$$

where the operator $d(t) = e^{i\bar{H}_0t/\hbar}De^{-i\bar{H}_0t/\hbar}$ and Baker-Hausdorff theorem (Mahan, 1993) has been used to simplify the non-commuting operator $\hat{N}$ and D. Eq. (A.20)

then becomes:

$$\frac{\mathrm{d}}{\mathrm{d}t}N_c(t) = \frac{1}{i\hbar^2}\sum_{1,2;3,4} P_{1,2}P_{3,4}\Delta^c_{1,2}\int_{-\infty}^{t} \mathrm{d}t' G^R(1,2;3,4;t-t')\; e^{i(t-t')(\mu_c\Delta^c_{1,2}+\mu_v\Delta^c_{1,2})/\hbar} \tag{A.23}$$

$$G^R(1,2;3,4;t) = \frac{1}{i}\theta(t)\langle[d_1(t), d_2(0)]\rangle_0. \tag{A.24}$$

As is shown, in Eq. (A.24) we used the commutator instead of the anti-commutator to define the three-particle retarded Green's function since the operator D consists of two fermions and one boson, making it a boson-like operator. $\theta(t)$ is the usual unit step function. The time-dependent retarded Green's function can be further manipulated by Fourier transform and Lehmann representation (Takeshima, 1982a). Performing the procedure of frequency summation in Green's function theory, we found that $G^R(t)$ can be expressed in terms of the imaginary part of its Fourier component, which is given by:

$$G^R(t) = -i\int_{-\infty}^{\infty} \mathrm{d}E' \mathrm{Im}G^R(E')e^{iE't/\hbar}. \tag{A.25}$$

Using the definition of the Dirac delta function: $\int_{-\infty}^{\infty} e^{i\omega t}\mathrm{d}t = 2\pi\delta(\omega)$, a direct evaluation of Eqs. (A.23), (A.24) and (A.25) yields:

$$\frac{\mathrm{d}}{\mathrm{d}t}N_c(t) = -\frac{2\pi}{\hbar}\sum_{1,2;3,4} P_{1,2}P_{3,4}\mathrm{Im}G^R(\mu_c - \mu_v). \tag{A.26}$$

We are now in the position to evaluate the retarded Green's function $\mathrm{Im}G^R(E)$ which will be derived from the Matsubara function of the system. Define the three-particle Matsubara function $g(\tau)$ as:

$$g(\tau) = -\left\langle T_\tau\left(C_1^\dagger(\tau)C_2(\tau)A(\tau)C_3^\dagger(0)C_4(0)A(0)\right)\right\rangle \tag{A.27}$$

with

$$C_i(\tau) = e^{\bar{H}\tau/\hbar} C_i e^{-\bar{H}\tau/\hbar}, \qquad \tau \equiv it. \tag{A.28}$$

where T_τ is the τ-ordering operator which arranges the operators with the earliest τ to the rightmost. Though Eq. (A.27) can be calculated to any order by expanding the corresponding Dyson's equation, we shall only retain the lowest order of $g(\tau)$ by using the free particle Green's function of photons to approximate each term obtained by applying Wick's theorem to Eq. (A.27) (Mahan, 1993). As a result, this approximation will ignore the polariton effect in the solid and consequently assuming that only electrons will interact with phonons. It should be noted that we did not approximate the electron Green's function with its free-particle one. Instead, we formally adopt the full Green's function of electrons (holes) whose form can be found in Chapter 2, Section 2.4. From Wick's theorem,

$$g(\tau) \simeq -\left\langle T_\tau \left(C_1^\dagger(\tau) C_2(\tau) C_3^\dagger(0) C_4(0) \right) \right\rangle \langle T_\tau A(\tau) A(0) \rangle \,. \tag{A.29}$$

The photon part is given by (Mahan, 1993):

$$g_{\rm ph}(\tau) \equiv \langle T_\tau A(\tau) A(0) \rangle \simeq \frac{\hbar}{2\omega_{\rm ph} v\epsilon} e^{-\tau\omega_{\rm ph}} \theta(\tau) \tag{A.30}$$

and the electron part can be formally written as:

$$\begin{aligned} g_2(\tau) &= -\left\langle T_\tau C_1^\dagger(\tau) C_2(\tau) \right\rangle \left\langle T_\tau C_3^\dagger(0) C_4 \right\rangle + \left\langle T_\tau C_4(0) C_1^\dagger(\tau) \right\rangle \left\langle T_\tau C_2(\tau) C_3^\dagger(0) \right\rangle \\ &= -\delta_{1,2}\delta_{3,4} n_1 n_3 + g_{2,3}(\tau) g_{4,1}(-\tau) \delta_{2,3}\delta_{4,1} \end{aligned} \tag{A.31}$$

with

$$g_{2,3}(\tau) = \left\langle T_\tau C_2(\tau) C_3^\dagger(0) \right\rangle, \quad \tau > 0 \tag{A.32}$$

$$g_{4,1}(\tau) = \left\langle T_\tau C_4(\tau) C_1^\dagger(0) \right\rangle, \quad \tau < 0. \tag{A.33}$$

Here, Eqs. (A.32) and (A.33) are the full Matsubara functions for one fermion (electron, hole). The same technique used in deriving Eq. (A.25) can also be used to represent Eqs. (A.32) and (A.33) in terms of the imaginary parts of the corresponding electron retarded Green's function. With $\tau > 0$, we find that:

$$g_{2,3}(\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}E_1' \mathrm{Im} G_{2,3}^R(E_1') e^{-\tau E_1'/\hbar} \Theta(E_1') \tag{A.34}$$

$$g_{4,1}(-\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}E_2' \mathrm{Im} G_{4,1}^R(E_2') e^{\tau E_2'/\hbar} \left[1 - \Theta(E_2')\right]. \tag{A.35}$$

Combining Eqs. (A.29), (A.30), (A.34) and (A.35), and substituting $i\omega_n$ with $\omega + i\delta$ to get the retarded Green's function from the Matsubara function, the system retarded Green's function in Eq. (A.26) becomes:

$$\begin{aligned} \mathrm{Im} G^R(E) = \frac{\hbar}{4\pi^2 \omega_{\mathrm{ph}} v \epsilon} \int_{-\infty}^{\infty} \mathrm{d}E_1' \int_{-\infty}^{\infty} \mathrm{d}E_2' \Theta(E_1') \left[1 - \Theta(E_2')\right] \\ \times \mathrm{Im} G_{2,3}^R(E_1') \mathrm{Im} G_{4,1}^R(E_2') \delta(E + E_2' - E_1' - \hbar\omega_{\mathrm{ph}}) \delta_{2,3} \delta_{4,1}. \end{aligned} \tag{A.36}$$

One should be aware that in the derivation above, we have always assumed that the time in the Green's function fulfills the condition $t > 0$. In the case of $t < 0$, however, the physical process reverses, i.e. the transition from band 1 to band 2 reverses as well. Consequently, when considering the total absorption for the material, both processes need to be involved and a reverse term should be added to Eq. (A.36). This will lead to a similar equation as Eq. (A.36) except that the Fermi factor now becomes

