Physics-based multiscale modeling of III-nitride light emitters

(Article begins on next page)

Tesi di Dottorato

Physics-based multiscale modeling of III-nitride light emitters

Xiangyu Zhou

Tutore	**Coordinatore del corso di dottorato**
Prof. Francesco Bertazzi	
Prof. Michele Goano	Prof. Giovanni Ghione

March 2016

Contents

Acknowledgments

Abstract

The application of computer simulations to scientific and engineering problems has evolved to an established phase over the last decades. In the field of semiconductor device physics, Technology CAD (TCAD) has been regarded as an indispensable tool for the interpretation and prediction of device behavior. More specifically, TCAD modeling and simulation of nanostructured III-nitride light emitters still have challenging problems and is currently a topic under active research.

This thesis devotes to the theoretical and numerical investigations of III-nitride bulk and quantum structures, following a bottom-up approach aimed at modeling and understanding photoluminescence and electroluminescence these structures. In the first part, the calculation of electronic bandstructure is addressed, where a novel $\mathbf{k} \cdot \mathbf{p}$ model derived from Non-local Empirical Pseudopotential method(NL-EPM) is presented. Optical properties are then calculated employing both Poisson-$\mathbf{k} \cdot \mathbf{p}$ and a density-matrix based approach, gain and luminescence spectra can be extracted by solving the semiconductor-Bloch equation numerically. The last part of this thesis deals with the microscopic quantum transport, within the framework of the quantum-statistical nonequilibrium Greens function formalism(NEGF). While classical drift-diffusion models assume that bound carriers hold their coherence in the confined direction and unbound carriers are completely incoherent, NEGF does not distinguish between bound and unbound states and treats them on equal footing. In addition, NEGF also provides intuitive insights into energy-resolved carrier distributions, currents and coherence loss mechanisms.

The numerical computations alongside this thesis can be computationally very involved, some code developed along with this thesis is deployed on the clusters and able to scale up to more than 1000 CPU cores, thanks to the parallel implementation technique such as OpenMP and MPI, as well as HPC infrastructures available at CINECA computing center.

Chapter 1

Introduction and outline

1.1 Background and Motivation

Compound semiconductor materials enable a wide range of novel devices in various areas, such as the energy-efficient solid-state lighting and displays, GaN-based power transistors and high-definition DVD players, of which nitride semiconductor nanostructures are often at the heart.

The continuous progress in epitaxial growth and photo-lithography technology provides possibilities to fabricate semiconductor hetero-structure which exhibits better quantum confinement of carriers, utilizing more sophisticated platforms. A well-established way to speed up and reduce the costs is appreciable thanks for the advent of technological computer aided design (TCAD), which allows for early assessment of new technology and extraction of important physical parameters that are unavailable from current experimental techniques. In general, TCAD physics-based modeling can be regarded as a forward engineering tool for premature technologies and devices, while data-driven modeling remains an effective inverse engineering apparatus in industrial mass-production.

A well-known problem in GaN-based LEDs called efficiency droop is still not well understood, i.e. the LED efficiency generally is highest at low currents, as the injection current increases, the efficiency decreases dramatically. Since correlated scattering mechanism dominate the device active region, quantum transport calculations need to be applied to clarify the contribution of various loss mechanism. Traditional approach such as drift-diffusion and its quantum-corrected variants do not work well for optoelectronic devices mainly due to lack of full-quantum description of coupled scattering mechanisms and

quantum interference effects, while density matrix based approach and Non-Equilibrium Green's Function(NEGF) are promising tools able to solve these issues.

1.2   Outline

This thesis aims at addressing these issues by proposing a physics-based multiscale modeling approach. Chapter 2 presents a reliable $\mathbf{k} \cdot \mathbf{p}$ band/subband model that is derived from NL-EPM full band structure with a first-principle manner. This ab-initio procedure allows for an accurate parameter extraction that is vital to the numerical robustness of the subband solver, also known as spurious-free $\mathbf{k} \cdot \mathbf{p}$ envelope function model. The $\mathbf{k} \cdot \mathbf{p}$ model remains a workhorse for the analysis of optoelectronic devices and is used along the way in the following chapters. Chapter 3 first deals with photoluminescence with the traditional Poisson-$\mathbf{k} \cdot \mathbf{p}$ solver, however, many body effects come into play even at modest injection level, this motivate us to solve the Semiconductor-Bloch equation in density matrix form. In addition, electron-electron scattering renders the computation very expensive, in which case a large number of high dimensional integrals have to be numerically evaluated in a efficient and accurate way. Chapter 4 tries to address the modeling of electroluminescence from a device oriented point of view, in which case NEGF is able to describe (quasi-)bound and unbound scattering states on equal footing, therefore probably being the best candidate for quantum transport calculations. We illustrate the NEGF approach in a heuristic way, trying to avoid tedious derivation and showing that the idea of boundary self-energy can be derived from QTBM in a physically sensible manner. Ballistic and scattering cases are both investigated, connections with respect to classical Boltzmann transport have been pointed out. Finally, we have applied our preliminary NEGF solver to some technologically relevant structures. Extremely expensive computational demand is required for a complete NEGF calculation of a realistic LED structure, which will motivate us to pursue more effective computation software and hardware infrastructures in the future.

Chapter 2

Electronic band structures of Semiconductors

2.1 Structural properties of semiconductors

The nature of a crystal is such that the surrounding of an atom repeats itself periodically in space. We can then build the entire crystal starting form a basic building block and, by suitable operations, repeat it in space. The dimension of such a building block can change and depends on the particular crystal. There are two important concepts to understand what is a crystal essentially. The first is the **lattice**, i.e. a set of points that form a perfect periodic structure. Each point sees exactly the same environment around itself. The second concept is the **basis**, a set of atoms attached to each lattice points, so that the crystal is produced. The combination of lattice and basis yields the crystal.

A basic concept in the description of any crystalline solid is the *Bravais lattice* , which specifies the arrangement of the repeated units of the crystal. We give two equivalent definitions of a *Bravais lattice* [1]:

1   A *Bravais lattice* is an infinite array of discrete points with an arrangement and orientation that appears exactly the same, from whichever of the points the array is viewed.

2   A (three-dimensional) *Bravais lattice* consists of all points with position vectors $\mathbf{R}$ of the form

$$\mathbf{R} = n_1 \mathbf{a_1} + n_2 \mathbf{a_2} + n_3 \mathbf{a_3} \tag{2.1}$$

where $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$ are any three vectors not all in the same plane, and n_1, n_2, and n_3 range through all integral values. Thus the point $\sum n_i \mathbf{a}_i$ is reached by moving n_i steps of length $\mathbf{a}_i$ in the direction of $\mathbf{a}_i$ for $i = 1$, 2, and 3, the vectors $\mathbf{a}_i$ are called *primitive vectors*.

Fig. 2.1 shows a portion of a two-dimensional *Bravais lattice* . The primitive vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ satisfying above definition of *Bravais lattice* are indicated in the figure. The term "*Bravais lattice* " is also used to represent the set of translations $\mathbf{R}$ determined by the vectors, rather than the vectors themselves. In addition, several terms should be clarified without confusion,
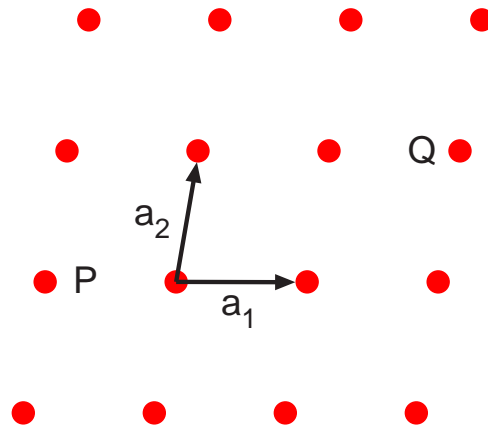
Figure 2.1: Example of 2-D Bravais lattice. All the lattice points can be computed with a linear combination of the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, for example, $P = -\mathbf{a}_1, Q = 2\mathbf{a}_1 + \mathbf{a}_2$

unit cell A unit cell is a region that fills space without any overlapping when translated through some subset of the vectors of a *Bravais lattice* , normally within the 14 Bravais systems.

primitive cell A primitive cell is smallest possible unit cell(one net lattice point per cell), there are alternative ways of choosing a primitive cell for a given *Bravais lattice* .

Wigner-Seitz cell A Wigner-Seitz cell of a lattice point is the region of space that is closer to that point than to any other lattice point, in this sense it is the smallest possible primitive cell, and now any point in space has a unique lattice point as its nearest neighbor will belong to the corresponding Wigner-Seitz cell of that particular lattice point.

In a simple cubic system, the unit and primitive cell could be the same cubes now and the Wigner-Seitz cell would be each single lattice point and its surroundings. Although

Wigner-Seitz is the smallest possible primitive cell, it doesn't mean that it is always representative of the crystal. For example, in a system with two different elements we can build the Wiger-Seitz cell, however, it won't be a primitive cell, in the sense that with only one kind of atoms we cannot reproduce the whole lattice by translational symmetry. In most analytic studies of periodic structures the reciprocal lattice plays a fundamental role. The **reciprocal lattice** is made of the set of all the **G** vectors which define plane waves having the same periodicity of the *Bravais lattice* , for any **r** and all **R**

$$e^{i\mathbf{G}\cdot(r+\mathbf{R})} = e^{i\mathbf{G}\cdot r}. \tag{2.2}$$

We can characterize the reciprocal lattice as the set of **G** vectors satisfying

$$e^{i\mathbf{G}\cdot\mathbf{R}} = 1 \tag{2.3}$$

for all the **R** vectors of the *Bravais lattice* .

The reciprocal lattice is thus itself a *Bravais lattice* , whose primitive vectors are

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \tag{2.4a}$$

$$\mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \tag{2.4b}$$

$$\mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \tag{2.4c}$$

We can simply verify that the above equations can provide a set of primitive vectors for the reciprocal lattice by noticing that

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij} \tag{2.5}$$

where δ_{ij} is the Kronecker delta symbol.

Now any vector **G** can be written as a linear combination of $\mathbf{b}_i$

$$\mathbf{G} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3 \tag{2.6}$$

Take in to account (2.1), then it follows from (2.5) that

$$\mathbf{G} \cdot \mathbf{R} = 2\pi(k_1 n_1 + k_2 n_2 + k_3 n_3) \tag{2.7}$$

Equation (2.7) can satisfy (2.3) only if the coefficients k_i are integers, in this sense we can say the reciprocal lattice is a *Bravais lattice* and the b_i can represent corresponding primitive vectors.

The most used semiconductors for electronic applications, besides the materials studied in this work, have two types of crystal lattice:

1. **diamond** or **zincblende** (Fig. 2.2);

2. **wurtzite** (Fig. 2.3).

In diamond or zincblende structures the basis set consists of two atoms, which are the
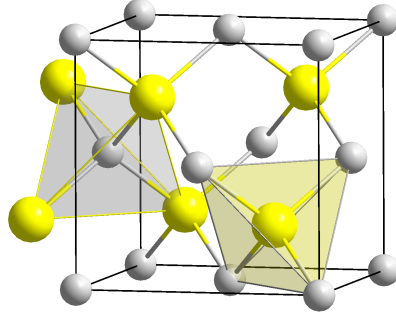
Figure 2.2: Zincblende structure.
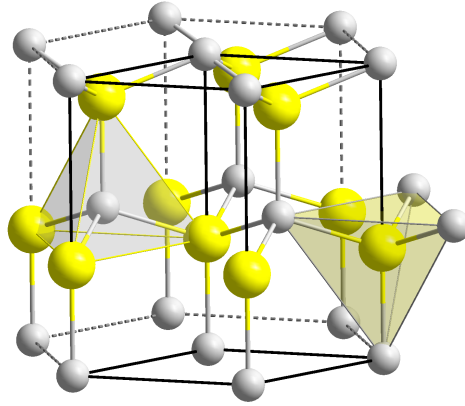
Figure 2.3: Wurtzite structure.

same in the former and different in the latter. Since we are more interested in compound semiconductor, for cubic material we will focus on the zincblende type. This lattice consists of two interpenetrating face-centered cubic (**fcc**) lattices, displaced from one another by one-fourth of the cube main diagonals. The zincblende lattice is not a Bravais lattice because the elementary cell contains two atoms, one located at the origin and the other at $\left(\frac{a}{4}, \frac{a}{4}, \frac{a}{4}\right)$, where a is the cube side length and is called the **lattice constant**. The reciprocal lattice of the Bravais lattice underlying the zincblende lattice (i.e. a fcc lattice) is a body centered cubic **bcc** lattice. A possible choice for the primitive lattice vectors of a fcc lattice is

$$\mathbf{a}_1 = \left(0, \frac{a}{2}, \frac{a}{2}\right) \tag{2.8a}$$

$$\mathbf{a}_2 = \left(\frac{a}{2}, 0, \frac{a}{2}\right) \tag{2.8b}$$

$$\mathbf{a}_3 = \left(\frac{a}{2}, \frac{a}{2}, 0\right). \tag{2.8c}$$

The primitive vectors of the reciprocal space are thus

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} = \left(-\frac{2\pi}{a}, \frac{2\pi}{a}, \frac{2\pi}{a}\right) \tag{2.9a}$$

$$\mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} = \left(\frac{2\pi}{a}, -\frac{2\pi}{a}, \frac{2\pi}{a}\right) \tag{2.9b}$$

$$\mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} = \left(\frac{2\pi}{a}, \frac{2\pi}{a}, -\frac{2\pi}{a}\right), \tag{2.9c}$$

which define a bcc lattice. Wurtzite structure is an hexagonal structure (**hcp**, hexagonal close packet), and is defined by two lattice constants a and c, and an additional internal parameter u. Possible choices for primitive lattice vectors are

$$\mathbf{a}_1 = \left(\frac{\sqrt{3}a}{2}, -\frac{a}{2}, 0\right)$$

$$\mathbf{a}_2 = \left(\frac{\sqrt{3}a}{2}, \frac{a}{2}, 0\right)$$

$$\mathbf{a}_3 = (0, 0, c)$$

or

$$\mathbf{a}_1 = (a, 0, 0)$$

$$\mathbf{a}_2 = \left(-\frac{a}{2}, \frac{\sqrt{3}a}{2}, 0\right)$$

$$\mathbf{a}_3 = (0, 0, c).$$

The correspondence between reciprocal lattice and families of lattice planes provides a convenient way to specify the orientation of a lattice plane, here we outline the brief sequence on how to specify the planes and directions by using the Miller indices.

To specify directions

1. Draw a vector through the origin.

2. Determine the coordinates of any point on the vector.

3. Multiply the resulting three numbers by a common factor to convert them to the smallest possible integers.

4. Enclose the resulting three integers in square brackets:[].

To specify planes

1. Choose a plane that does not contain the origin.

2. Determine the intercepts of the plane on the three axes.

3. Take the reciprocals of the intercepts.

4. Multiply the reciprocals by the smallest common factor that will clear all fractions.

5. Enclose the resulting integers in parentheses:().

Following the above steps, we can use miller indices to denote the specific planes in hexagonal lattices for instance, in this case it is convenient to use four basis vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $\mathbf{c}$, as seen in Fig. 2.4 we can use the $(hijk)$ notation to represent different planes, note that the relation among $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ is $-i = h + k$.
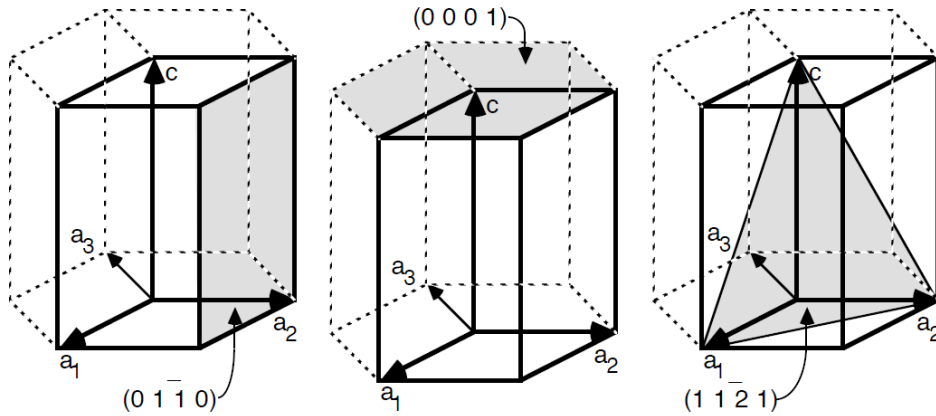
Figure 2.4: Different planes: $(01\bar{1}0), (0001), (11\bar{2}1)$ of hexagonal lattices denoted with Miller indices.