$\Theta(E_1') - \Theta(E_2')$ to account for the net absorption from band 1 to band 2. In the case of the radiative recombination rate, only one process should be included and Eq. (A.36) gives the correct Green's function. Using these relations:

$$\alpha(\hbar\omega_{\rm ph}) = \frac{{\rm d}N_c}{{\rm d}t} \Big/ \left(\frac{I}{\hbar\omega_{\rm ph}}\right) \tag{A.37}$$

with $I = \hbar\omega_{\rm ph} n c_0/4v$ and remembering that the subscript 1, 2, ... represent a group of states in the band structure $(l, \mathbf{k}, \sigma)$, the final expression for the net absorption coefficient between band 1 and 2 (assuming that $E_2(\mathbf{k}) > E_1(\mathbf{k})$) is:

$$\alpha_{12}(\hbar\omega_{\rm ph}) = \frac{2\pi}{n_r c_0 v \omega_{\rm ph} \epsilon_0} \sum_{\mathbf{k}} |P_{1,2}(\mathbf{k})|^2 \times \int {\rm d}E_1' \int {\rm d}E_2' \left(\Theta(E_1') - \Theta(E_2')\right)$$
$$\times \delta(\mu_c - \mu_v + E_2' - E_1' - \hbar\omega_{\rm ph}) \times {\rm Im}G_{l_1}^R(\mathbf{k}, E_1') {\rm Im}G_{l_2}^R(\mathbf{k}, E_2'). \tag{A.38}$$

The corresponding radiative recombination rate per unit volume per energy interval from band 2 to band 1, therefore becomes:

$$R_{21}(\hbar\omega_{\rm ph}) = \frac{2 n_r \omega_{\rm ph}}{\pi \hbar c_0^3 v \epsilon_0} \sum_{\mathbf{k}} |P_{1,2}(\mathbf{k})|^2 \times \int {\rm d}E_1' \int {\rm d}E_2' \Theta(E_2') \left(1 - \Theta(E_1')\right)$$
$$\times \delta(\mu_c - \mu_v + E_2' - E_1' - \hbar\omega_{\rm ph}) \times {\rm Im}G_{l_1}^R(\mathbf{k}, E_1') {\rm Im}G_{l_2}^R(\mathbf{k}, E_2'). \tag{A.39}$$

Appendix B

Discretization of Continuum Charge Distribution in Tetrahedron Mesh for Solving Poisson Equation using FEM

In the common mesh generator (e.g. Synopsys Sentaurus Mesh), the doping information in the device has been embedded in the mesh structure which is defined at each nodes in the mesh. In order to correctly use this information to solve Poisson equation in the Monte-Carlo simulation, one need to convert this doping information into a form that is compatible with Eq. (2.82). In this appendix, the last term in Eq. (2.82) representing the continuum charge distribution in the space is derived in terms of the charge density $\rho_j(\mathbf{r}_p)$ defined at each node in the tetrahedron mesh.

To evaluate the integral $\int_\Omega \rho_j N_j(\mathbf{r}) N_i(\mathbf{r})\,\mathrm{d}\mathbf{r}$, a master tetrahedron (Huebner et al., 1995) can be used to transform the integration from an arbitrary domain Ω to the master tetrahedron domain Ω_m. The master tetrahedron is defined using three normalized coordinates ξ, η, ζ, which is enclosed by four planes:

$$\begin{aligned} &\xi = 0, \quad \eta = 0, \quad \zeta = 0, \\ &\xi + \eta + \zeta = 1. \end{aligned} \tag{B.1}$$

It is easy to find out that the normalized coordinates are actually the natural coordi-

nates in the tetrahedron. Consequently, we have:

$$N_1 = \xi, \;\; N_2 = \eta, \;\; N_3 = \zeta, \;\; N_4 = 1 - \xi - \eta - \zeta. \text{ with } \xi,\ \eta,\ \zeta \in [0,\ 1] \tag{B.2}$$

The integration therefore can be represented as:

$$\int_{\Omega} \rho_j N_j(\mathbf{r}) N_i(\mathbf{r})\, \mathrm{d}\Omega = \begin{cases} \rho_j \displaystyle\int_{\Omega_m} \xi^2\, |J|\, \mathrm{d}\Omega_m, & \text{if } j = i, \\ \rho_j \displaystyle\int_{\Omega_m} \xi\, \eta\, |J|\, \mathrm{d}\Omega_m, & \text{if } j \neq i. \end{cases} \tag{B.3}$$

with $|J|$ being the Jacobian of the coordinate transformation:

$$|J| = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)}. \tag{B.4}$$

Using Eq. (B.2), the two integral in Eq. (B.3) can respectively be evaluated as:

$$\begin{aligned} \rho_j \int_{\Omega_m} \xi^2\, |J|\, \mathrm{d}\Omega_m &= \rho_j\, |J| \int_{\Omega_m} \xi^2\, \mathrm{d}\Omega_m \\ &= \rho_j\, |J| \int_0^1 \xi^2\, \mathrm{d}\xi \int_0^{1-\xi} \mathrm{d}\eta \int_0^{1-\xi-\eta} \mathrm{d}\zeta \\ &= \frac{1}{60}\, \rho_j\, |J|, \end{aligned} \tag{B.5}$$

and

$$\begin{aligned} \rho_j \int_{\Omega_m} \xi\, \eta\, |J|\, \mathrm{d}\Omega_m &= \rho_j\, |J| \int_{\Omega_m} \xi\, \eta\, \mathrm{d}\Omega_m \\ &= \rho_j\, |J| \int_0^1 \xi\, \mathrm{d}\xi \int_0^{1-\xi} \eta\, \mathrm{d}\eta \int_0^{1-\xi-\eta} \mathrm{d}\zeta \\ &= \frac{1}{120}\, \rho_j\, |J|. \end{aligned} \tag{B.6}$$

As a result, we can obtain the integral in Eq. (B.3) as:

$$\int_{\Omega} \rho_j N_j(\mathbf{r}) N_i(\mathbf{r})\, \mathrm{d}\Omega = \begin{cases} \dfrac{1}{60}\, \rho_j\, |J|, & \text{if } j = i, \\ \dfrac{1}{120}\, \rho_j\, |J|, & \text{if } j \neq i. \end{cases} \tag{B.7}$$

As a test, we assume a uniform doping in the space which indicates that $\rho_j = \rho_0$ as a constant. Notice that Eq. (B.4) can be evaluated as $|J| = 6\, V$ where V is the volume of the tetrahedron, Eq. (B.7) for one tetrahedron element gives:

$$\begin{aligned} &\sum_{j}^{i,j\in\Omega} \int_{\Omega} \rho_j N_j(\mathbf{r}) N_i(\mathbf{r})\, \mathrm{d}\Omega \\ &= \rho_0 \left(\frac{1}{60}\, |J| + 3 \times \frac{1}{120}\, |J| \right) \\ &= \frac{1}{4} \rho_0\, V. \end{aligned} \tag{B.8}$$

This result indicates that in the case of uniform doping, each node in the tetrahedron mesh will get 1/4 of the total charge enclosed by the tetrahedron that has this node as its vertex. Obviously, this result agrees with the actual charge discretization by simply averaging the charges in all the tetrahedrons to all the nodes. Eq. (B.7), as a result, can be further used to solve the Poisson equation Eq. (2.82) in the framework of FEM.

References

Adachi, S. (2005). *Properties of Group-IV, III-V and II-VI Semiconductors*. John Wiley & Sons Ltd.

Agarwal, S., Xiu, K., Bajaj, M., Johnson, J. B., Furkay, S., Oldiges, P., and Murali, K. V. R. M. (2015). Finite element based three dimensional Schrodinger solver for nano-scale devices. *Journal of Computational Electronics*, 14:163–166.

Ahrenkiel, R. K., Ellingson, R., Johnston, S., and Wanlass, M. (1998). Recombination lifetime of $In_{0.53}Ga_{0.47}As$ as a function of doping density. *Applied Physics Letters*, 72:3470–3472.

Aldaghri, O., Ikoni, Z., and Kelsall, R. W. (2012). Optimum strain configurations for carrier injection in near infrared Ge lasers. *Journal of Applied Physics*, 111:053106.

Aldegunde, M., Garcia-Loureiro, A., Martinez, A., and Kalna, K. (2008). Tetrahedral elements in self-consistent parallel 3D Monte Carlo simulations of MOSFETs. *Journal of Computational Electronics*, 7:201–204.

Aldegunde, M. and Kalna, K. (2015). Energy conserving, self-force free Monte Carlo simulations of semiconductor devices on unstructured meshes. *Computer Physics Communications*, 189:31–36.

Aldegunde, M., Seoane, N., Garcia-Loureiro, A. J., and Kalna, K. (2010). Reduction of the self-forces in Monte Carlo simulations of semiconductor devices on unstructured meshes. *Computer Physics Communications*, 181:24–34.

Allen, P. B. and Cardona, M. (1983). Temperature dependence of the direct gap of Si and Ge. *Physical Review B: Condensed Matter and Materials Physics*, 27:4760–4769.

Arslan, Y., Oguz, F., and Besikci, C. (2015). Extended wavelength SWIR InGaAs focal plane array: Characteristics and limitations. *Infrared Physics & Technology*, 70:134–137.

Bardyszewski, W. and Yevick, D. (1985). Compositional dependence of the auger coefficient for InGaAsP lattice matched to InP. *Journal of Applied Physics*, 58:2713–2723.

Bardyszewski, W. and Yevick, D. (1989). Stimulated recombination in highly excited GaAs. *Physical Review B: Condensed Matter and Materials Physics*, 39:10839–10851.

Battaglia, J., Brubaker, R., Ettenberg, M., and Malchow, D. (2007). High speed short wave infrared (SWIR) imaging and range gating. *Proceedings of SPIE*, 6541:654106.

Beattie, A. R. and Landsberg, P. T. (1959). Auger effect in semiconductors. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical*, 249:16–29.

Bellaiche, L. and Vanderbilt, D. (2000). Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites. *Physical Review B: Condensed Matter and Materials Physics*, 61:7877–7882.

Bellotti, E. and D'Orsogna, D. (2006). Numerical analysis of HgCdTe simultaneous two-color photovoltaic infrared detectors. *IEEE Journal of Quantum Electronics*, 42:418–426.

Berolo, O., Woolley, J. C., and Vechten, J. A. V. (1973). Effect of disorder on the conduction-band effective mass, valence-band spin-orbit splitting, and the direct band gap in III-V alloys. *Physical Review B: Condensed Matter and Materials Physics*, 8:3794–3798.

Bertazzi, F., Goano, M., and Bellotti, E. (2007). Electron and hole transport in bulk ZnO: A full band monte carlo study. *Journal of Electronic Materials*, 36:857–863.

Bertazzi, F., Goano, M., and Bellotti, E. (2011). Calculation of auger lifetimes in HgCdTe. *Journal of Electronic Materials*, 40:1663–1667.

Bertazzi, F., Goano, M., and Bellotti, E. (2012). Numerical analysis of indirect auger transitions in InGaN. *Applied Physics Letters*, 101:011111.

Bertazzi, F., Moresco, M., and Bellotti, E. (2009). Theory of high field carrier transport and impact ionization in wurtzite GaN. part I: A full band monte carlo model. *Journal of Applied Physics*, 106:063718.

Bloom, S. and Bergstresser, T. (1968). Band structure of a-Sn, InSb and CdTe including spin-orbit effects. *Solid State Communications*, 6:465–467.

Borrego, J. M., Saroop, S., Gutmann, R. J., Charache, G. W., Donovan, T., Baldasaro, P. F., and Wang, C. A. (2001). Photon recycling and recombination processes in 0.53 eV p-type InGaAsSb. *Journal of Applied Physics*, 89:3753–3759.

Boucaud, P., Kurdi, M. E., Ghrib, A., Prost, M., de Kersauson, M., Sauvage, S., Aniel, F., Checoury, X., Beaudoin, G., Largeau, L., Sagnes, I., Ndong, G., Chaigneau, M., and Ossikovski, R. (2013). Recent advances in germanium emission. *Photonics Research*, 1:102–109.

Boztug, C., Sanchez-Perez, J., Sudradjat, F., Jacobson, R., Paskiewicz, D., Lagally, M., and Paiella, R. (2013). Tensilely strained germanium nanomembranes as infrared optical gain media. *Small*, 9:622–630.

Brennan, K. F. (1999). *The Physics of Semiconductors With Applications to Optoelectronic Devices*. Cambridge University Press.

Bude, J., Sano, N., and Yoshii, A. (1992). Hot-carrier luminescence in silicon. *Physical Review B: Condensed Matter and Materials Physics*, 45(11):5848–5856.

Cao, G., Li, T., Tang, H., Shao, X., Li, X., and Gong, H. (2014). Performance of extended wavelength InGaAs/InAsP SWIR detector. *Proceedings of SPIE*, 9284:928406.

Capellini, G., Kozlowski, G., Yamamoto, Y., Lisker, M., Wenger, C., Niu, G., Zaumseil, P., Tillack, B., Ghrib, A., de Kersauson, M., Kurdi, M. E., Boucaud, P., and Schroeder, T. (2013). Strain analysis in SiN/Ge microstructures obtained via Si-complementary metal oxide semiconductor compatible approach. *Journal of Applied Physics*, 113:013513.

Capper, P. (1994). *Properties of narrow gap Cadmium-based compounds*. INSPEC/lnstitution of Electrical Engineers, London, UK.

Carroll, L., Friedli, P., Neuenschwander, S., Sigg, H., Cecchi, S., Isa, F., Chrastina, D., Isella, G., Fedoryshyn, Y., and Faist, J. (2012). Direct-gap gain and optical absorption in germanium correlated to the density of photoexcited carriers, doping, and strain. *Physical Review Letters*, 109:057402.

Chang, Y., Grein, C. H., and Sivananthan, S. (2006). Narrow gap HgCdTe absorption behavior near the band edge including nonparabolicity and the Urbach tail. *Applied Physics Letters*, 89:062109.

Chazapis, V., Blom, H. A., Vodopyanov, K. L., Norman, A. G., and Phillips, C. C. (1995). Midinfrared picosecond spectroscopy studies of auger recombination in InSb. *Physical Review B: Condensed Matter and Materials Physics*, 52:2516–2521.

Chen, A. R., Akinwande, A. I., and Lee, H.-S. (2005). CMOS-based microdisplay with calibrated backplane. *IEEE Journal of Solid-State Circuits*, 40:2746–2755.

Chuang, S. (2009). *Physics of Photonic Devices*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Chynoweth, A. G. (1958). Ionization rates for electrons and holes in silicon. *Physical Review*, 109:1537–1540.

Ciesla, C. M., Murdin, B. N., Pidgeon, C. R., Stradling, R. A., Phillips, C. C., Livingstone, M., Galbraith, I., Jaroszynski, D. A., Langerak, C. J. G. M., Tang, P. J. P., and Pullin, M. J. (1996). Suppression of auger recombination in arsenic-rich $InAs_{1-x}Sb_x$ strained layer superlattices. *Journal of Applied Physics*, 80:2994–2997.

Clementi, E. and Roetti, C. (1974). Roothaan-Hartree-Fock atomic wavefunctions: Basis functions and their coefficients for ground and certain excited states of neutral and ionized atoms, $z < 54$. *Atomic Data and Nuclear Data Tables*, 14:177–478.