Specifically, for Wurtzite GaN, which is of great interest to us, the different planes, namely, c-plane, a-plane, r-plane and m-plane[2] are shown in Fig. 2.5 together with their miller indices and polarity.

The polarity handling of III-V nitride semiconductors is another critical issue, since the crystal is not inversion symmetric with respect to the c-axis, the result of this anisotropy is a permanent polarization of the crystal, i.e. the so called **spontaneous polarization**. It is then straightforward that GaN grown on (0001) planes may have a pair of opposite polarity, denoted as Ga(+c) or N(-c) face polarity highlighted in Fig. 2.6 [3]. This property offer a possibility to experimentally control the polarity, usually films grown by MOCVD and MBE have +c and -c polarity, respectively.

Plane	Index	Polarity
c-plane	(0001)	polar
a-plane	(11-20)	nonpolar
r-plane	(1-102)	semi-polar
m-plane	(10-10)	nonpolar
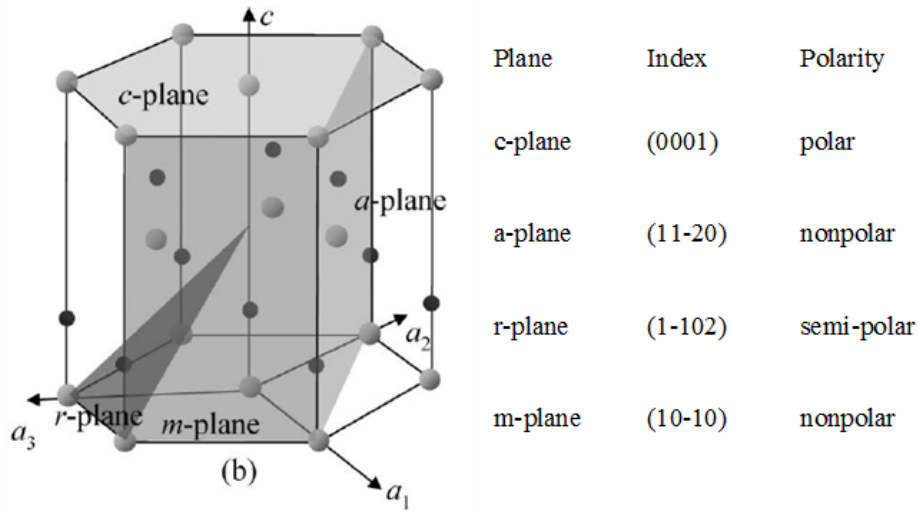
Figure 2.5: Different planes: $c-plane, a-plane, r-plane, m-plane$ of Wurtzite structure with Miller indices and polarity [2].
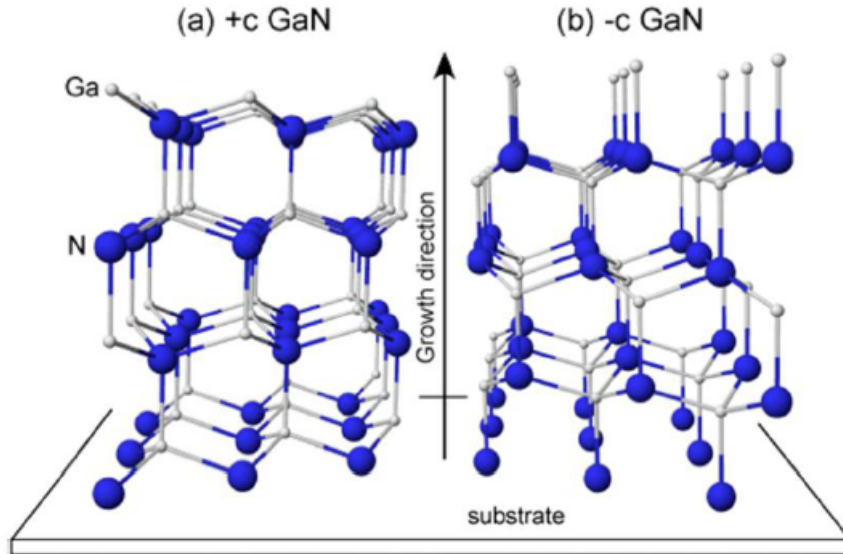
Figure 2.6: The small and large spheres indicate Ga and N, respectively. GaN with Ga-face (+c) polarity on left side and GaN with N-face (-c) polarity on right side [3]

2.2 Pseudopotential and EPM

The motion of a particle within the framework of quantum mechanics can be well described by the **Schrödinger equation** in the abbreviated form

$$H\psi(r) = E(k)\psi(r). \tag{2.10}$$

This is the time-independent form, where H is the Hamiltonian of the system and $E(k)$ is the energy. However, a solid system has many atoms, each atom consists of a positive

9

nuclei surrounded by a set of electrons, the general Hamiltonian is written as

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{i,I} \frac{Z_I e^2}{|r_i - R_I|} + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|r_i - r_j|} - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J e^2}{|R_I - R_J|} \quad (2.11)$$

where $\hbar$ is the Planck constant divided by 2π, m_e is the electron mass, M the nuclei mass, Z the atomic number of the nuclei, e the electronic charge, r the electronic position and R the nuclei position. I and J denote quantities related to the nuclei while i and j are related to the electrons. We have three interaction terms, ion-ion, electron-electron and the electron-ion coupling term, as well as kinetic energy of ions and electrons, as such a complicated coupled system.

To simplify, we need to make a number of simplifications in order to set up a soluble problem.

1   Of all the electrons in the solid, only a fraction will determine most of the properties, namely, the valence electrons in the outmost incompletely filled shells. The left core electrons are those tightly bounded to the nuclei, so they can be lumped with the nuclei to form the so-called ion cores. This approximation is known as the **frozen-core approximation**.

2   The ions are much heavier than the electrons, so they move much more slowly. As a result, electrons can respond to ionic motion almost instantaneously and, on the other hand, ions cannot follow the motion of the electrons and they see only a time-averaged adiabatic electronic potential. Thus, we can omit the fourth term of (2.11) (kinetic energy of the nuclei), and consider the nuclei as a fixed external potential acting on the electrons. The next approximation invoked is the **Born-Oppenheimer** or **adiabatic** approximation.

 After we have introduced the above approximations, we come to the simplified Hamiltonian, which can be expressed as

$$H_e = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{i,I} \frac{Z_I e^2}{|r_i - R_I|} + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|r_i - r_j|}. \quad (2.12)$$

The next step is to further simplify this Hamiltonian by using what is known as the **mean-field approximation**: We can describe the motion of a single electron assuming that it feels an average force $V(r)$ due to the vibrating lattice and all the other screening particles in the systems. Thus the Schrödinger equations describing the motion of each electron will be identical and given by

$$H_{1e}\phi_n(r) = \left(-\frac{\hbar^2}{2m_e}\nabla^2 + V(r)\right)\phi_n(r) = E_n\phi_n(r), \quad (2.13)$$

where H_{1e}, $\phi_n(r)$ and E_n denote, respectively, the one-electron Hamiltonian, and the wave function and energy of an electron in an eigenstate labeled by n. Also there exists more complex quantum mechanical formalism for many-body problems which is beyond the scope of this thesis, even this single-electron problem is difficult in general because of complex spatial and temporal variations of the potential energy $\boldsymbol{V(r)}$.

The potential energy $V(r)$ felt by a single valence electron can be formally separated into a macroscopic part U and a microscopic part V_L, where the former usually comes from the externally applied voltage or heterostructure band edge, the latter is due to the periodic crystal potential, this concept will become more concrete in the following chapters.

The starting point is the **nearly free electron** model basically considering the electrons delocalized in the whole crystal, based on which the electron states can be described by a superposition of plane waves. However, the plane wave basis set is not effective in describing the wavefunctions close to the core where they are highly oscillating. It is then worth considering the form of the potential term in Schrödinger equation, is it possible if we apply some change such that the wavefunctions are modeled effectively and still the potential accounted for the chemical bond is accurately described to some extent. Based on this thought, Phillips and Kleinman[4] and developed their original **Pseudopotential**, formally we have

$$H\phi_{n\mathbf{k}}(r) + \int V_R(rr')\phi_{n\mathbf{k}}(r')dr' = E_n\phi_{n\mathbf{k}}(r) \tag{2.14}$$

where the non-local potential

$$V_R(rr') = \sum_p \sum_q u_j^*(\mathbf{k}, r' - R_p)[E_n - E_j]u_j(\mathbf{k}, r - R_q) \tag{2.15}$$

futher seprate the Hamiltonian to the kinetic part and periodic potential we get

$$T\phi_{n\mathbf{k}}(r) + \int V_{ps}(rr')\phi_{n\mathbf{k}}(r')dr' = E_n\phi_{n\mathbf{k}}(r) \tag{2.16}$$

where pseudopotential V_{ps} is the superimposition of V_L and V_R

$$V_{ps}(rr') = V_L\delta(rr') + V_R(rr') \tag{2.17}$$

The short range repulsive potential $V_R(rr')$ then kind of cancels the long range attractive potential V_L, resulting in the pseudopotential V_{ps}, which justifies the nearly free model. Fig. 2.7 shows qualitatively how a generic pseudopotential varies with distance r in real space from the nucleus.

In this spirit of plane waves, we can expand the periodic part of the Bloch function

$$\Phi_{n\mathbf{k}} = \frac{1}{\sqrt{V}}e^{i\mathbf{k}r}u_{n\mathbf{k}}(r) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}r}\sum_{\mathbf{G}_i}a_{\mathbf{G}_i}e^{i\mathbf{G}_ir} \tag{2.18}$$
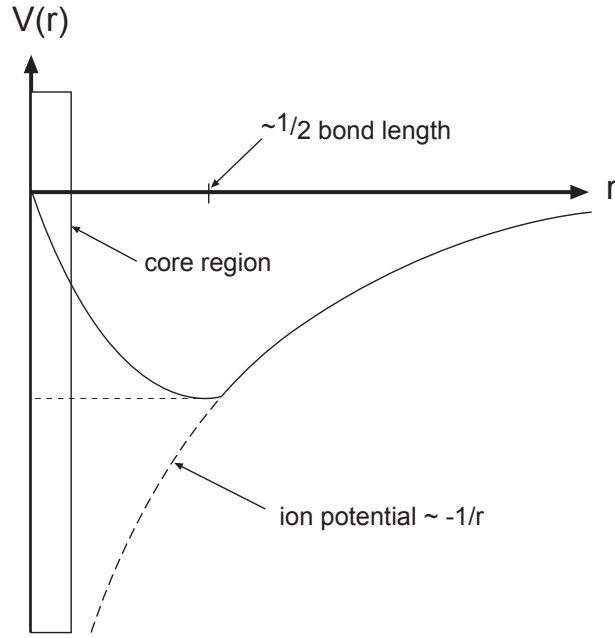
Figure 2.7: Generic atomic pseudopotential in real space

We substitute this into the Schrödinger equation(Eq. 2.13)

$$\{-\frac{\hbar^2}{2m}\nabla^2 + V(r)\}\Phi_{n\mathbf{k}}(r) = E\Phi_{n\mathbf{k}}(r) \tag{2.19}$$

after expanding we obtain

$$\frac{\hbar^2}{2m}|\mathbf{G}_i + \mathbf{k}|^2 \frac{1}{\sqrt{V}}\sum_{\mathbf{G}_i} a_{\mathbf{G}_i}e^{i(\mathbf{G}+\mathbf{k})r} + V(r)\frac{1}{\sqrt{V}}\sum_{\mathbf{G}_i} a_{\mathbf{G}_i}e^{i(\mathbf{G}+\mathbf{k})r} = \frac{1}{\sqrt{V}}\sum_{\mathbf{G}_i} a_{\mathbf{G}_i}e^{i(\mathbf{G}+\mathbf{k})r} \tag{2.20}$$

multiply $\frac{1}{\sqrt{V}}e^{-i(\mathbf{G}+\mathbf{k})r}$, and integrate over the unit cell

$$\frac{\hbar^2}{2m}|\mathbf{G}_i + \mathbf{k}|^2\sum_{\mathbf{G}_i} a_{\mathbf{G}_i}\frac{1}{V}\int e^{i(\mathbf{G}_i-\mathbf{G}_j)r}dr + \sum_{\mathbf{G}_i} a_{\mathbf{G}_i}\frac{1}{V}\int V(r)e^{i(\mathbf{G}_i-\mathbf{G}_j)r}dr \tag{2.21}$$

$$= E\sum_{\mathbf{G}_i} a_{\mathbf{G}_i}\frac{1}{V}\int e^{i(\mathbf{G}_i-\mathbf{G}_j)r}dr$$

with the normalization rule

$$\frac{1}{V}\int e^{i(\mathbf{G}_i-\mathbf{G}_j)r}dr = \frac{1}{\Omega}\int_\Omega e^{i(\mathbf{G}_i-\mathbf{G}_j)r} = \delta_{ij} \tag{2.22}$$

finally we obtain

$$\{\frac{\hbar^2}{2m}|\mathbf{G}_i + \mathbf{k}|^2 + \sum_{\mathbf{G}_i} V(\mathbf{G}_i - \mathbf{G}_j)\}a_{\mathbf{G}_j} = E_n a_{\mathbf{G}_j} \tag{2.23}$$

$$\sum_j \{(\frac{\hbar^2}{2m}|\mathbf{G}_i + \mathbf{k}|^2 - E)\delta_{ij} + V(\mathbf{G}_i - \mathbf{G}_j)\}a_{\mathbf{G}_j} = 0$$

with

$$V(\mathbf{G}_i - \mathbf{G}_j) = \frac{1}{\Omega} \int_\Omega e^{-i(\mathbf{G}_i - \mathbf{G}_j)r} dr \tag{2.24}$$

Here $V(\mathbf{G}_i - \mathbf{G}_j)$ is the Fourier transform of unknown potential corresponding to a specific crystal system we are interested in. Now within the scheme of EPM, we firstly need to determine the specific form of the potential matrix element. For a generic material with N atoms in the primitive cell, we can write the potential inside the crystal as the superposition of the potential V_α of the single atoms α in position d_α:

$$V(r) = \sum_\alpha V_\alpha(r - d_\alpha). \tag{2.25}$$

The matrix element of the potential is obtained from its Fourier transform in the real space

$$V(q) = \frac{1}{\Omega} \int_\Omega V(r) e^{iq \cdot r} dr \tag{2.26}$$

where q is a reciprocal lattice vector and Ω the volume of the primitive lattice cell. Substituting (2.25) in (2.26) we get

$$V(q) = \frac{1}{\Omega} \int_\Omega \sum_\alpha V_\alpha(r - d_\alpha) e^{iq \cdot r} dr, \tag{2.27}$$

from which

$$V(q) = \sum_\alpha e^{iq d_\alpha} \frac{1}{\Omega} \int_\Omega V_\alpha(r) e^{iq \cdot r} dr. \tag{2.28}$$

We can therefore write the effective potential in the momentum space as

$$V(q) = \sum_\alpha S_\alpha(q) V_\alpha(q). \tag{2.29}$$

The term

$$S_\alpha(q) = \frac{1}{N} e^{iq \cdot d_\alpha}$$

is the structure factor for the α atom, where N is the number of unit cells, and

$$V_\alpha(q) = \frac{N}{\Omega} \int_\Omega V_\alpha(r) e^{iq \cdot r} dr$$

is the Fourier transform of the effective potential of the α atom. When $q = \mathbf{G}$, $V(q)$ is called form factor of the atom (Fig. 2.8).