Connelly, B. C., Metcalfe, G. D., Shen, H., and Wraback, M. (2010). Direct minority carrier lifetime measurements and recombination mechanisms in long-wave infrared type II superlattices using time-resolved photoluminescence. *Applied Physics Letters*, 97:251117.

Dargam, T. G., Capaz, R. B., and Koiller, B. (1997). Critical analysis of the virtual crystal approximation. *Brazilian Journal of Physics*, 27:299–304.

Dash, W. C. and Newman, R. (1955). Intrinsic optical absorption in single-crystal germanium and silicon at 77K and 300K. *Physical Review*, 99:1151–1155.

de Guevara, H. P. L., Rodriguez, A. G., Navarro-Contreras, H., and Vidal, M. A. (2004). Determination of the optical energy gap of $Ge_{1-x}Sn_x$ alloys with $0 < x < 0.14$. *Applied Physics Letters*, 84:4532–4534.

Devillers, O., Pion, S., and Teillaud, M. (2001). Walking in a triangulation. Technical report, Unite de recherche INRIA Sophia Antipolis.

DeWames, R., Maloney, P., Billman, C., and Pellegrino, J. (2011). Electro-optical characteristics of MWIR and LWIR planar hetero-structure p+n HgCdTe photodiodes limited by intrinsic carrier recombination processes. *Proceedings of SPIE*, 8012:801239.

Dixon, J. R. and Ellis, J. M. (1961). Optical properties of n-type indium arsenide in the fundamental absorption edge region. *Physical Review*, 123:1560–1566.

Dollfus, P., Bournel, A., Galdin, S., Barraud, S., and Hesto, P. (2004). Effect of discrete impurities on electron transport in ultrashort MOSFET using 3D MC simulation. *IEEE Transactions on Electron Devices*, 51:749–756.

D'Souza, A. I., Ionescu, A. C., Salcido, M., Robinson, E., Dawson, L., and Okerlund, D. L. (2011). InAsSb detectors for visible to MWIR high operating temperature applications. *Proceedings of SPIE*, 8012:80122S.

du Plessis, M., Aharoni, H., and Snyman, L. W. (2002). Spatial and intensity modulation of light emission from a silicon LED matrix. *IEEE Photonics Technology Letters*, 14:768–770.

du Plessis, M., Venter, P., and Bellotti, E. (2013). Spectral characteristics of hot electron electroluminescence in silicon avalanching junctions. *IEEE Journal of Quantum Electronics*, 49:570–577.

du Plessis, M., Wen, H., and Bellotti, E. (2015). Temperature characteristics of hot electron electroluminescence in silicon. *Optics Express*, 23:12605–12612.

Dutt, B., Sukhdeo, D., Nam, D., Vulovic, B., Yuan, Z., and Saraswat, K. (2012). Roadmap to an efficient germanium-on-silicon laser: Strain vs. n-type doping. *IEEE Photonics Journal*, 4:2002–2009.

Edwall, D., Dewames, R. E., Mclevige, W. V., Pasko, J. G., and Arias, J. M. (1998). Measurement of minority carrier lifetime in n-type MBE HgCdTe and its dependence on annealing. *Journal of Electronic Materials*, 27:698–702.

Fichtner, W., Rose, D. J., and Bank, R. E. (1983). Semiconductor device simulation. *SIAM Journal on Scientific and Statistical Computing*, 4:391–415.

Finkman, E. and Nemirovsky, Y. (1979). Infrared optical absorption of $Hg_{1-x}Cd_xTe$. *Journal of Applied Physics*, 50:4356–4361.

Fischetti, M. V. and Laux, S. E. (1996). Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys. *Journal of Applied Physics*, 80:2234–2252.

Fuchs, G., Schiedel, C., Hangleiter, A., Harle, V., and Scholz, F. (1993). Auger recombination in strained and unstrained InGaAs/lnGaAsP multiple quantum-well lasers. *Applied Physics Letters*, 62:396.

Galkin, G. N., Kharakhorin, F. F., and Shatkovskii, E. V. (1971). Recombination of nonequilibrium carriers in indium arsenide at high excitation levels. *Soviet physics: Semiconductors*, 5:387–391.

Geller, C. B., Wolf, W., Picozzi, S., Continenza, A., Asahi, R., Mannstadt, W., Freeman, A. J., and Wimmer, E. (2001). Computational band-structure engineering of III-V semiconductor alloys. *Applied Physics Letters*, 79:368.

Gnani, E., Reggiani, S., Rudan, M., and Baccarani, G. (2002). Extraction method for the impact-ionization multiplication factor in silicon at large operating temperatures. *IEEE Proceedings of the 32nd European Solid-State Device Research Conference (ESSDERC)*, pages 227–230.

Goano, M., Bellotti, E., Ghillino, E., Ghione, G., and Brennan, K. F. (2000). Band structure nonlocal pseudopotential calculation of the III-nitride wurtzite phase materials system. part I. binary compounds GaN, AlN, and InN. *Journal of Applied Physics*, 88:6467–6475.

Green, M., Zhao, J., Wang, A., Reece, P., and Gal, M. (2001). Efficient silicon light-emitting diodes. *Nature*, 412:805–808.

Grein, C. H., Ehrenreich, H., and Runge, E. (1997). Radiative lifetime in semiconductors: Influence of photon recycling. *Proceedings of SPIE*, 2999:277–786X.

Grein, C. H., Flatte, M. E., and Chang, Y. (2008). Modeling of recombination in HgCdTe. *Journal of Electronic Materials*, 37:1415–1419.

Grein, C. H., Young, P. M., Flatte, M. E., and Ehrenreich, H. (1995). Long wavelength InAs/InGaSb infrared detectors: Optimization of carrier lifetimes. *Journal of Applied Physics*, 78:7143–7152.

Haddadi, A., Chen, G., Chevallier, R., Hoang, A. M., and Razeghi, M. (2014). InAs/$InAs_{1-x}Sb_x$ type II superlattices for high performance long wavelength infrared detection. *Applied Physics Letters*, 105:121104.

Hall, R. N. (1952). Electron-hole recombination in germanium. *Physical Review*, 87:387.

Hansch, W. (1991). *The Drift Diffusion Equation and Its Applications in MOSFET Modeling*. Springer-Verlag Wien New York.

Hansen, M. P. and Malchow, D. S. (2008). Overview of SWIR detectors, cameras, and applications. *Proceedings of SPIE*, 6939:693901.

Harrison, D., Abram, R. A., and Brand, S. (1999). Impact ionization rate calculations in wide band gap semiconductors. *Journal of Applied Physics*, 85:8178–8185.

Haug, A. (1983). Auger recombinationin direct-gapsemiconductors: band-structure effects. *Journal of Physics C: Solid State Physics*, 16:4159–4172.

Hausser, S., Fuchs, G., Hangleiter, A., and Streubel, K. (1990). Auger recombination in bulk and quantum well InGaAs. *Applied Physics Letters*, 56:913.

He, G. and Atwater, H. A. (1997). Interband transitions in Sn_xGe_{1-x} alloys. *Physical Review Letters*, 79:1937–1940.

Herman, F., Kuglin, C. D., Cuff, K. F., and Kortum, R. L. (1963). Relativistic corrections to the band structure of tetrahedrally bonded semiconductors. *Physical Review Letters*, 11:541–545.

Hess, K. (1988). *Advanced Theory of Semicondutor Devices*. Prentice Hall.

Hess, K., editor (1991). *Monte Carlo Device Simulation: Full Band and Beyond*. Kluwer Academic Publishers.

Hesthaven, J. S. and Warburton, T. (2008). *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer Science+Business Media, LLC.

Hockney, R. W. and Eastwood, J. W. (1988). *Computer Simulation Using Particles*. Institute of Physics Publishing, Bristol and Philadelphia.

Huebner, K. H., Thornton, E. A., and Byrom, E. G. (1995). *The Finite Element Method for Engineers, 3rd Edition*. Jogn Wiley & Sons, Inc.

Humphreys, R. G. (1983). Radiative lifetime in semiconductors for infrared detection. *Infrared Physics*, 23:171–175.

Ioffe. http://www.ioffe.ru/SVA/NSM/Semicond/index.html.

Ivanov, S. V., Semenov, A. N., Solov'ev, V. A., Lyublinskaya, O. G., Terent'ev, Y. V., Meltser, B. Y., Prokopova, L. G., Sitnikova, A. A., Usikova, A. A., Toropov, A. A., and Kop'ev, P. S. (2005). Molecular beam epitaxy of type II InSb/InAs nanostructures with InSb sub-monolayers. *Journal of Crystal Growth*, 278:72–77.

Jacoboni, C. and Lugli, P. (1989). *The Monte-Carlo Method for Semiconductor Device Simulation*. Springer-Verlag Wien, New York, USA.

Jiang, Y., Teich, M. C., and Wang, W. I. (1991). Carrier lifetimes and threshold currents in HgCdTe double heterostructure and multiquantum-well lasers. *Journal of Applied Physics*, 69:6869–6875.

Jozwikowski, K., Kopytko, M., and Rogalski, A. (2011). Numerical estimations of carrier generation-recombination processes and photon recycling effect in 3-μm n-on-p HgCdTe photodiodes. *Optical Engineering*, 50:061003.

Jozwikowski, K., Kopytko, M., and Rogalski, A. (2012). Numerical estimations of carrier generation–recombination processes and the photon recycling effect in HgCdTe heterostructure photodiodes. *Journal of Electronic Materials*, 41:2766–2774.

Jungemann, C. and Meinerzhagen, B. (2003). *Hierarchical Device Simulation The Monte-Carlo Perspective*. Springer Verlag/Wien, New York, USA.

Kadantsev, E. S. and Hawrylak, P. (2011). Absolute deformation potentials and robust ab initio model for band shifts induced by (001) biaxial strain in group IIIA-VA semiconductors. *Applied Physics Letters*, 98:023108.

Keasler, C. A. and Bellotti, E. (2011). A numerical study of broadband absorbers for visible to infrared detectors. *Applied Physics Letters*, 99:091109.

Kim, J. and Fischetti, M. V. (2010). Electronic band structure calculations for biaxially strained Si, Ge, and III-V semiconductors. *Journal of Applied Physics*, 108:013710.

Kinch, M. A. (2007). *Fundamentals of Infrared Detector Materials*. SPIE Press.

Kinch, M. A., Brau, M. J., and Simmons, A. (1973). Recombination mechanisms in 8-14-μ HgCdTe. *Journal of Applied Physics*, 44:1649–1663.

Kioupakis, E., Rinke, P., Delaney, K. T., and de Walle, C. G. V. (2011). Indirect auger recombination as a cause of efficiency droop in nitride light-emitting diodes. *Applied Physics Letters*, 98:161107.

Kizilyalli, I. C., Edwards, A. P., Nie, H., Disney, D., and Bour, D. (2013). High voltage vertical GaN p-n diodes with avalanche capability. *IEEE Transactions on Electron Devices*, 60:3067–3070.

Klipstein, P., Klin, O., Grossman, S., Snapi, N., Lukomsky, I., Aronov, D., Yassen, M., Glozman, A., Fishman, T., Berkowicz, E., Magen, O., Shtrichman, I., and Weiss, E. (2011). XBn barrier photodetectors based on InAsSb with high operating temperatures. *Optical Engineering*, 50:061002.

Klipstein, P., Klin, O., Grossman, S., Snapi, N., Yaakobovitz, B., Brumer, M., Lukomsky, I., Aronov, D., Yassen, M., Yofis, B., Glozman, A., Fishman, T., Berkowicz, E., Magen, O., Shtrichman, I., and Weiss, E. (2010). MWIR InAsSb XBn detectors for high operating temperatures. *Proceedings of SPIE*, 7660:76602Y–1.

Kodama, M., Sugimoto, M., Hayashi, E., Soejima, N., Ishiguro, O., Kanechika, M., Itoh, K., Ueda, H., Uesugi, T., and Kachi, T. (2008). GaN-based trench gate metal oxide semiconductor field-effect transistor fabricated with novel wet etching. *Applied Physics Express*, 1:021104.

Krishnamurthy, S. (1996). Near band edge absorption spectra of narrow-gap III–V semiconductor alloys. *Journal of Applied Physics*, 80:4045–4048.

Krishnamurthy, S., Berding, M., and Yu, Z. (2006). Minority carrier lifetimes in HgCdTe alloys. *Journal of Electronic Materials*, 35:1369–1378.

Krishnamurthy, S., Berding, M. A., Yu, Z. G., Swartz, C. H., Myers, T. H., Edwall, D., and Dewames, R. (2005). Model for minority carrier lifetimes in doped HgCdTe. *Journal of Electronic Materials*, 34:873–879.

Krishnamurthy, S., Sher, A., and Chen, A.-B. (1997). Accurate calculation of auger rates in infrared materials. *Journal of Applied Physics*, 82:5540–5546.

Kuo, C. P., Vong, S. K., Cohen, R. M., and Stringfellow, G. B. (1985). Effect of mismatch strain on band gap in III-V semiconductors. *Journal of Applied Physics*, 57:5428–5432.

Kurdi, M. E., Bertin, H., Martincic, E., Kersauson, M. D., Fishman, G., Sauvage, S., Bosseboeuf, A., and Boucaud, P. (2010). Control of direct band gap emission of bulk germanium by mechanical tensile strain. *Applied Physics Letters*, 96:041909.

Kwong, N., Rupper, G., and Binder, R. (2009). Self-consistent t-matrix theory of semiconductor light-absorption and luminescence. *Physical Review B: Condensed Matter and Materials Physics*, 79:155205.

Laks, D. B. and Neumark, G. F. (1990). Accurate interband-auger-recombination rates in silicon. *Physical Review B: Condensed Matter and Materials Physics*, 42:5176–5185.

Laux, S. E. (1995). On particle mesh coupling in Monte Carlo semiconductor device simulation. Technical report, IBM, Engineering and Technology.

Lawson, W. D., Nielsen, S., Putley, E. H., and Young, A. S. (1959). Preparation and properties of HgTe and mixed crystals of HgTe-CdTe. *Journal of Physics and Chemistry of Solids*, 9:325–329.

Lee, S., Kwon, T. S., Nahm, K., and Kim, C. K. (1990). Band structure of ternary compound semiconductors beyond the virtual crystal approximation. *Journal of Physics: Condensed Matter*, 2:3253–3257.

Lehmann, G. and Taut, M. (1972). On the numerical calculation of the density of states and related properties. *Physica Status Solidi B: Basic Solid State Physics*, 54:469–477.

Lindberg, J., Aldegunde, M., Nagy, D., Dettmer, W. G., Kalna, K., Garcia-Loureiro, A. J., and Peric, D. (2014). Quantum corrections based on the 2D Schrodinger equation for 3D finite element Monte Carlo simulations of nanoscaled FinFETs. *IEEE Transactions on Electron Devices*, 61:423–429.