For a binary compound we can define symmetric and antisymmetric form factor (V^S and V^A) and structure factor (S^S and S^A). The potential is thus defined as

$$V(\mathbf{G}) = S^S(\mathbf{G}) V^S(\mathbf{G}) + S^A(\mathbf{G}) V^A(\mathbf{G}) \tag{2.30}$$

13
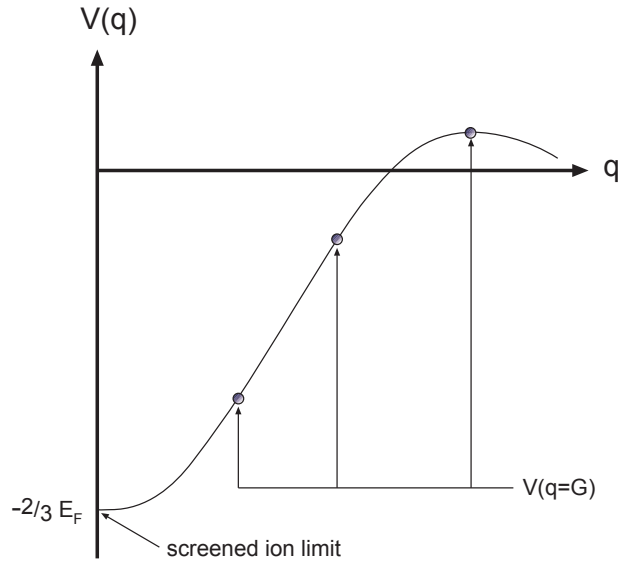
Figure 2.8: Pseudopotential in momentum space

or, in real space,

$$V(r) = \sum_{\mathbf{G}}[S^S(\mathbf{G})V^S(\mathbf{G}) + iS^A(\mathbf{G})V^A(\mathbf{G})]e^{i\mathbf{G}\cdot r}. \tag{2.31}$$

V^S and V^A are calculated as

$$V^S(\mathbf{G}) = (V_{\text{atom a}}(\mathbf{G}) + V_{\text{atom b}}(\mathbf{G}))/2 \tag{2.32a}$$

$$V^A(\mathbf{G}) = (V_{\text{atom a}}(\mathbf{G}) - V_{\text{atom b}}(\mathbf{G})/2, \tag{2.32b}$$

while S^S and S^A depend on the crystal structure. For zincblende materials they are

$$S^S = \cos\left[\pi\left(\frac{l}{4} + \frac{m}{4} + \frac{n}{4}\right)\right] \tag{2.33a}$$

$$S^A = \sin\left[\pi\left(\frac{l}{4} + \frac{m}{4} + \frac{n}{4}\right)\right], \tag{2.33b}$$

while for wurtzite they are

$$S^S = \cos\left[2\pi\left(\frac{l}{6} + \frac{m}{6} + \frac{n}{4}\right)\right]\cos(2\pi nu/2) \tag{2.34a}$$

$$S^A = \cos\left[2\pi\left(\frac{l}{6} + \frac{m}{6} + \frac{n}{4}\right)\right]\sin(2\pi nu/2), \tag{2.34b}$$

where u is the internal parameter and l, m and n are three integers (they are not related to the angular momentum).

An EPM calculation proceeds as follows [5, 6]:

- starting $V(\mathbf{G})$s for elements involved are chosen;

14

- the structure is included via $S(\mathbf{G})$;

- Schrödinger equation is solved, and energies and masses are compared to known values;

- if good agreement is not achieved, $V(\mathbf{G})$s are altered;

- the process is repeated until satisfactory agreement is obtained.

Based on the spatial dependence of $V(r)$, we can have two different EPM contributes:

- **local**, in which the potential is assumed to depend only on radial coordinates $V(r) = V(|r|)$, so we are neglecting the angular momentum of the atoms;

- **nonlocal**, where the potential depends on all the spatial coordinates $V(r) = R(r)Y_{l,m}(\theta, \phi)$.

Choice for one out of the two contributions depends on the considered atomic system: the cancellation theorem reveals that, in order to have an effective cancellation of the nuclei potential, the local potential must have the same symmetry of the valence electron potential. If this assumption is not true, a nonlocal correction must be introduced. In particular, if the cores do not contain electrons of a certain angular momentum involved in the sum, there is no repulsive potential for that component. As an example, the core of carbon is $1s^2$, and therefore carbon will not have a p-repulsive potential. To model the local pseudopotential several approaches have been proposed (e.g. in [7], [8] and [9]). For our calculations we have used the analytic expression presented in [10]. A nonlocal contribution for atomic species α is used to account for the electrons with different angular momentum l. For example, in oxygen the electrons in the $2s$ shell feel attractive potential weaker than $2p$ electrons. An analogous effect can be found in d electrons of metals, in which the l component of order 0 and 1 is weaker than component 2. The nonlocal contribute for our calculations is taken from the work of Chelikowsky and Cohen [11]. Unlike the local case, in the nonlocal calculation it is not possible to rely only on a few parameters, but a potential model must be introduced, the detail is beyond the scope of this thesis and we briefly show the result for full band electronic structure for wurtzite GaN here, see Fig. 2.9.
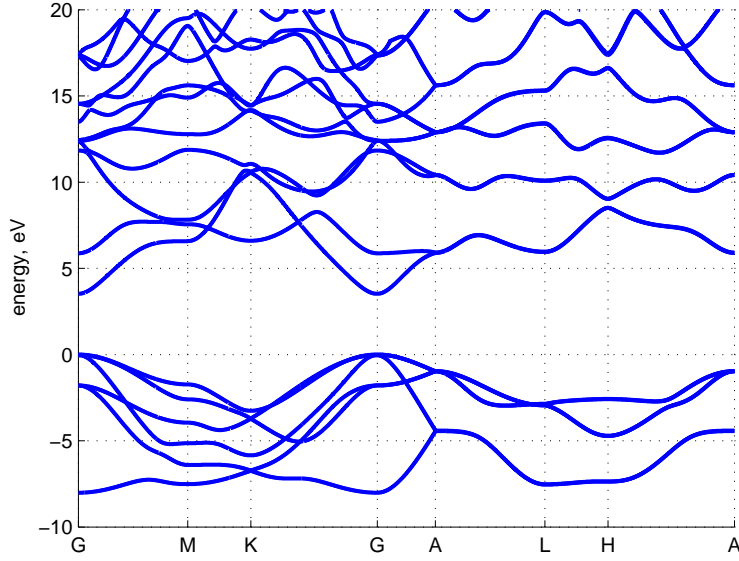
Figure 2.9: Schematic plot for full band dispersion relation of wurtzite GaN

2.3 First principle derivation of $\mathbf{k} \cdot \mathbf{p}$ from NL-EPM model

Previous section gives a brief introduction to EPM, now we try to develop a multiband $\mathbf{k} \cdot \mathbf{p}$ model that is deeply based on EPM[12]. Multiband $\mathbf{k} \cdot \mathbf{p}$ envelope function (EFA) models continue to play a key role in the design of III-nitride optoelectronic devices owing to their fair compromise between accuracy and computational cost, the size and complexity of technologically relevant structures[13, 14] currently exceeding the capabilities of first-principles electronic-structure tools. An unwelcome feature of the multiband EFA method is the appearance of spurious, unphysical solutions of the Schrödinger equation playing havoc with subband dispersions. The exact envelope function theory developed by Burt [15] and Foreman [16] clarified that the proper ordering of the wavenumber operators is a critical ingredient for the stability of the envelope equations. Mireles and Ulloa [17, 18] derived an ordered valence-band Hamiltonian for wurtzite nanostructures, showing that the (formally incorrect) symmetrized Hamiltonian is not appropriate for structures with relevant discontinuities of material parameters at interfaces.

In addition to operator ordering issues, an ill-advised choice of band parameters has also been blamed for the appearance of spurious solutions within the EFA formalism. Focusing on direct band gap zinc-blende heterostructures, Veprek and coworkers[19] noted that 8×8 models based on experimental $\mathbf{k} \cdot \mathbf{p}$ parameters generally produce spurious solutions even with Burt-Foreman (BF) ordering. For the III-nitride system, many of

the key band parameters have not been conclusively determined yet, despite the extensive research effort.[20, 21, 22] In a comprehensive review,[20] Vurgaftman and Meyer summarized the field of III-nitrides and recommended up-to-date band parameters for all common compounds and their alloys by combining different values from the most reliable experimental and theoretical sources. With the aim of providing consistent sets of band parameters for AlN, GaN, and InN, Rinke et al.[22] revisited this work employing many-body perturbation theory in the G_0W_0 approximation. Few notable exceptions aside,[23, 24] $\mathbf{k} \cdot \mathbf{p}$ parameters are obtained from electronic structure calculations or experimental data through fitting procedures. Due to the lower symmetry of wurtzite crystals, which implies a larger number of parameters compared to zinc-blende semiconductors, potential inconsistencies may arise from such fitting approaches, since different parameter sets may produce a very similar fit to the target crystal electronic structure.

The evaluation of Auger transitions requires the use of fully coupled 8×8 $\mathbf{k} \cdot \mathbf{p}$ models instead of the more common 6×6 formulations.[25, 26, 27, 28] More generally, the non-parabolicity of the conduction band, arising from the coupling between conduction and valence bands, should be included in carrier transport[29] and optical spectra[30] calculations. An often overlooked issue in 8×8 models is related to the replacement of the fitted parameters by their corresponding versions after Löwdin renormalization.[31] Improper renormalization arising from inconsistencies in the band parameters may lead to pathological features of the electronic structure. As a notable example of the subtleties of the renormalization procedure, we mention that the widely used InN band parameters recommended by Vurgaftman and Meyer[20] were derived from inconsistent sources (valence band parameters were obtained from early EPM calculations that assumed a gap around $2.0\,\mathrm{eV}$,[9] while the conduction band effective mass was reassessed according to more recent references that predicted a much narrower gap in the range of $0.7 - 0.8\,\mathrm{eV}$) and their use results in incorrect band bending after renormalization.

Many strategies for eliminating spurious solutions have been proposed[32, 33, 34, 35, 36, 19, 37, 38, 39] but none has yet found wide acceptance. Although the ellipticity analysis proposed by Veprek et al.[19, 37, 38] provides convincing stability criteria for 6×6 models, its extension to the 8×8 case is arguable since the intrinsic non-convexity of the associated bilinear form requires a decoupled analysis for conduction and valence Hamiltonians. It is interesting to notice that the Schrödinger equation is related (by a Wick rotation in Minkowski space) to the advection-diffusion-reaction problem,[40] which has well known numerical instabilities when the advection term is dominant.[41] Therefore it is not surprising that upwind approaches[42] originally proposed for the stabilization of

advection-dominated problems have striking similarities with numerical techniques independently developed to eliminate spurious solutions from band structure calculations. A close inspection of the discretized equations reveals that the implementation of BF operator ordering within a finite element formulation of the multiband $\mathbf{k} \cdot \mathbf{p}$ EFA equations corresponds to the upwind discretization of the advection-diffusion equation, whereas a central difference treatment of the latter, equivalent to symmetric operator ordering, leads to spurious oscillations in advection dominated regimes.

With a view of the above remarks, we argue that numerical stability may be simply achieved by deriving both operator ordering and material parameters from first-principles, without resorting to artificial diffusion[34, 35] or other arbitrary stabilization techniques.[33, 36] Therefore, we propose a rigorous numerical procedure to obtain a unique set of band parameters from wave functions and energy eigenvalues computed at the center of the Brillouin zone with a nonlocal empirical pseudopotential technique (NL-EPM).[10] Some complications arise from the nonlocal nature of the pseudopotentials involved, as conventional $\mathbf{k} \cdot \mathbf{p}$ formulations for wurtzite crystals assume the commutability of the potential operator with the coordinate operator. In treating nonlocal corrections to static and frequency-dependent dielectric response functions within density functional theory (DFT), Refs. [43, 44, 45] reported that the velocity operator

$$\hat{v} = \frac{i}{\hbar}[\hat{H}, \hat{r}] \tag{2.35}$$

is no longer equivalent to the commutator of the kinetic energy with the coordinates, i.e. the momentum operator, an additional commutator that arises from the nonlocal potential entering the formalism. Despite the relevance of nonlocal potentials in modern electronic structure calculations, the issue has rarely been addressed in $\mathbf{k} \cdot \mathbf{p}$ perturbation theories.[46, 23, 47] The $\mathbf{k} \cdot \mathbf{p}$ approach described in the following can be considered an extension of the model proposed by Chuang and Chang[48] for bulk wurtzite crystals with amendements to account for operator ordering and nonlocal potentials. Having separated local and nonlocal components of the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{H}_{loc} + \hat{H}_{nl}, \tag{2.36}$$

the relevant commutators displaying its wavevector dependence are

$$\frac{\partial \hat{H}}{\partial k} = [\hat{H}, i\hat{r}] = \frac{\hbar}{m_0}\hat{p} + [\hat{H}_{nl}, i\hat{r}] \tag{2.37}$$

$$\frac{1}{2}\frac{\partial^2 \hat{H}}{\partial k^2} = \frac{1}{2}[[\hat{H}, i\hat{r}], i\hat{r}] = \frac{\hbar^2}{2m_0} + \frac{1}{2}[[\hat{H}_{nl}, i\hat{r}], i\hat{r}]. \tag{2.38}$$

The resulting $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian takes the following form

$$H_{k \cdot p} = -\frac{\hbar^2}{2m}\nabla^2 + \frac{\hbar}{m_0}k \cdot p + \frac{\hbar^2 k^2}{2m_0} + H_{loc}$$
$$+ k \cdot [\hat{H}_{nl}, i\hat{r}] + \frac{1}{2}k^2[[\hat{H}_{nl}, i\hat{r}], i\hat{r}] \tag{2.39}$$

which reduces to a well known expression (see e.g. Eq. (1) in Ref. [48]) for purely local Hamiltonians. Dividing band-edge functions into weakly interacting classes $\mathcal{A}$ and $\mathcal{B}$, with class $\mathcal{A}$ including the states

$$|1\rangle = |iS \uparrow\rangle, |2\rangle = |iS \downarrow\rangle,$$

$$|3\rangle = \left|-\frac{1}{\sqrt{2}}(X + iY) \uparrow\right\rangle, |4\rangle = \left|\frac{1}{\sqrt{2}}(X - iY) \uparrow\right\rangle,$$

$$|5\rangle = |Z \uparrow\rangle, |6\rangle = \left|\frac{1}{\sqrt{2}}(X - iY) \downarrow\right\rangle$$

$$|7\rangle = \left|-\frac{1}{\sqrt{2}}(X + iY) \downarrow\right\rangle, |8\rangle = |Z \downarrow\rangle, \tag{2.40}$$

and class $\mathcal{B}$ all the other bands (the remote bands), quasi-degenerate perturbation theory (see Ref. [49]) leads to the following eight-band ordered Hamiltonian

$$H_{\mathbf{k} \cdot \mathbf{p}}^{8 \times 8} = \begin{bmatrix} E_c & 0 & -\frac{k_+ P_2}{\sqrt{2}} & \frac{k_- P_2}{\sqrt{2}} & k_z P_1 & 0 & 0 & 0 \\ 0 & E_c & 0 & 0 & 0 & \frac{k_- P_2}{\sqrt{2}} & -\frac{k_+ P_2}{\sqrt{2}} & k_z P_1 \\ -\frac{P_2 k_-}{\sqrt{2}} & 0 & F & -(K^\dagger)^* & -(H_+^\dagger)^* & 0 & 0 & 0 \\ \frac{P_2 k_+}{\sqrt{2}} & 0 & -K^* & G & H_-^\dagger & 0 & 0 & \sqrt{2}\Delta_3 \\ P_1 k_z & 0 & -H_+^* & H_- & \Lambda & 0 & \sqrt{2}\Delta_3 & 0 \\ 0 & \frac{P_2 k_+}{\sqrt{2}} & 0 & 0 & 0 & F^* & -K^* & H_-^\dagger \\ 0 & -\frac{P_2 k_-}{\sqrt{2}} & 0 & 0 & \sqrt{2}\Delta_3 & -(K^\dagger)^* & G^* & -(H_+^\dagger)^* \\ 0 & P_1 k_z & 0 & \sqrt{2}\Delta_3 & 0 & H_- & -H_+^* & \Lambda \end{bmatrix} \tag{2.41}$$

where

$$k_\pm = k_x \pm i k_y,$$

$$E_c = E_g + \Delta_1 + \Delta_2$$
$$+ \frac{\hbar^2}{2m}[k_x A_t k_x + k_y A_t k_y + k_z A_z k_z],$$

$$\Lambda = \frac{\hbar^2}{2m}[k_z A_1 k_z + k_x A_2 k_x + k_y A_2 k_y],$$

$$\Theta = \frac{\hbar^2}{2m}[k_z A_3 k_z + k_x A_4 k_x + k_y A_4 k_y],$$

$$F = \Delta_1 + \Delta_2 + \Lambda + \Theta$$
$$+ \frac{\hbar^2}{2m}[-ik_y(A_5^+ - A_5^-)k_x + ik_x(A_5^+ - A_5^-)k_y],$$

19

$$G = \Delta_1 - \Delta_2 + \Lambda + \Theta$$
$$+ \frac{\hbar^2}{2m}[-ik_x(A_5^+ - A_5^-)k_y + ik_y(A_5^+ - A_5^-)k_x],$$
$$K = \frac{\hbar^2}{2m}[k_x A_5 k_x - i(k_x A_5 k_y + k_x A_5 k_y) - k_y A_5 k_y],$$
$$H_\pm = \frac{\hbar^2}{2m}[k_z A_6^+(k_x - ik_y) + (k_x - ik_y)A_6^- k_z]$$
$$\mp i A_7(k_x - ik_y). \tag{2.42}$$

In the expressions above, $\Delta_1 = \Delta_{cr}$ is the crystal field splitting energy and $\Delta_2 = \Delta_3 = \Delta_{so}/3$ where Δ_{so} is the spin-orbit splitting energy. Although the numerical results reported below are for quantum wells with the confining direction along the z-axis, the general form of the Hamiltonian defined by Eqs. (2.41)-(2.42) allows the treatment of quantum systems of any dimensionality. Notice that the dagger notation flips the ordering of the terms, e.g. $(k_m A k_n)^\dagger = k_n A^* k_m$. The $\mathbf{k} \cdot \mathbf{p}$ parameters introduced above are rigorously derived from NL-EPM calculations as

$$P_1^2 = \left| \frac{\hbar}{m_0} \left\langle iS \left| \hat{p}_z + \frac{m_0}{\hbar}[\hat{H}_{nl}, i\hat{r}_z] \right| Z \right\rangle \right|^2,$$

$$P_2^2 = \left| \frac{\hbar}{m_0} \left\langle iS \left| \hat{p}_x + \frac{m_0}{\hbar}[\hat{H}_{nl}, i\hat{r}_x] \right| X \right\rangle \right|^2,$$

$$A_{t,z} = \frac{2m_0}{\hbar^2}C_{t,z}, \quad A_1 = \frac{2m_0}{\hbar^2}L_2, \quad A_2 = \frac{2m_0}{\hbar^2}M_3,$$

$$A_3 = \frac{2m_0}{\hbar^2}(M_2 - L_2), \quad A_4 = \frac{2m_0}{\hbar^2}\left(\frac{L_1 + M_1}{2} - M_3\right),$$

$$A_5^\pm = \frac{2m_0}{\hbar^2}\frac{N_1^\pm}{2}, \quad A_6^\pm = \frac{2m_0}{\hbar^2}\frac{N_2^\pm}{\sqrt{2}},$$

$$A_7 = \frac{-i\hbar}{m_0\sqrt{2}}\left\langle X \left| \hat{p}_x + \frac{m_0}{\hbar}[\hat{H}_{nl}, i\hat{r}_x] \right| Z \right\rangle, \tag{2.43}$$

where $A_5 = A_5^+ + A_5^-$, $A_6 = A_6^+ + A_6^-$, and the Luttinger-Kohn parameters are defined as

$$C_{t,z} = \frac{\hbar^2}{2m}\left(1 + \sum_{\gamma \subset \mathcal{B}} \frac{2\hat{P}_{S\gamma}^{t,z}\hat{P}_{\gamma S}^{t,z}}{m_0(E_S - E_\gamma)}\right),$$

$$L_1 = \frac{\hbar^2}{2m}\left(1 + \left\langle X \left| \frac{m_0}{\hbar^2}[[\hat{H}_{nl}, i\hat{r}_x], i\hat{r}_x] \right| X \right\rangle + \sum_{\gamma \subset \mathcal{B}} \frac{2\hat{P}_{X\gamma}^x\hat{P}_{\gamma X}^x}{m_0(E_X - E_\gamma)}\right),$$

$$L_2 = \frac{\hbar^2}{2m}\left(1 + \left\langle Z \left| \frac{m_0}{\hbar^2}[[\hat{H}_{nl}, i\hat{r}_z], i\hat{r}_z] \right| Z \right\rangle + \sum_{\gamma \subset \mathcal{B}} \frac{2\hat{P}_{Z\gamma}^z\hat{P}_{\gamma Z}^z}{m_0(E_Z - E_\gamma)}\right),$$

$$M_1 = \frac{\hbar^2}{2m}\left(1 + \left\langle X \left| \frac{m_0}{\hbar^2}[[\hat{H}_{nl}, i\hat{r}_y], i\hat{r}_y] \right| X \right\rangle + \sum_{\gamma \subset \mathcal{B}} \frac{2\hat{P}_{X\gamma}^y\hat{P}_{\gamma X}^y}{m_0(E_X - E_\gamma)}\right),$$

$$M_2 = \frac{\hbar^2}{2m}\left(1 + \left\langle X \left| \frac{m_0}{\hbar^2}[[\hat{H}_{nl}, i\hat{r}_z], i\hat{r}_z] \right| X \right\rangle + \sum_{\gamma \subset \mathcal{B}} \frac{2\hat{P}_{X\gamma}^z\hat{P}_{\gamma X}^z}{m_0(E_X - E_\gamma)}\right),$$

$$M_3 = \frac{\hbar^2}{2m}\left(1 + \left\langle Z\left|\frac{m_0}{\hbar^2}[[\hat{H}_{nl}, i\hat{r}_x], i\hat{r}_x]\right|Z\right\rangle + \sum_{\gamma \subset \mathcal{B}}\frac{2\hat{P}^x_{Z\gamma}\hat{P}^x_{\gamma Z}}{m_0(E_Z - E_\gamma)}\right),$$

$$N_1^+ = \frac{\hbar^2}{2m}\sum_{\gamma \subset \mathcal{B}}\frac{2\hat{P}^x_{X\gamma}\hat{P}^y_{\gamma Y}}{m_0(E_X - E_\gamma)} + \frac{1}{2}\left\langle X\left|[[\hat{H}_{nl}, i\hat{r}_x], i\hat{r}_y]\right|Y\right\rangle$$

$$N_1^- = \frac{\hbar^2}{2m}\sum_{\gamma \subset \mathcal{B}}\frac{2\hat{P}^y_{X\gamma}\hat{P}^x_{\gamma Y}}{m_0(E_X - E_\gamma)} + \frac{1}{2}\left\langle X\left|[[\hat{H}_{nl}, i\hat{r}_y], i\hat{r}_x]\right|Y\right\rangle$$

$$N_2^+ = \frac{\hbar^2}{2m}\sum_{\gamma \subset \mathcal{B}}\hat{P}^x_{X\gamma}\hat{P}^z_{\gamma Z}\left(\frac{1}{m_0(E_X - E_\gamma)} + \frac{1}{m_0(E_Z - E_\gamma)}\right) + \frac{1}{2}\left\langle X\left|[[\hat{H}_{nl}, i\hat{r}_x], i\hat{r}_z]\right|Z\right\rangle$$

$$N_2^- = \frac{\hbar^2}{2m}\sum_{\gamma \subset \mathcal{B}}\hat{P}^z_{X\gamma}\hat{P}^x_{\gamma Z}\left(\frac{1}{m_0(E_X - E_\gamma)} + \frac{1}{m_0(E_Z - E_\gamma)}\right) + \frac{1}{2}\left\langle X\left|[[\hat{H}_{nl}, i\hat{r}_z], i\hat{r}_x]\right|Z\right\rangle$$

$$\hat{P}^\alpha_{X\gamma} = \left\langle X\left|\hat{p}_\alpha + \frac{m_0}{\hbar}[\hat{H}_{nl}, i\hat{r}_\alpha]\right|\gamma\right\rangle \tag{2.44}$$

and $\hat{p}_\alpha = -i\hbar\partial/\partial\alpha$, with $\alpha = x, y, z$. We remark that Eq. (2.43) includes a rigorous definition of the coupling parameters P_1, P_2 taking into account nonlocal effects, beyond the approximations commonly adopted in the literature (Ref. [48], Eq. (18)) which may lead to inconsistencies in parameter renormalization techniques. The equations above simplify to the $\mathbf{k} \cdot \mathbf{p}$ theory developed in Ref. [48] for wurtzite crystals when nonlocal potentials are neglected, with only minor modifications due to the more rigorous quasi-degenerate perturbation theory adopted here, see e.g. the expression for $N_2^\pm$ in Eq. (2.44) and the corresponding definition in Ref. [24], Eq. (5). Nonlocal corrections can be conveniently evaluated in reciprocal space as in Ref. [45]

$$\left\langle\mathbf{K}\left|[\hat{H}_{nl}, i\hat{r}]\right|\mathbf{K}'\right\rangle = (\nabla_\mathbf{K} + \nabla_{\mathbf{K}'})H_{nl}(\mathbf{K}, \mathbf{K}'), \tag{2.45}$$

where $H_{nl}(\mathbf{K}, \mathbf{K}') = \left\langle\mathbf{K}\left|\hat{H}_{nl}\right|\mathbf{K}'\right\rangle$ are the matrix elements beween plane waves required by NL-EPM calculations, and $\mathbf{K} = \mathbf{k} + \mathbf{G}$, $\mathbf{K}' = \mathbf{k} + \mathbf{G}'$. Deformation potentials, as formulated by Bir and Pikus,[50] can also be directly derived from nonlocal pseudopotential calculations,[23] although this approach will not be pursued in this thesis.

Fig. 2.10 compares the electronic structure of wurtzite GaN and AlN computed with NL-EPM and the present NL-EPM-derived $\mathbf{k} \cdot \mathbf{p}$ model. An excellent agreement near the Brillouin zone center can be observed. As accurate $\mathbf{k} \cdot \mathbf{p}$ bands have to be built upon a nonlocal full-Brillouin-zone description, be it DFT or NL-EPM, a local approximation of the Luttinger-Kohn parameters listed in Eq. (2.44) may lead to inconsistent results. In order to assess the role of nonlocality in the numerical procedure described above, we have compared NL-EPM-derived $\mathbf{k} \cdot \mathbf{p}$ bands obtained with and without nonlocal corrections; Fig. 2.10 clearly shows that, if band parameters are to be extracted from nonlocal descriptions, the additional terms involving $\hat{H}_{nl}$ in Eq. (2.44) should not be neglected.
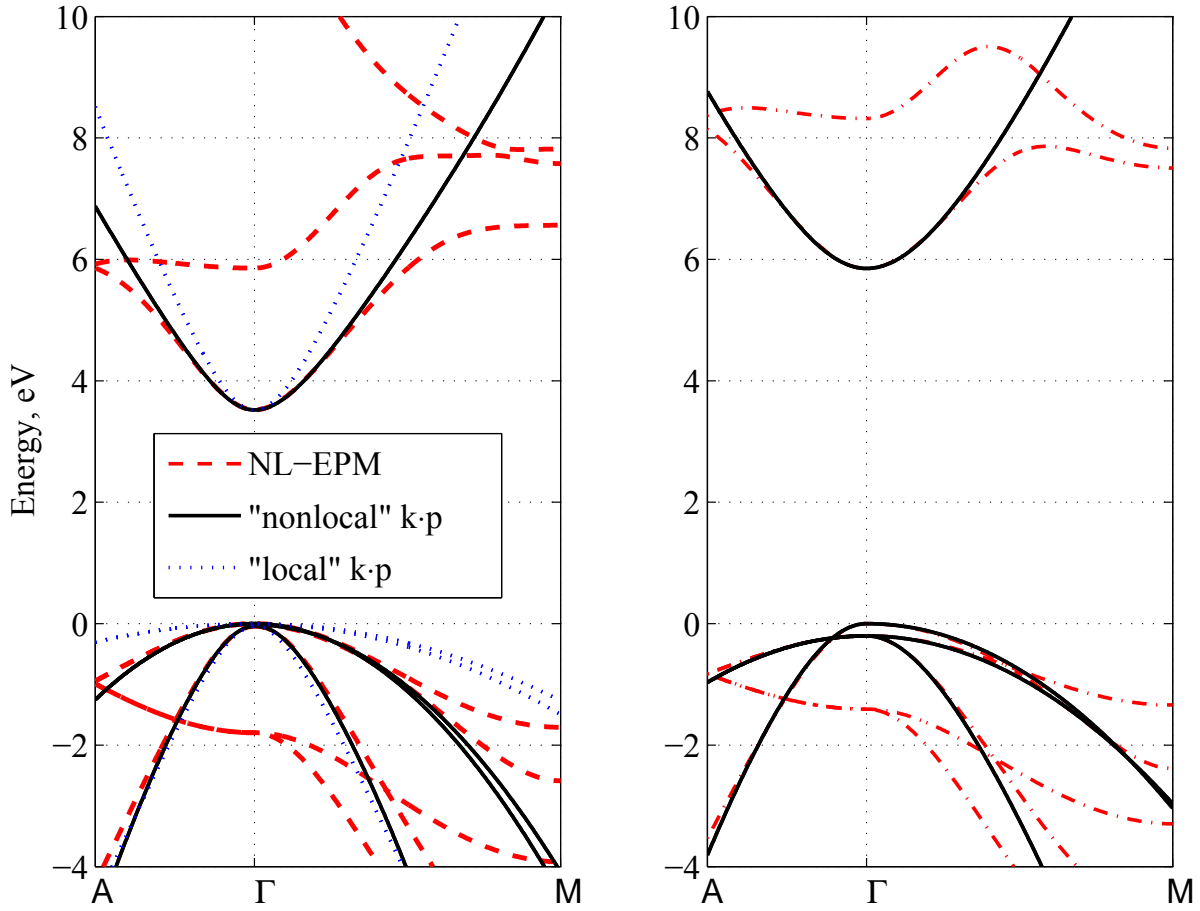
Figure 2.10: Band structure of GaN (left) and AlN (right), computed with NL-EPM (dashed red lines) and the present $\mathbf{k} \cdot \mathbf{p}$ model with parameters derived from the corresponding NL-EPM bands (solid black lines). The $\mathbf{k} \cdot \mathbf{p}$ bands obtained for GaN in the local approximation are reported as dotted blue lines.