Lindle, J. R., Meyer, J. R., Hoffman, C. A., and Bartoli, F. J. (1995). Auger lifetime in InAs, InAsSb, and InAsSb-InAlAsSb quantum wells. *Applied Physics Letters*, 67:3153–3155.

Liu, J., Sun, X., Camacho-Aguilera, R., Kimerling, L., and Michel, J. (2010). Ge-on-Si laser operating at room temperature. *Optics Letters*, 35:679–681.

Liu, J., Sun, X., Pan, D., Wang, X., Kimerling, L., Koch, T., and Michel, J. (2007). Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si. *Optics Express*, 15:11272–11277.

Lopes, V. C., Syllaios, A. J., and Chen, M. C. (1993). Minority carrier lifetime in mercury cadmium telluride. *Semiconductor Science and Technology*, 8:824–841.

MacDonald, A., Vosko, S., and Coleridge, P. (1979). Extensions of the tetrahedron method for evaluating spectral properties of solids. *Journal of Physics C: Solid State Physics*, 12:2991–3002.

Madelung, O., editor (1996). *Semiconductors-Basic Data (2nd. Edition)*. Springer-Verlag, Germany.

Madelung, O., editor (2004). *Semiconductors: Data Handbook (3rd. Edition)*. Springer-Verlag, Germany.

Mahan, G. D. (1993). *Many-Particle Physics, 2nd Edition*. Plenum Press, New York, USA.

Malone, B. D. and Cohen, M. L. (2013). Quasiparticle semiconductor band structures including spin-orbit interactions. *Journal of Physics: Condensed Matter*, 25:105503.

Maranowski, K. D., Peterson, J. M., Johnson, S. M., Varesi, J. B., Childs, A. C., Bornfreund, R. E., Buell, A. A., Radford, W. A., de Lyon, T. J., and Jensen, J. E. (2001). MBE growth of HgCdTe on silicon substrates for large format MWIR focal plane arrays. *Journal of Electronic Materials*, 30:619–622.

Marchetti, S., Martinelli, M., and Simili, R. (2001). The InSb Auger recombination coefficient derived from the IR-FIR dynamical plasma reflectivity. *Journal of Physics C: Condensed Matter*, 13:7363–7369.

Marchetti, S., Martinelli, M., and Simili, R. (2002). The Auger recombination coefficient in InAs and GaSb derived from the infrared dynamical plasma reflectivity. *Journal of Physics C: Condensed Matter*, 14:3653–3656.

Masovic, D. and Vukajlovic, F. (1983). Band structure calculations of cubic metals, elementary semiconductors and semiconductor compounds with spin-orbit interaction. *Computer Physics Communications*, 30:207–217.

McCauley, J. W. (2013). Materials by design and protection materials. Technical report, Army Research Laboratory.

Meyer, J. R., Felix, C. L., Bewley, W. W., Vurgaftman, I., Aifer, E. H., Olafsen, L. J., Lindle, J. R., Hoffman, C. A., Yang, M.-J., Bennett, B. R., and Shanabrook, B. V. (1998). Auger coefficients in type II InAs/$Ga_{1-x}In_xSb$ quantum wells. *Applied Physics Letters*, 73:2857–2859.

Moontragoon, P., Pengpit, P., Burinprakhon, T., Maensiri, S., Vukmirovic, N., Ikonic, Z., and Harrison, P. (2012). Electronic properties calculation of $Ge_{1-x-y}Si_xSn_y$ ternary alloy and nanostructure. *Journal of Non-Crystalline Solids*, 358:2096–2098.

Muller, R. S. and Kamins, T. I. (2003). *Device Electronics for Integrated Circuits, 3rd Edition*. Johan Wiley & Sons. Inc.

Newman, R. and Tyler, W. W. (1957). Effect of impurities on free-hole infrared absorption in p-type germanium. *Physical Review*, 105:885–886.

Norton, P. (2002). HgCdTe infrared detectors. *Opto-electronics Review*, 10:159–174.

Olson, B. V., Shaner, E. A., Kim, J. K., Klem, J. F., Hawkins, S. D., Flatte, M. E., and Boggess, T. F. (2013). Identification of dominant recombination mechanisms in narrow-bandgap InAs/InAsSb type II superlattices and InAsSb alloys. *Journal of Applied Physics*, 103:052106.

Onat, B. M., Huang, W., Masaun, N., Lange, M., Ettenberg, M. H., and Dries, C. (2007). Ultra low dark current InGaAs technology for focal plane arrays for low-light level visible-shortwave infrared imaging. *Proceedings of SPIE*, 6542:65420L.

Picozzi, S. (2002). Accurate first-principles detailed-balance determination of auger recombination and impact ionization rates in semiconductors. *Physical Review Letters*, 89:197601.

Pratt, R. G. and Hewett, J. (1983). Minority carrier lifetime in n-type bridgman grown $Hg_{1-x}Cd_xTe$. *Journal of Applied Physics*, 54:5152–5157.

Ram-Mohan, L. R. (2002). *Finite Element and Boundary Element Applications in Quantum Mechanics*. Oxford University Press.

Rees, H. (1968). Calculation of steady state distribution functions by exploiting stability. *Physics Letters A*, 26:416–417.

Rees, H. (1969). Calculation of distribution functions by exploiting the stability of the steady state. *Journal of Physics and Chemistry of Solids*, 30:643–655.

Reggiani, S., Gnani, E., Rudan, M., Baccarani, G., Corvasce, C., Barlini, D., Ciappa, M., Fichtne, W., Denison, M., Jensed, N., Groos, G., and Stecher, M. (2004a). Experimental extraction of the electron impact-ionization coefficient at large operating temperatures. *IEEE IEDM Technical Digest*, pages 407–410.

Reggiani, S., Rudan, M., Gnani, E., and Baccarani, G. (2004b). Investigation about the high-temperature impact-ionization coefficient in silicon. *Proceeding of the 34th European Solid-State Device Research conference (ESSDERC)*, pages 245 – 248.

Ridley, B. K. (1999). *Quantum Processes in Semiconductors, 4th Edition*. Clarendon Press, Oxford.

Rogalski, A. (1994). *New Ternary Alloy Systems for Infrared Detectors*. SPIE, The International Society for Optical Engineering.

Rogalski, A. (2005). HgCdTe infrared detector material: history, status and outlook. *Reports on Progress in Physics*, 68:2267–2336.

Sasaki, T., Norman, A. G., Romero, M. J., Al-Jassim, M. M., Takahasi, M., Kojima, N., Ohshita, Y., and Yamaguchi, M. (2013). Defect characterization in compositionally graded InGaAs layers on GaAs (001) grown by MBE. *Physica Status Solidi C*, 10:1640–1643.

Schuster, J. and Bellotti, E. (2012). Analysis of optical and electrical crosstalk in small pitch photon trapping HgCdTe pixel arrays. *Applied Physics Letters*, 101:261118.

Schuster, J. and Bellotti, E. (2013). Numerical simulation of crosstalk in reduced pitch HgCdTe photon-trapping structure pixel arrays. *Optics Express*, 12:14712–14727.

Schuster, J., Keasler, C. A., Reine, M., and Bellotti, E. (2012). Numerical simulation of InAs nBn back-illuminated detectors. *Journal of Electronic Materials*, 41:2981–2991.

S*üe*ss, M., Geiger, R., Minamisawa, R. A., Schiefler, G., ans D. Chrastina, J. F., Isella, G., Spolenak, R., Faist, J., and Sigg, H. (2013). Analysis of enhanced light emission from highly strained germanium microbridges. *Nature Photonics*, 7:466–472.