Bulk dispersions obtained from other $\mathbf{k} \cdot \mathbf{p}$ parametrizations are shown in Fig. 2.11. A complete listing of band parameters calculated with the present technique and with previous approaches [22, 20, 24] is presented in Table 2.1. The inverse mass parameters A_i recommended by Vurgaftman and Meyer[20] for GaN and AlN were fitted on empirical pseudopotential calculations[51, 52] and on DFT calculations within the local density approximation,[53] respectively. The parameters proposed by Dugdale et al.[24] were derived from pseudopotential calculations,[9] while Rinke et al.[22] determined their parameters from a fitting to G_0W_0 DFT calculations. The scatter of $\mathbf{k} \cdot \mathbf{p}$ parameters reported in the literature mainly arises from the different levels of approximation by which the original bulk dispersions were obtained. While a consensus over the main large-scale features of III-nitride electronic structures seems to be gradually emerging among DFT practitioners[54, 55, 56] (some of these features being related to relevant carrier

Table 2.1: Band parameters for wurtzite GaN and AlN. The 8×8 parameters obtained in the present approach have been renormalized for comparison with the 6×6 parameters reported in the literature.

|	GaN				AlN			
	present work	Ref. [22]	Ref. [20]	Ref. [24]	present work	Ref. [22]	Ref. [20]	Ref. [24]
A_1	-6.501	-5.947	-7.21	-7.979	-3.899	-3.991	-3.86	-4.711
A_2	-0.828	-0.528	-0.44	-0.581	-0.616	-0.311	-0.25	-0.476
A_3	5.562	5.414	6.68	7.291	3.325	3.671	3.58	4.176
A_4	-2.29	-2.512	-3.46	-3.289	-1.366	-1.147	-1.32	-1.816
A_5	-2.214	-2.510	-3.40	-3.243	-1.424	-1.329	-1.47	-1.879
A_6	-2.63	-3.202	-4.90	-4.281	-1.684	-1.952	-1.64	-2.355
A_7, eV Å	0.408	0.046	0.0937	0.179	0.138	0.026	0	0.096
E_g, eV	3.52	3.24	3.437		6.05	6.47	6.00	
Δ_{cr}, eV	0.042	0.034	0.010	0.022	-0.202	-0.295	-0.227	-0.093
$m_e^{\parallel}$	0.228	0.209	0.21	0.15	0.34	0.329	0.32	0.25
$m_e^{\perp}$	0.185	0.186	0.20	0.14	0.29	0.322	0.30	0.24
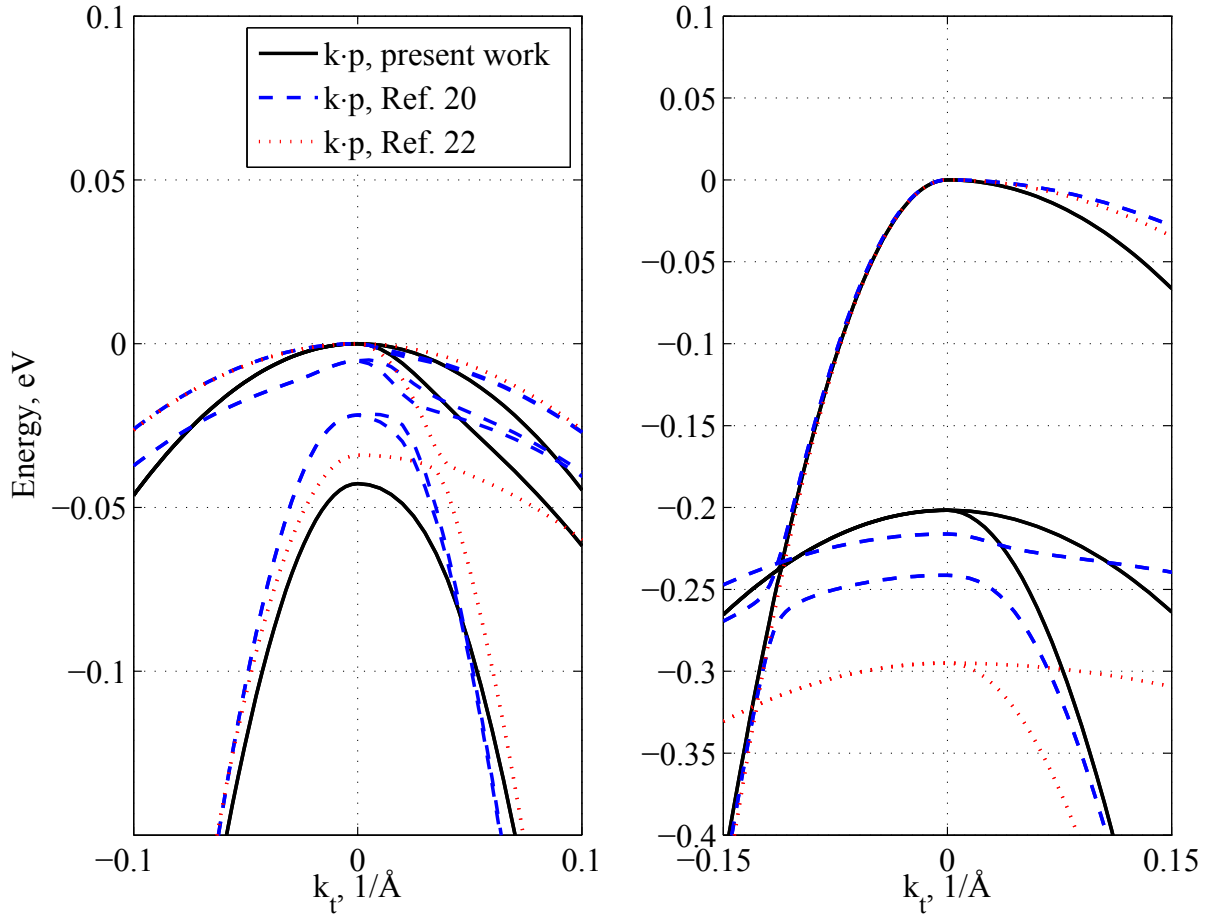
Figure 2.11: Valence band structure of GaN (left) and AlN (right), computed with the present NL-EPM-derived $\mathbf{k} \cdot \mathbf{p}$ model (solid black lines) and with the parameter sets proposed in Ref. [20] (dashed blue lines) and Ref. [22] (dotted red lines).

transport[29, 57] and recombination[26, 28] properties), finer details of the energy dispersions near the band edges are less established. In particular, the crystal-field and spin-orbit splittings Δ_{cr}, Δ_{so} have a strong impact on the resulting subband dispersions. A review of experimental exciton data suggested $\Delta_{cr} = 10\,\text{meV}$ and $\Delta_{so} = 17\,\text{meV}$ in Ref.[20], while G_0W_0 DFT predicted $\Delta_{cr} = 34\,\text{meV}$ in GaN;[22] our NL-EPM optimization[10, 58, 59] was based on several DFT calculations,[60, 61, 62, 63] and the GaN crystal field splitting $\Delta_{cr} = 42\,\text{meV}$ was selected according to Ref.[63] (spin-orbit corrections were neglected there as well as in Ref.[22]). A discussion of the accuracy of these values is beyond the scope of this thesis, the aim here is to provide a rigorous procedure for the extraction of $\mathbf{k} \cdot \mathbf{p}$ parameters eliminating possible inconsistencies in calculations of subband dispersions, optical properties and Auger transitions in quantum-confined nanostructures.

2.4 Finite element implementation of $\mathbf{k} \cdot \mathbf{p}$ for quantum wells

In this thesis, the numerical technique employed to solve the differential problems is the Galerkin method [64, 65], within the framework of Finite Element Method(FEM).
The general problem we will deal with is a coupled partial differential problem. As we will see, the real problem related with multiband $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian is of the form

$$H_{\mathbf{k} \cdot \mathbf{p}} = -\sum_{i,j=x,y,z} h_{ij}^{(2)} k_i k_j + \sum_{i=x,y,z} h_i^{(1)} k_i + h^{(0)} \tag{2.46}$$

when applied to the specific case of quantum well system, assume k_z is the quantized direction, then k_z can be substituted by $-i\partial_z$, resulting in a coupled PDE problem, where $\mathbf{k}_t$ is along the unconfined in-plane direction

$$H_{\mathbf{k} \cdot \mathbf{p}} = -\sum_{i,j} \partial_i h_{ij}^{(2)}(r, \mathbf{k}_t)\partial_j + \sum_i (h_{iL}^{(1)}(r, \mathbf{k}_t)\partial_i + \partial_i h_{iR}^{(1)}(r, \mathbf{k}_t)) + h^{(0)}(r, \mathbf{k}_t) \tag{2.47}$$

We illustrate here first how FEM can be applied to a single band Schrödinger equation in the effective mass approximation, then it could be trivially extended to suit the more complicated $\mathbf{k} \cdot \mathbf{p}$ model(system of coupled PDEs). Based on the Hamiltonian in Eq. 2.13, we write the one-dimensional Schrödinger equation as

$$[-\frac{\hbar^2}{2}\frac{\partial}{\partial x}\frac{1}{m^*}\frac{\partial}{\partial x} + U(x)]\Psi(x) = E\Psi(x) \tag{2.48}$$

then we apply the Galerkin preocedure to discretize Eq. 2.48

$$-\frac{\hbar^2}{2}\int N_i \frac{\partial}{\partial x}\frac{1}{m^*}\frac{\partial}{\partial x}\Psi(x)dx + \int N_i U(x)\Psi(x)dx = E\int N_i \Psi(x)dx \tag{2.49}$$

where $\{N_i(x)\}$ are so called weighting functions, integral by part futher yields

$$\frac{\hbar^2}{2} \int \frac{\partial N_i(x)}{\partial x} \frac{1}{m^*} \frac{\partial \Psi(x)}{\partial x} dx + \int N_i(x) U(x) \psi(x) dx = E \int N_i(x) \Psi(x) dx \qquad (2.50)$$

Now the 1D computational domain is discretized into line elements, in each of which the wavefunction Ψ is expanded in terms of Lagrange polynomials

$$\Psi(x) = \sum_j \psi_j N_j(x) \qquad (2.51)$$

where $\{N_i(x)\}$ here is the testing functions, the same with the weighting functions in Eq. 2.49, substituting Ψ in Eq. 2.50 we have

$$\frac{\hbar^2}{2} \sum_j \psi_j \int \frac{\partial N_i}{\partial x} \frac{1}{m^*} \frac{\partial N_j}{\partial x} dx + \sum_j \psi_j \int U N_i N_j dx = E \sum_j \psi_j \int N_j N_j dx \qquad (2.52)$$

assuming in each element the effective mass and potential are constant, we obtain

$$\frac{\hbar^2}{2m^*} \sum_j \psi_j \int \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} dx + U \sum_j \psi_j \int N_i N_j dx = E \sum_j \psi_j \int N_j N_j dx \qquad (2.53)$$

finnally it can be cast in matrix form

$$\{\frac{\hbar^2}{2m^*}[T]_e + U[\mu]_e\}\{\psi\}_e = E[\mu]_e\{\psi\}_e \qquad (2.54)$$

By summing up the matrix representation for each element, we can construct the generalized eigenvalue problem.

$$\sum_e \{\frac{\hbar^2}{2m^*}[T]_e + U[\mu]_e\}\{\psi\}_e = E \sum_e [\mu]_e\{\psi\}_e \qquad (2.55)$$

where

$$[T]_e = \int_e \frac{\partial \{N\}^T}{\partial x} \frac{\partial \{N\}}{\partial x} dx, \quad in\ unit\ \frac{1}{l} \qquad (2.56)$$

$$[\mu]_e = \int_e \{N\}^T \{N\} dx, \qquad in\ unit\ l \qquad (2.57)$$

once we perform a real implementation, we write Eq. 2.55 as

$$[A]\{\psi\} = E[B]\psi \qquad (2.58)$$

where

$$[A] = \sum_e \{\frac{\hbar^2}{2m^*}[T]_e + U[\mu]_e\} \qquad (2.59)$$

$$[B] = \sum_e [\mu]_e \qquad (2.60)$$

in a simple problem we could use the first-order Lagrange polynomials, hence we need only two basis in each element, do a simple integral over each line element(of length l_e), we will find $[T]_e$ and $[\mu]_e$, more detailed element matrices are documented in Appendix.A

$$[T]_e = \frac{1}{l_e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{2.61}$$

$$[\mu]_e = \frac{l_e}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{2.62}$$

Regarding the case of multiband $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian, We assume the following ansatz for the nanostructure wave function

$$\psi(r,z) = e^{ik\cdot r} \sum_{m=1}^{M} u_{m0}(r,z)\zeta_{km}(z) \tag{2.63}$$

where the coordinates z and r are the symmetry broken and the translational invariant directions, respectively, k represents the transversal crystal momentum, and $u_{m0}(r,z)$ are $\mathbf{k}\cdot\mathbf{p}$ orthonormal lattice-periodic functions. The slowly-varying envelopes $\zeta_{km}(z)$ describe, at every position in the symmetry broken direction z, how the lattice-periodic functions are mixed together. The bands included in the present analysis are the light and heavy holes, the split-off band and the first conduction band ($M = 8$).

By substituting Eq. 2.63 into Eq. 2.47, one obtains the coupled equation system for nanostructure envelopes $\zeta(z) = \{\zeta_1, \zeta_2, ..., \zeta_M\}^T$

$$\left\{ -\sum_{i,j} \partial_i H^{(2)}(k,z)\partial_j - \sum_i \left[H_L^{(1)}(k,z)i\partial_i + i\partial_i H_R^{(1)}(k,z) \right] + H^{(0)}(k,z) \right\} \zeta(z) = E\zeta(z),$$
$$\tag{2.64}$$

parametrized for the wavevector k in the translational invariant space. Discretizing the z axis into N points $(z_1, z_2, ...z_N)$ and representing the envelopes in the finite-element basis $\zeta(z) = \sum_j \zeta_j N_j(z)$, the application of Galerkin procedure leads to the weak form of Eq. (2.64)

$$\sum_e \int_{\Omega_e} \sum_j \zeta_j \partial N_i H^{(2)}(k,z)\partial N_j dz$$

$$-\sum_e \int_{\Omega_e} \sum_j \zeta_j \left(iN_i H_L^{(1)}(k,z)\partial N_j - i\partial N_i H_R^{(1)}(k,z)N_j \right) dz$$

$$+\sum_e \int_{\Omega_e} \sum_j \zeta_j N_i H^{(0)}(k,z)N_j dz = E\sum_e \int_{\Omega_e} \sum_j \zeta_j N_i N_j dz. \tag{2.65}$$