Sermage, B., Heritage, J. P., and Dutta, N. K. (1985). Temperature dependence of carrier lifetime and auger recombination in 1.3 μm InGaAsP. *Journal of Applied Physics*, 57:5443.

Shewchuk, J. R. (2001). Delaunay refinement algorithms for triangular mesh generation. Technical report, Department of Electrical Engineering and Computer Science, University of California at Berkeley.

Shockley, W. and Read, W. T. (1952). Statistics of the recombinations of holes and electrons. *Physical Review*, 87:835–842.

Soref, R., Hendrickson, J., and Cleary, J. (2012). Mid- to long-wavelength infrared plasmonic-photonics using heavily doped n-Ge/Ge and n-GeSn/GeSn heterostructures. *Optics Express*, 20:3814–3824.

Sporken, R., Chen, Y. P., Sivananthan, S., Lange, M. D., and Faurie, J. P. (1992). Current status of direct growth of CdTe and HgCdTe on silicon by molecularbeam epitaxy. *Journal of Vacuum Science & Technology B*, 10:1405–1409.

Steenbergen, E. H., Connelly, B. C., Metcalfe, G. D., Shen, H., Wraback, M., Lubyshev, D., Qiu, Y., Fastenau, J. M., Liu, A. W. K., Elhamri, S., Cellek, O. O., and Zhang, Y.-H. (2011). Significantly improved minority carrier lifetime observed in a long-wavelength infrared III-V type II superlattice comprised of InAs/InAsSb. *Applied Physics Letters*, 99:251110.

Stringfellow, G. B. and Greene, P. E. (1971). Liquid phase epitaxial growth of $InAs_{1-x}Sb_x$. *Journal of The Electrochemical Society*, 118:805–810.

Taflove, A. and Hagness, S. C. (2005). *Computational Electromagnetics The Finite-Difference Time-Domain Method, 3rd Edition*. Artech House.

Takeshima, M. (1981). Theory of phonon-assisted auger recombination in semiconductors. *Physical Review B: Condensed Matter and Materials Physics*, 23:6625.

Takeshima, M. (1982a). Green's-function formalism of band-to-band auger recombination in semiconductors: Correlation effect. *Physical Review B: Condensed Matter and Materials Physics*, 26:917–930.

Takeshima, M. (1982b). Unified theory of the impurity and phonon scattering effects on auger recombination in semiconductors. *Physical Review B: Condensed Matter and Materials Physics*, 25:5390–5414.

Takeshima, M. (1984a). Intervalence-band absorption in relation to auger recombination in laser materials. *Japanese Journal of Applied Physics*, 23:428–435.

Takeshima, M. (1984b). Theory of thermal behavior of laser operation in $In_{0.53}Ga_{0.47}As$. *Journal of Applied Physics*, 56:49–56.

Tidhar, G. A. and Segal, R. (2011). New applications with a SWIR imager employing long wavelengths. *Proceedings of SPIE*, 8012:801207.

Trupke, T., Green, M. A., Wurfel, P., Altermatt, P. P., Wang, A., Zhao, J., and Corkish, R. (2003). Temperature dependence of the radiative recombination coefficient of intrinsic crystalline silicon. *Journal of Applied Physics*, 94:4930–4937.

van Roosbroeck, W. and Shockley, W. (1954). Photon-radiative recombination of electrons and holes in germanium. *Physical Review*, 94:1558–1560.

Venter, P. J., Alberts, A. C., du Plessis, M., Joubert, T. H., Goosen, M. E., van Rensburg, C. J., Rademeyer, P., and Faure, N. M. (2013a). A CMOS microdisplay with integrated controller utilizing improved silicon hot carrier luminescent light sources. *Proceedings of SPIE*, 8643:864309.

Venter, P. J. and du Plessis, M. (2014). A 128x96 pixel CMOS microdisplay utilizing hot carrier electroluminescence from junctions in reach through. *IEEE Journal of Display Technology*, 10:721–728.

Venter, P. J., du Plessis, M., Bogalecki, A. W., and van Rensburg, C. J. (2013b). Nanoscale SOI silicon light source design for improved efficiency. *Proceedings of SPIE*, 8628:86280A.

Ventura, D., Gnudi, A., Baccarani, G., and Odeh, F. (1992). Multidimensional spherical harmonics expansion of Boltzmann equation for transport in semiconductors. *Applied Mathematics Letters*, 5:85–90.

Vinter, B. (2002). Auger recombination in narrow-gap semiconductor superlattices. *Physical Review B: Condensed Matter and Materials Physics*, 66:045324.

Virgilio, M., Manganelli, C., Grosso, G., Pizzi, G., and Capellini, G. (2013). Radiative recombination and optical gain spectra in biaxially strainedn-type germanium. *Physical Review B: Condensed Matter and Materials Physics*, 87:235313.

Vodopyanov, K. L., Graener, H., Phillips, C. C., and Tate, T. J. (1992). Picosecond carrier dynamics and studies of auger recombination processes in indium arsenide at room temperature. *Physical Review B: Condensed Matter and Materials Physics*, 46:13194–13200.

Walter, J. P. and Cohen, M. L. (1972). Frequency- and wave-vector-dependent dielectric function for silicon. *Physical Review B: Condensed Matter and Materials Physics*, 5:3101–3110.

Walter, J. P., Cohen, M. L., Petroff, Y., and Balkanski, M. (1970). Calculated and measured reflectivity of ZnTe and ZnSe. *Physical Review B: Condensed Matter and Materials Physics*, 1:2661–2667.

Wang, C. C., Shin, S. H., Chu, M., Lanir, M., and Vanderwyck, A. H. B. (1980). Liquid phase growth of HgCdTe epitaxial layers. *Journal of The Electrochemical Society*, 127:175–178.

Weiler, M. H. (1981). *Semiconductors and Semimetals vol. 16*. Academic, New York.

Weisz, G. (1966). Band structure and fermi surface of white tin. *Physical Review*, 149:504–518.

Wen, H. and Bellotti, E. (2014). Numerical analysis of radiative recombination in narrow-gap semiconductors using the green's function formalism. *Journal of Electronic Materials*, 43:2841–2848.

Wen, H. and Bellotti, E. (2015a). Optical absorption and intrinsic recombination in relaxed and strained $InAs_{1-x}Sb_x$ alloys for mid-wavelength infrared application. *Applied Physics Letters*, 107:222103.

Wen, H. and Bellotti, E. (2015b). Rigorous theory of the radiative and gain characteristics of silicon and germanium lasing media. *Physical Review B: Condensed Matter and Materials Physics*, 91:035307.

Wen, H., Pinkie, B., and Bellotti, E. (2015). Direct and phonon-assisted indirect auger and radiative recombination lifetime in HgCdTe, InAsSb, and InGaAs computed using green's function formalism. *Journal of Applied Physics*, 118:015702.

W.G. Spitzer, F. T. and Logan, R. (1961). Properties of heavily doped n-type germanium. *Journal of Applied Physics*, 32:1822–1830.

Willardson, R. K. and Beer, A. C., editors (1967). *Semiconductors and Semimetals, Volume 3, Optical Properties of III-V Compounds*. Academic Press, New York.

Woolley, J. C. and Warner, J. (1964). Optical energy-gap variation in InAs-InSb alloys. *Canadian Journal of Physics*, 42:1879–1885.