From a theoretical point of view, there is no fundamental difference between solving the above equation system and Eq. 2.53. In practice, the assembling unit becomes a matrix

instead of a scalar(effective mass hamiltonian), therefore leading to a much larger sparse linear system.
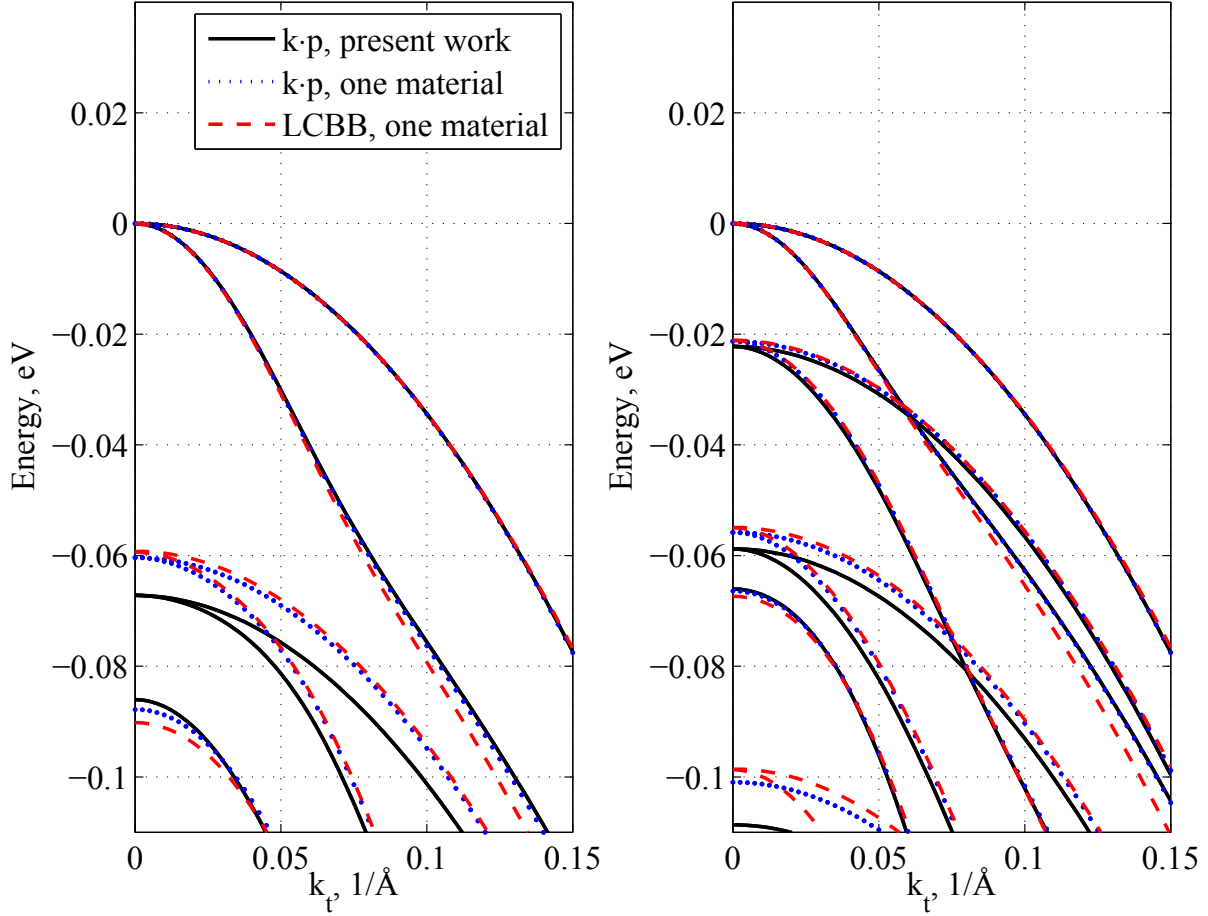
Figure 2.12: Valence-subband structure of $GaN/Al_{0.25}Ga_{0.75}N$ QWs with $30\,\text{Å}$ (left) and $60\,\text{Å}$ (right) well width according to the present $\mathbf{k} \cdot \mathbf{p}$ model (solid black lines), compared with LCBB and $\mathbf{k} \cdot \mathbf{p}$ calculations performed in the one-material approximation (dashed red and dotted blue lines).

We obtain the conduction and valence subband structure after solving the sparse linear system(Eq. 2.65) as a general eigenvalue problem. The valence subband dispersion of unstrained $GaN/Al_{0.25}Ga_{0.75}N$ quantum wells computed with the present NL-EPM-derived $\mathbf{k} \cdot \mathbf{p}$ model is shown in Fig. 2.12 (solid lines) for a well width of $30\,\text{Å}$ (left) and $60\,\text{Å}$ (right). In order to assess the accuracy of the proposed perturbative approach, we compare our results with a full-Brillouin-zone NL-EPM model based on the linear combination of bulk bands (LCBB).[28] Originally proposed in Ref.[66] to investigate the electronic structure of quantum dots, LCBB avoids the decomposition of the wave function into envelope functions by expanding the states of the quantum structure in terms of the full-zone Bloch eigenstates of the constituent bulk crystals. The original LCBB formulation, which assumes that all materials composing the nanostructure share
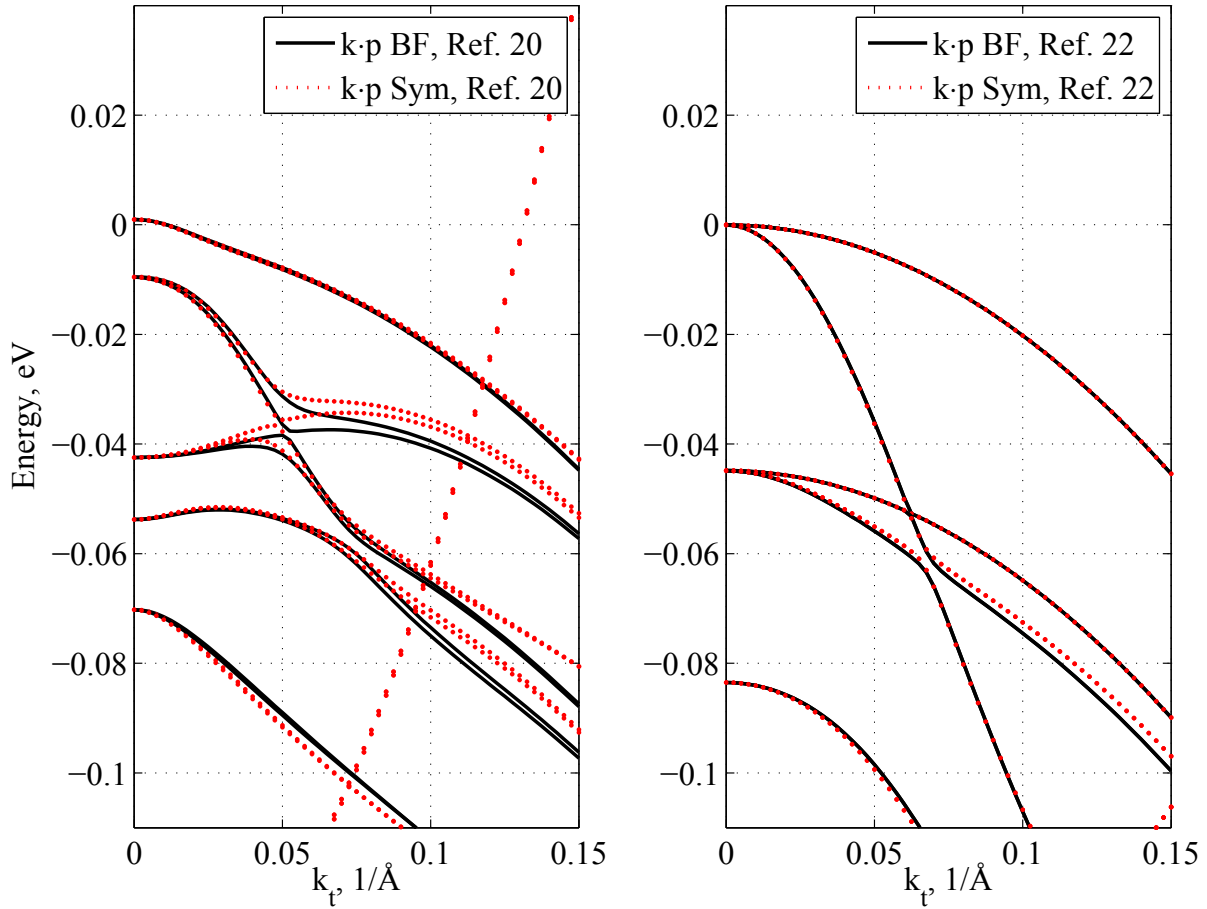
Figure 2.13: (left) Valence-subband structure of the $30\,\text{Å}$-wide $\text{GaN}/\text{Al}_{0.25}\text{Ga}_{0.75}\text{N}$ QW considered in Fig. 2.12(left), computed with the $\mathbf{k} \cdot \mathbf{p}$ parameters from Ref. [20] (left) and Ref. [22] (right) using BF (solid black lines) and symmetric (dotted red lines) ordering.

the same Bravais lattice, was generalized in Ref. [67] to include the description of strain effects. However, the extended formalism is considerably more involved and does not include the treatment of nonlocal potentials. Therefore, we resort here to a one-material approximation,[68, 28] modeling band-edge discontinuities with an appropriate confining potential. Within this approximation, which appears satisfactory for well-confined bound states, the LCBB formulation greatly simplifies, [28] and a comparison between $\mathbf{k} \cdot \mathbf{p}$ calculations (dotted lines) and LCBB results (dashed lines) proves that properly derived envelope function models afford full-zone-quality subband dispersions in a specified region of the Brillouin zone. (This conclusion does not apply in general to all material systems, a notable example being the subband structure of silicon inversion layers, whose full-zone analysis has shown features that escape zone-center theories.[69]) A more detailed comparison to assess the limitations of the present $\mathbf{k} \cdot \mathbf{p}$ approach versus an extended LCBB model accounting e.g. for discontinuities at interfaces, nonlinear strain and other atomistic details will be addressed in a future work, some of these effects being beyond

the reach of envelope function techniques.

The comparison reported in Fig. 2.13 between the present subband dispersions and similar calculations performed adopting the band parameters from Refs. [20, 22] underscores the discrepancies among some of the most widely used descriptions of the GaN valence band near Γ_6^v. The potential effect of these differences in the calculation of radiative and non-radiative properties of III-nitride QWs has not been investigated in detail so far. Additional work is necessary to obtain a consistent set of material parameters from first-principles electronic structure calculations to enable the reliable simulation of III-nitride quantum structures.
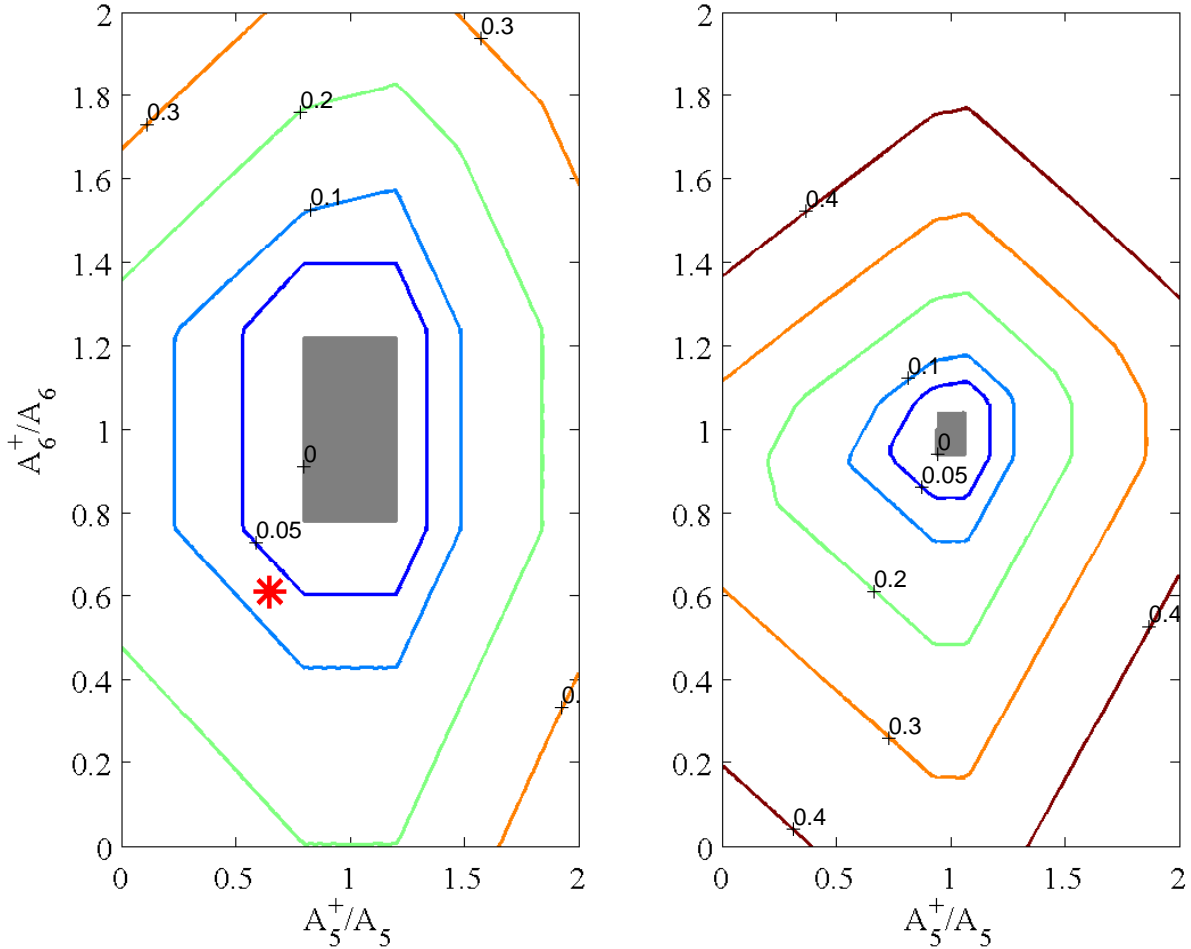
Figure 2.14: Degree of non-ellipticity ρ for different splitting ratios A_5^+/A_5 and A_6^+/A_6 in GaN with $\mathbf{k} \cdot \mathbf{p}$ parameters from the present NL-EPM analysis (left) and from Ref. [20] (right). The shaded regions mark the parameter space where the bilinear form is strictly convex. The star denotes the splitting ratios obtained from the NL-EPM-derived operator ordering.

Additional consideration concerns the numerical stability of the present approach. We have addressed this issue by performing the non-ellipticity analysis proposed in Ref. [38], restricting our attention to the fully quantized limit in a 6×6 model. The degree of

non-ellipticity of the EFA equations, estimated through the ρ ratio defined in Ref. [38], is shown in Fig. 2.14 as a function of the splitting ratios A_5^+/A_5 and A_6^+/A_6 for different $\mathbf{k} \cdot \mathbf{p}$ parameter sets. Shaded regions mark the parameter space where the EFA bilinear form is strictly convex ($\rho = 0$). The splitting ratios resulting from the present NL-EPM analysis define a point (marked by a star in Fig. 2.14, left) close to the convexity region. The use of the $\mathbf{k} \cdot \mathbf{p}$ parameters of Ref. [20] leads to higher non-ellipticity in the whole parameter space (Fig. 2.14, right). Although applicable only to fully quantized systems (quantum dots), the present stability analysis indicates that the NL-EPM-derived EFA equations are well-posed. Inspection of the results obtained for the technologically relevant examples discussed above supports this conclusion. In fact, the subband structure calculated from the band parameters of Ref. [20] is affected by spurious solutions when a symmetrized Hamiltonian is employed, see Fig. 2.13 (dotted lines). On the other hand, symmetric operator ordering ($A_5 = 2A_5^+$, $A_6 = 2A_6^+$) does not lead to spurious solutions or appreciable deviations when $\mathbf{k} \cdot \mathbf{p}$ parameters are directly derived from the NL-EPM bands. We remark that incorrect operator ordering does not lead in general to incorrect results depending on the stability margins of the band parameters; in particular, the equation system resulting from Ref. [20] is elliptic only for nearly complete asymmetric splitting ($A_5 \approx A_5^+$, $A_6 \approx A_6^+$).[38]

Chapter 3

Optical properties in III-nitride nanostructures

3.1 Free-carrier theory and optical gain

Calculations of optical transitions in this chapter are based upon a set of eigen-solutions from Possion-Schrödinger self-consistent solutions. Previous work for III-nitride systems are mostly based on solving $6 \times 6\, \mathbf{k} \cdot \mathbf{p}$ model for the valence band and using effective mass approximation for the conduction band separately[70, 71, 72], while this work pursues a coupled approach using $8 \times 8\, \mathbf{k} \cdot \mathbf{p}$ model.

Additional complications come from the fact that most III-nitride materials are piezo-electric material, i.e. external force(e.g. strain at heterogeneous interfaces) can induce charge or vice versa. This effect results in the piezoelectric polarization ($\mathbf{P}_{pz}$) which is normally induced by a elastic perturbation, while an intrinsic asymmetry of the bonding in the equilibrium crystal structure leads to the spontaneous polarization ($\mathbf{P}_{sp}$). In the absence of an external electric field, the corresponding Maxwells equation becomes(assume polarization $\mathbf{P}$ is time-independent)

$$\nabla \cdot \mathbf{D} = \rho_0 \tag{3.1}$$

$$\mathbf{D} = \varepsilon_b \mathbf{E} + \mathbf{P}, \ \mathbf{P} = \mathbf{P}_{pz} + \mathbf{P}_{sp} \tag{3.2}$$

$$E = -\nabla \phi_{sc} \tag{3.3}$$

where ϕ_{sc} is the self-consistent electro-static potential, ε_b is the dielectric constant and ρ_0

is the charge density. Eq. 3.1-Eq. 3.3 further lead to

$$-\nabla(\varepsilon_b \nabla \phi_{sc}) = -\nabla \cdot \mathbf{P} + \rho_0 \tag{3.4}$$

In this work, numerical model employs spontaneous polarization with linear interpolation[13], and the model for piezoelectric polarization can be represented as,

$$\mathbf{P}_{pz} = 2d_{31}(C_{11} + C_{12} - \frac{2C_{13}^2}{C_{33}})\varepsilon_{xx} \tag{3.5}$$

In 1D case, the charge carrier distributions can be computed from the wave function

$$\rho(z) = |e|(p(z) - n(z)) \tag{3.6}$$

$$n(z) = \sum_c |\psi_c(z)|^2 N_c \tag{3.7}$$

$$p(z) = \sum_v |\psi_v(z)|^2 N_v \tag{3.8}$$

where $p(z)$ and $n(z)$ are integrated envelope function in the longitudinal direction. N_c and N_v can be calculated by integrating the statistical factor over in-plane k_t,

$$N_c = \frac{1}{\mathcal{L}} \int_0^\infty \frac{2\pi k_t}{(2\pi)^2} f_{ck_t} dk_t \tag{3.9}$$

$$N_v = \frac{1}{\mathcal{L}} \int_0^\infty \frac{2\pi k_t}{(2\pi)^2} f_{vk_t} dk_t \tag{3.10}$$

with

$$f_{ck_t} = \frac{1}{1 + e^{[E_c(k_t)-F_c]/k_B t}} \quad , f_{vk_t} = \frac{1}{1 + e^{[F_v-E_v(k_t)]/k_B t}} \tag{3.11}$$

where F_c and F_v are constant levels computed from a given injection level[73] and $\mathcal{L}$ is the volume in the quantized direction. As the k_t dependence of the distribution function cannot be determined analytically, the inversion of the above equation requires a iterative computation such as the Newton procedure. The charge densities $n(z)$ and $p(z)$ are then used as input for the Poisson equation

$$-\nabla(\varepsilon_b \nabla \phi_{sc}) = -\nabla \cdot \mathbf{P_{sp}} - \nabla \cdot \mathbf{P_{pz}} + |e|(p(z) - n(z)) \tag{3.12}$$

The self-consistent loop between Schrödinger and Poisson steps is coupled through a quasi-newton method: the Broyden method(see Appendix.C), which incorporates a numerically efficient rank-1 update. In practice, we find this Broyden update is especially robust for the Possion-Schrödinger type self-consistent calculations, normally it takes less than 10 iterations for a modest charge carrier density. Even for the case of Multiple quantum

wells(MQW), such as shown in Fig. 3.2, the Broyden method is able to achieve convergence in less than $20 \sim 30$ iterations. The new potential after each Broyden update becomes

$$\phi_{sc}^{new} = \phi_{sc}^{old} + r \cdot \phi_{sc}^{update} \tag{3.13}$$

where ϕ_{sc} is the self-consistent potential and $r \leq 1$ is the under-relaxation parameter, however, in this work r is always chosen to be 1 in contrast to other work using newton-like method. This new potential ϕ^{new} is then used as the input for the $\mathbf{k} \cdot \mathbf{p}$ solver in the next iteration, and this procedure is repeated iteratively until ϕ_{sc} doesn't change any more.

Figure 3.1: 3nm InGaN quantum well, sheet charge density: $1e^{13}cm^{-2}$. The solid lines represent the conduction/valence band profile used in the current Poisson equation, green lines stand for the updated potential profile for the next iteration, the overlapping indicates the self-consistency has been achieved.

As can be seen in Fig. 3.1 and Fig. 3.2, the potential profile at the two ends are pinned at the same position, which is a numerical approximation for MQW at device active regions. From a physical point of view, this scenario corresponds to a photoluminescence test case characterized by a closed boundary condition. Notice that the 2D sheet charge density in a MQW structure is not well-defined in the literature, here we assume the active region of of the MQW structure is thin enough to suit a 2D description and choose a modest value such that the injection level in an individual quantum well is comparable to the single quantum well case. The Gaussian-like electron-hole profiles are the integrated(in k_t) charge carrier distribution in the longitudinal direction.

After the converged potential profile as well as the quantized energy level and associated wave functions are obtained, we have all the ingredients for computing the optical gain and solving Semiconductor-Bloch equations in the following.
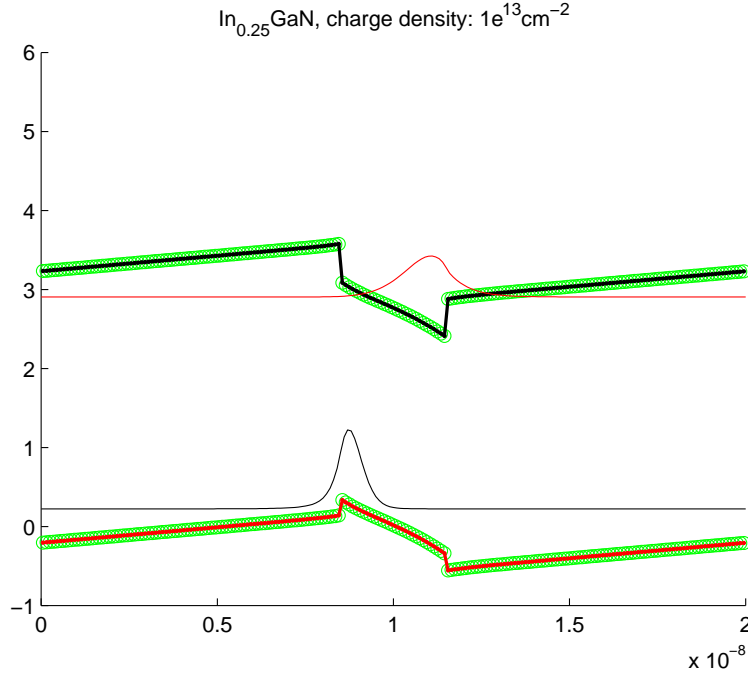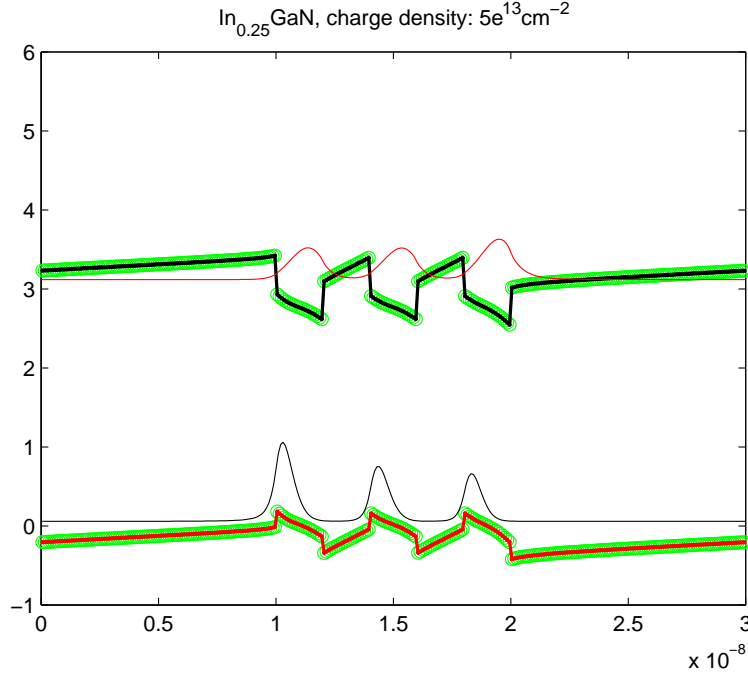

Figure 3.2: InGaN MQW structure, sheet charge density: $5e^{13}cm^{-2}$. The solid black/red lines represent the conduction/valence band profile used in the current Poisson equation, green lines stand for the updated potential profile for the next iteration, the overlapping indicates the self-consistency has been achieved.

A thorough analysis of optical gain has been addressed by Chuang[74] based on Fermi's golden rule, which is equivalent to the approach of solving Semiconductor-Bloch equation at the free carrier theory(FCT) level, see Fig. 3.7. Here we outline the main results using Chuang's method, the gain spectrum can be computed as

$$g(\omega) = C_0 \cdot \frac{2\pi}{\mathcal{L}} \sum_{m,n} \int_0^\infty \frac{k_t dk_t}{2\pi} |M_{mn}^{k_t}|^2 \cdot \frac{\gamma/\pi}{[E_{vm}^{cn}(k_t) - \hbar\omega]^2 + \gamma^2} (f_c^n - f_v^m) \qquad (3.14)$$

where,

$$C_0 = \frac{\pi e^2}{n_r c \epsilon_0 m_0^2 \omega}, \quad E_{vm}^{cn}(k_t) = E_{cn}(k_t) - E_{vm}(k_t) + E_g \qquad (3.15)$$

$E_{vm}^{cn}(k_t)$ represents the energy difference between the n-th conduction subband and the m-th valence subband at k_t, and E_g is the energy gap. The momentum matrix element $M_{mn}^{k_t}$ can be calculated as within the envelope function approximation,

$$M_{mn}^{k_t} = \sum_{i,j} \langle u_{i0} | \hat{p} | u_{j0} \rangle_{V_c} \int_{\mathcal{L}} dz \zeta_{m,i}^*(k_t) \zeta_{n,j}(k_t) \qquad (3.16)$$

34
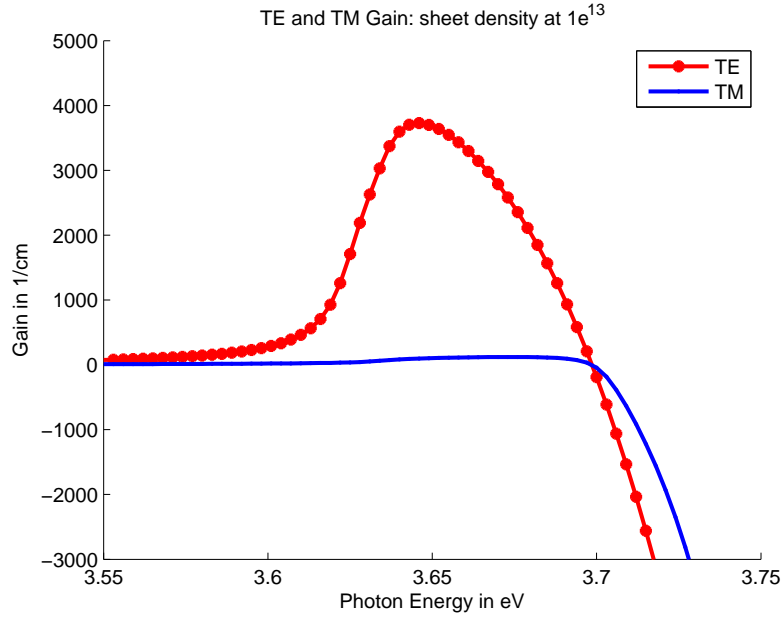
Figure 3.3: 3nm $Al_{0.25}GaN$ Quantum well: gain spectra for a sheet charge density of $N_{2D} = 1.0 \times 10^{13}\,cm^{-2}$, TE and TM mode are plotted respectively.

where u_{i0} and u_{j0} are the zone-center Bloch functions in $\mathbf{k} \cdot \mathbf{p}$ theory, V_c is the crystal cell and $\zeta_{m,i}/\zeta_{n,j}$ are the envelope functions. The TE and TM modes depend on the specific matrix element described by $\langle u_{i0} \,|\hat{p}|\, u_{j0}\rangle_{V_c}$, their gain spectra for a quantum well structure are show in Fig. 3.3, respectively. For brevity, we will drop the subscript of momentum vector k_t, simply use k for the following part.

3.2   Density matrix and Semiconductor-Bloch equations

In the emerging field of semiconductor quantum optics, where the focus is on ultrafast transients, a density matrix approach is usually applied to analyze the correlation dynamics in the coupled quantized light-semiconductor system[75]. In this chapter, Density Matrix method will be applied to analyze the correlations dynamics in the coupled quantized light-semiconductor system, by solving the semiconductor-bloch equation numerically.

Historically density matrix is an import apparatus used to connect quantum mechanics(include second quantization form) with statistic mechanics[76]. One has to distinguish between the so called density operator and density matrix, in simple terms, density matrix is a representation of density operator in some basis. First, we discuss the density

operator

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle \langle \psi_k| \tag{3.17}$$

where ψ is some pure state vector, notice that ρ is not some physical observable but represents a mixed state of the system. We can expand $|\psi\rangle$ using some complete orthonormal basis set ϕ_m

$$|\psi_k\rangle = \sum_m C_m^k |\phi_m\rangle \tag{3.18}$$

therefore,

$$\begin{aligned}
\hat{\rho} &= \sum_k p_k (\sum_m C_m^k |\phi_m\rangle)(\sum_n (C_n^k)^* \langle \phi_n|) \\
&= \sum_{m,n} (\sum_k p_k C_m^k (C_n^k)^*) |\phi_m\rangle \langle \phi_n|
\end{aligned} \tag{3.19}$$

then it is trivial to write the matrix element as

$$\rho_{mn} = \langle \phi_m |\hat{\rho}| \phi_n\rangle = \sum_k p_k C_m^k (C_n^k)^* = \overline{C_m C_n^*} \tag{3.20}$$

where $\overline{C_m C_n^*}$ is the ensemble average of the overlap $C_m C_n^*$. The diagonal part of the density matrix ρ_{mm} gives the probability of finding the system in particular basis state ϕ_m, or, in more detailed way, with $C_m^k (C_m^k)^* \equiv |C_m^k|^2$ being the probability to find the pure state $\rho = |k\rangle \langle k|$ in basis state m, with $\sum_k p_k C_m^k (C_m^k)^*$ being the overall probability to find the system(in mixed state) in basis state m. The other important property of density matrix is the trace operation, in a trivial case we have

$$Tr(\rho) = \sum_m \rho_{mm} = \sum_m \sum_k p_k C_m^k (C_m^k)^* = \sum_k p_k \sum_m |C_m^k|^2 = \sum_k p_k = 1 \tag{3.21}$$

we can also generalize the trace operation to more practical case

$$\begin{aligned}
Tr(\hat{\rho}\hat{A}) &= \sum_q \langle \phi_q |\hat{\rho}\hat{A}| \phi_q\rangle = \sum_q \sum_k p_k (\sum_m (C_m^k)^* \langle \phi_q \mid \phi_m\rangle)(\sum_n C_n^k \langle \phi_n |\hat{A}| \phi_q\rangle) \\
&= \sum_k p_k (\sum_m (C_m^k)^*) \cdot (\sum_n C_n^k \langle \phi_n |\hat{A}| \phi_m\rangle) \\
&= \sum_k p_k \langle \psi^k |\hat{A}| \psi^k\rangle = \overline{\langle A\rangle}
\end{aligned} \tag{3.22}$$

notice that if the system is in a pure state, $\hat{\rho} = |\psi\rangle \langle \psi|$, then

$$Tr(\hat{\rho}\hat{A}) = \langle \psi |\hat{A}| \psi\rangle = \langle A\rangle \tag{3.23}$$

Therefore, we arrive the conclusion that acting the density operator to any operator $\hat{A}$ and performing the trace operation amounting to take the ensemble average of the expectation value of the operator $\hat{A}$ with respect to some mixed state system.