CURRICULUM VITAE

Hanqing Wen

hanqing@bu.edu
Department of Electrical and Computer Engineering, Boston University
8 Saint Mary's Street, Boston, MA 02215
(857) 600 - 5316

EDUCATION

- *Boston University, College of Engineering*, Boston, MA
  - Ph.D. in Electrical Engineering Sept. 2011 - May 2016 (expected)
  - M.S. in Electrical Engineering Sept. 2011 - May 2014
- *Peking University, School of Physics*, Beijing, China
  - B.S. in Physics Sept. 2007 - July 2011

PROFESSIONAL EXPERIENCE

Boston University, Department of Electrical and Computer Engineering
Graduate Research Assistant Jun. 2012 - present

- Developing a 3D Monte-Carlo Simulation Program with Quantum Effect Included
  - A 3D self-consistent Schrodinger-Poisson solver based on the 3D real space tetrahedral mesh is being developed as a part of a 3D Monte-Carlo simulation program, which is capable of simulating semiconductor devices with complex geometry.
  - Finite element method and linear shape functions are used to discretize the Poisson and Schrodinger equations. Particle-mesh force in the unstructured tetrahedral mesh is suppressed by correcting the calculated potential with a set of pre-calculated particle self-potential.
  - LU decomposition direct solver and FGMRES iterative solver with ILU preconditioner are used in the program.
  - Phonon scattering, impurity scattering and dislocation scattering mechanism from analytical band model is included in the program.

- Modeling of Radiative and Auger Recombination in Narrow-Gap Semiconductor Materials and Devices
  - Developed and implemented a numerical many-body theory (based on Green's function theory) to model the optical absorption and Auger recombination processes in both direct and indirect band-gap semiconductor materials. Both direct and phonon-assisted indirect processes including acoustic, piezoelectric, non-polar and polar optical phonon scatterings are numerically studied with different lattice temperatures and doping concentrations.
  - Modeled the full band structures of InAsSb and InGaAs alloys using the empirical pseudo-potential method with spin-orbit coupling and strain effect included.
  - Hybrid MPI-OpenMP parallel computing technique is implemented in the code to maximize the use of the resources on multi-core cluster while limiting the total memory usage.
- Developing An Adaptive Tetrahedral Mesh Generator in Finite-Element Calculations
  - Developed a non-uniform adaptive tetrahedral mesh generator to discretize the Brillouin zone of material in the finite element analysis.
  - The mesh generator can refine the tetrahedral mesh according to the band structure and the density of state of the materials.
  - The output tetrahedral mesh supports the "walking algorithm" which allows quick search in the unstructured mesh.
- Modeling the Electronic Properties of SiGe LED
  - Simulated the carrier and electric field distribution in a SiGe PIN diode under different bias conditions using drift-diffusion and multiple-quantum-well model combined with Poisson equation.
  - Constructed the geometry models of the fabricated SiGe nanowires and simulated the spectra of photoluminescence under reach-through and punch-through bias conditions.
  - Temperature and doping dependence of the photoluminescence in the SiGe LED is also investigated.

Peking University, School of Physics
Undergraduate Research Assistant — Oct. 2009 - May 2011

- Investigation of the Far-Field Polarizations of L-shaped Aperture Optical Antennas
  - Simulated different L-shaped aperture optical antennas using FDTD method. Obtained the far-field electric field distributions of the optical antennas.

  - Experimentally verified the simulation results using a phase-locked optical chopper and a CCD imaging system.

- Simulated the optical electric field enhancement of Archimedean spiral antenna illuminated by circularly polarized light using FDTD method.
- Simulated the near-field enhancement effects of H-shaped optical antennas with slot coupling.

RESEARCH INTERESTS

- Numerical Modeling of Semiconductor Materials and Devices
- Finite Element Method
- Space Triangulation

TECHNICAL REPORT

1. Enrico Bellotti, F. Bertazzi, S. Shishehchi and H. Wen. Theoretical Study of Carrier Transport, Impact Ionization And Radiative Recombination in Si/SiGe Nanostructures. Final Report for University of Pretoria, Pretoria, South Africa.

PUBLICATIONS

Peer-Reviewed Publications

1. Hanqing Wen, Jing Yang, Weiwei Zhang, and Jiasen Zhang. Optical resonant Archimedean spiral antennas. *Journal of Nanophotonics* **5**, no. 1 (2011): 053523-053523.

2. Xu Han, Xiang Ji, Hanqing Wen, and Jiasen Zhang. H-shaped resonant optical antennas with slot coupling. *Plasmonics* **7**, no. 1 (2012): 7-11.

3. Hanqing Wen, and Enrico Bellotti. Numerical Analysis of Radiative Recombination in Narrow-Gap Semiconductors Using the Green's Function Formalism. *Journal of Electronic Materials* **43**, no. 8 (2014): 2841-2848.

4. Hanqing Wen, and Enrico Bellotti. Rigorous theory of the radiative and gain characteristics of silicon and germanium lasing media. *Physical Review B* **91**, no. 3 (2015): 035307.

5. Monuko Du Plessis, Hanqing Wen, and Enrico Bellotti. Temperature characteristics of hot electron electroluminescence in silicon. *Optics Express* **23**, no. 10 (2015): 12605-12612.

6. Hanqing Wen, Benjamin Pinkie, and Enrico Bellotti. Direct and phonon-assisted indirect Auger and radiative recombination lifetime in HgCdTe, InAsSb, and InGaAs computed using Green's function formalism. *Journal of Applied Physics* **118**, no. 1 (2015): 015702.

7. Hanqing Wen, and Enrico Bellotti. Optical absorption and intrinsic recombination in relaxed and strained $InAs_{1-x}Sb_x$ alloys for mid-wavelength infrared application. *Applied Physics Letters* **107**, no. 22 (2015): 222103.

Conference Presentations

1. Hanqing Wen and E. Bellotti. Numerical Analysis of Radiative Recombination in Direct and Indirect Narrow-Gap Semiconductors Using the Green's Function Formalism, *The U.S. Workshop on the Physics and Chemistry of II-VI Materials*, Chicago, IL (2013).

2. Enrico Bellotti, Hanqing Wen, Sara Shishehchi and Masahiko Matsubara. Semiconductor Transport Properties and Power Devices Using III-Nitrides, *38th WOCSDICE*, Delphi, Greece (2014).

3. Hanqing Wen, Benjamin Pinkie, and Enrico Bellotti. Comparison of Auger and radiative recombination lifetime in HgCdTe, InAsSb, and InGaAs. *The U.S. Workshop on the Physics and Chemistry of II-VI Materials*, Baltimore, Maryland (2014).

4. Enrico Bellotti, Hanqing Wen, Benjamin Pinkie, Masahiko Matsubara, and Francesco Bertazzi. Full-band structure modeling of the radiative and non-radiative properties of semiconductor materials and devices. *SPIE Nanoscience and Engineering*, San Diego, California (2015).

5. Hanqing Wen and Enrico Bellotti. Numerical Study of Intrinsic Recombination Lifetime in Extended SWIR: A Comparison between InGaAs and HgCdTe, *The U.S. Workshop on the Physics and Chemistry of II-VI Materials*, Chicago, IL (2015).

6. Sara Shishehchi, Hanqing Wen and Enrico Bellotti. Investigation of the channel mobility in vertical and lateral normally-off GaN MOSFETs using Monte-Carlo methods, *SPIE Photonics West*, San Francisco, CA (2016).

7. Hanqing Wen, Sara Shishehchi and Enrico Bellotti. A study of GaN-based semiconductor devices using 3D self-consistent Monte-Carlo approach incorporating quantization effects, *SPIE Photonics West*, San Francisco, CA (2016).