The equation that governs the time evolution of the density matrix is the so called von Neumann equation, which holds within Schrödinger picture and it is immediate to prove in quantum physics. Indeed, as ρ is an incoherent superposition of pure states, one has,

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k| \tag{3.24}$$

Hence, ρ evolves in time as a result of the standard Schrödinger evolution of pure states of the mixture,

$$\hat{\rho}(t) = \sum_k p_k U_t |\psi_k\rangle\langle\psi_k| U_t^\dagger = U_t \rho U_t^\dagger \tag{3.25}$$

where $U_t = e^{\frac{t}{i\hbar}\hat{H}}$ is the usual time evolutor. This identity immediately leads to Liouville-von Neumann equation,

$$\begin{aligned}\frac{\mathrm{d}\hat{\rho}(t)}{\mathrm{d}t} &= \sum_j p_k \left(\frac{\mathrm{d}\,|\psi_k\rangle}{\mathrm{d}t}\langle\psi_k| + |\psi_k\rangle\frac{\mathrm{d}\,\langle\psi_k|}{\mathrm{d}t}\right)\\ &= \sum_j p_k\left(\frac{1}{i\hbar}\hat{H}\,|\psi_k\rangle\,\langle\psi_k| + |\psi_k\rangle\,\langle\psi_k|\,\hat{H}\frac{-1}{i\hbar}\right) = \frac{1}{i\hbar}[\hat{H},\rho]\end{aligned} \tag{3.26}$$

It would be interesting to compare the Liouville-von Neumman equation with the Heisenberg equation of motion, since formally they only differ by a minus sign. Say, if $\rho(t)$ is a time-dependent observable in Schrödinger picture, moving to Heisenberg picture we have

$$\frac{d\rho_H}{dt} = \frac{\partial\rho_H}{\partial t} + \frac{i}{\hbar}[H,\rho_H] = 0 \tag{3.27}$$

if given,

$$\frac{d\rho_H}{dt} = 0 \tag{3.28}$$

Indeed this is trivial without computing any derivative, since

$$\rho_H(t) = U_t^\dagger \rho(t) U_t = U_t^\dagger U_t \rho U_t^\dagger U_t = \rho \tag{3.29}$$

The conserved quantity would be ρ itself, however it's not exactly physical in the sense that ρ is a pure state, not an observable and it can not be governed by Heisenberg evolution, since a state is universally constant in Heisenberg picture. Further discussions

on expectation values in Schrödinger and Heisenberg pictures are given in Appendix. B, there it is demonstrated they are equivalent.

Semiconductor-Bloch equations exhibit themselves in the form of operator-valued equations, within the framework of density matrix theory. We start with the single-particle Hamiltonian describing the free carriers interacting with a classical field

$$H = H_0 + H_{cf} + H_{cc} + H_{cp} \tag{3.30}$$

with the following expressions for the free-carrier and phonon contribution H_0, the carrier–field H_{cf}, the carrier–carrier H_{cc} and the carrier–phonon interaction H_{cp}

$$H_0 = \sum_{l} \varepsilon_l a_l^\dagger a_l + \sum_{u} \hbar \omega_Q \left(b_u^\dagger b_u + \frac{1}{2} \right) \tag{3.31}$$

$$H_{cf} = -\sum_{l_1,l_2} \mu_{l_1,l_2} E\left(t\right) a_{l_1}^\dagger a_{l_2} \tag{3.32}$$

$$H_{cc} = \frac{1}{2} \sum_{A,B,C,D} V_{C\ D}^{A\ B} a_A^\dagger a_B^\dagger a_D a_C \tag{3.33}$$

$$H_{cp} = \sum_{l_1 l_2} \sum_{u} \left[g_u^{l_1\ l_2} a_{l_1}^\dagger a_{l_2} b_u + g_u^{l_1 l_2 *} a_{l_2}^\dagger a_{l_1} b_u^\dagger \right] \tag{3.34}$$

where ε_l is the single-particle energy, the compound index $l = (k, \lambda, s)$ contains the electron momentum k, the band index λ and the spin s. The operator $b_u^\dagger$ (b_u) creates (annihilates) a bulk phonon $u = (q, q_z, j)$ with wave vector (q, q_z) in phonon mode j. Hereafter we will drop spin and phonon mode indices for clarity. The unscreened Coulomb matrix element is

$$V_{c\ d}^{a\ b} = \frac{e^2}{4\pi\varepsilon_0} \int_{\mathcal{V}} \int_{\mathcal{V}} \psi_a^*(r,z)\psi_b^*(r',z') \frac{1}{|(r,z)-(r',z')|} \psi_d(r',z)\psi_c(r,z) dr dr' dz dz' \tag{3.35}$$

Substituting the ansatz

$$\psi(r,z) = e^{ik\cdot r} \sum_{m=1}^{M} u_{m0}(r,z)\zeta_{km}(z) \tag{3.36}$$

we have

$$V_{c\ d}^{a\ b} = \frac{e^2}{4\pi\varepsilon_0} \int_{\mathcal{L}} \int_{\mathcal{L}} dz dz' \zeta_a^*(z)\zeta_b^*(z')\zeta_d(z')\zeta_c(z)$$

$$\times \int_{\mathcal{A}} \int_{\mathcal{A}} \frac{1}{\sqrt{|s|^2 + (z-z')^2}} e^{i(k_c + k_d - k_a - k_b)\cdot r} e^{i(k_b - k_d)\cdot s} dr ds \tag{3.37}$$

The integration over r yields the momentum conserving function $\delta_{k_a + k_b, k_c + k_d}$, reflecting the in the plane translational invariance of the system. Replacing $\frac{1}{\sqrt{|s|^2+(z-z')^2}}$ with its Fourier expansion $\sum_q \frac{2\pi}{\mathcal{A}q} e^{-q|z-z'|} e^{iq\cdot s}$, we have

$$V_{c\ d}^{a\ b} = \frac{e^2}{4\pi\varepsilon_0} \int_{\mathcal{L}} \int_{\mathcal{L}} dz dz' \zeta_a^*(z)\zeta_b^*(z')\zeta_d(z')\zeta_c(z) \int_{\mathcal{A}} \sum_{q} \frac{2\pi}{\mathcal{A}q} e^{-q|z-z'|} e^{iq\cdot s} e^{i(k_2 - k_3)\cdot s} ds \tag{3.38}$$

The integral over s yields $\mathcal{A}\delta_{q,k_b-k_d}$ and we are left with

$$V_{c\ d}^{a\ b} = \frac{e^2}{2\varepsilon_s}\frac{1}{q}\iint \zeta_a^*(z)\zeta_b^*(z')e^{-q|z-z'|}\zeta_d(z')\zeta_c(z)dzdz' \tag{3.39}$$

We restrict ourselves to the interaction matrix elements of electrons with longitudinal optical bulk phonons (three-dimensional) and assume a constant energy dispersion relation. The Frohlic matrix element is

$$g_{q,q_z}^{ab} = -i\sqrt{\frac{e^2\hbar\omega_{LO}}{2\mathcal{A}\mathcal{L}(q^2+q_z^2)}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)}\int_\mathcal{A}\int_\mathcal{L}\psi_a^*(r,z)e^{iq\cdot r}e^{iq_z\cdot z}\psi_b(r,z)drdz \tag{3.40}$$

Again, integration over r yields the momentum conserving function

$$g_{q,q_z}^{ab} = -i\sqrt{\frac{e^2\hbar\omega_{LO}}{2\mathcal{L}(q^2+q_z^2)}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)}\int_\mathcal{L}\zeta_a^*(z)\zeta_b(z)e^{iq_z z}dz \tag{3.41}$$

In general, we are interested in the explicit expression for $g_{q,q_z}^{ab}\left(g_{q,q_z}^{dc}\right)^*$, which, after integration on q_z, takes the following form

$$g_q^{ab}\left(g_q^{dc}\right)^* = \frac{e^2\hbar\omega_{LO}}{2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)\frac{\pi}{q}\iint \zeta_a^*(z)\zeta_c^*(z')e^{-q|z-z'|}\zeta_d(z')\zeta_b(z)dzdz' \tag{3.42}$$

The optical matrix element is given by the expectation value of the scalar product between the momentum operator and the vector potential $A(t)$

$$M_{ab} = -\frac{e_0}{m_0}\langle\psi_a|p\cdot A|\psi_b\rangle \tag{3.43}$$

In the approximation of slowly varying envelopes, the derivatives of the envelopes are small and the matrix element is dominated by

$$M_{ab} \approx -\frac{e_0}{m_0}A\cdot\sum_{ij}\langle u_{i0}|p|u_{j0}\rangle\int \zeta_{ai}^*(z)\zeta_{bj}dz. \tag{3.44}$$

Applying the Heisenberg equation of motion

$$i\hbar\frac{dO}{dt} = [O,H] \tag{3.45}$$

we obtain the dynamics of the microscopic polarization separately for different components of the Hamilton operator

$$\frac{d}{dt}\langle a_1^\dagger a_2\rangle|_{H_0} = \frac{i}{\hbar}(\varepsilon_1-\varepsilon_2)\langle a_1^\dagger a_2\rangle \tag{3.46}$$

$$\frac{d}{dt}\langle a_1^\dagger a_2\rangle|_{H_{cf}} = -\frac{i}{\hbar}\mu_k[\langle a_2^\dagger a_2\rangle-\langle a_1^\dagger a_1\rangle]E(z,t) \tag{3.47}$$

$$\frac{d}{dt}\langle a_1^\dagger a_2\rangle|_{H_{cc}} = +\frac{i}{2\hbar}\sum_{A,B,D}V_{1\ D}^{A\ B}\langle a_A^\dagger a_B^\dagger a_D a_2\rangle-\frac{i}{2\hbar}\sum_{A,B,C}V_{C\ 1}^{A\ B}\langle a_A^\dagger a_B^\dagger a_C a_2\rangle$$

$$+\frac{i}{2\hbar}\sum_{A,C,D}V^{A\;2}_{C\;D}\langle a^\dagger_1 a^\dagger_A a_D a_C\rangle - \frac{i}{2\hbar}\sum_{B,C,D}V^{2\;B}_{C\;D}\langle a^\dagger_1 a^\dagger_B a_D a_C\rangle \tag{3.48}$$

$$\frac{d}{dt}\langle a^\dagger_1 a_2\rangle|_{H_{cp}} = -\frac{i}{2\hbar}\sum_{3u}\left[g_{23}\langle a^\dagger_1 a_3 b_u\rangle - g_{31}\langle a^\dagger_3 a_2 b_u\rangle + g^*_{32}\langle a^\dagger_1 a_3 b^\dagger_u\rangle - g^*_{13}\langle a^\dagger_3 a_2 b^\dagger_u\rangle\right] \tag{3.49}$$

where we have used the compound indices $1 = (k, v)$ and $2 = (k, c)$.

In general, the Semiconductor-Bloch equations can be written in the following form[77]

$$\frac{d}{dt}p_k = -i\tilde{\omega}_k p_k - i\Omega_k(n_{ek} + n_{hk} - 1) + \frac{d}{dt}p_k|_{col}$$
$$\frac{d}{dt}\rho_{ek} = i[\Omega_k p^*_k - \Omega^*_k p_k] + \frac{d}{dt}\rho_{ek}|_{col}$$
$$\frac{d}{dt}\rho_{hk} = i[\Omega_k p^*_k - \Omega^*_k p_k] + \frac{d}{dt}\rho_{hk}|_{col} \tag{3.50}$$

In the following sections we will explore the individual equation as well as the collision terms in the all-electron picture[78].

3.3 Semiconductor-Bloch equations: first order

The first order terms derived in this section will address the explicit form of the terms $\tilde{\omega}$ and Ω_k in Eg. 3.50. We omit the carrier-phonon interaction H_{cp} for the moment without loss of generality, the many-particle Hamiltonian consists of the free-carrier part H_0, carrier-field part H_{cf} and carrier-carrier part H_{cc},

$$H = H_0 + H_{cf} + H_{cc}$$
$$= \sum_l \varepsilon_l a^\dagger_l a_l - \sum_{l,l',k}\mu^{l,l'}_k E(z,t) a^\dagger_{l,k} a_{l',k} + \frac{1}{2}\sum_{A,B,C,D}V^{A\;B}_{C\;D} a^\dagger_A a^\dagger_B a_D a_C \tag{3.51}$$

Applying the Heisenberg equation(Append. B) of motion to a generic two operator term, we have

$$\frac{d}{dt}a^\dagger_1 a_2 = \frac{i}{\hbar}\left[H, a^\dagger_1 a_2\right]$$
$$= \frac{i}{\hbar}(\varepsilon_1 - \varepsilon_2)a^\dagger_1 a_2 - \frac{i}{\hbar}\mu_k\left(a^\dagger_2 a_2 - a^\dagger_1 a_1\right)E(z,t)$$
$$+ \frac{i}{2\hbar}\sum_{A,B,D}V^{A\;B}_{1\;D} a^\dagger_A a^\dagger_B a_D a_2 - \frac{i}{2\hbar}\sum_{A,B,C}V^{A\;B}_{C\;1} a^\dagger_A a^\dagger_B a_C a_2$$
$$+ \frac{i}{2\hbar}\sum_{A,C,D}V^{A\;2}_{C\;D} a^\dagger_1 a^\dagger_A a_D a_C - \frac{i}{2\hbar}\sum_{B,C,D}V^{2\;B}_{C\;D} a^\dagger_1 a^\dagger_B a_D a_C \tag{3.52}$$

Indices renaming and Coulomb matrix element symmetry are frequently used to simplify the expressions in the all-particle picture, in this case it consists of three steps,

$$h_{cc} = \frac{i}{2\hbar}\sum_{A,B,D}V^{A\;B}_{1\;D} a^\dagger_A a^\dagger_B a_D a_2 - \frac{i}{2\hbar}\sum_{A,B,C}V^{A\;B}_{C\;1} a^\dagger_A a^\dagger_B a_C a_2 \tag{3.53}$$

First we exploit the symmetry $V_{C\ D}^{A\ B} = V_{D\ C}^{B\ A}$ (the other universally valid symmetry rule is $V_{C\ D}^{A\ B} = V_{A\ B}^{C\ D\,*}$)

$$h_{cc} = \frac{i}{2\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2 - \frac{i}{2\hbar} \sum_{A,B,C} V_{1\ C}^{B\ A}\ a_A^\dagger a_B^\dagger a_C a_2 \tag{3.54}$$

Since the indices A, B, C, D are generic and can represent any quantum state, we are free to rename them, by applying transformation $B \to A, A \to B, C \to D$ to the second term, we have

$$h_{cc} = \frac{i}{2\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2 - \frac{i}{2\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_B^\dagger a_A^\dagger a_D a_2 \tag{3.55}$$

Finally we invoke the anti-commutation relation, and collect these two terms,

$$\begin{aligned}
h_{cc} &= \frac{i}{2\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2 + \frac{i}{2\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2 \\
&= \frac{i}{\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2
\end{aligned} \tag{3.56}$$

This looks as if the relation $V_{C\ D}^{A\ B} = -V_{C\ D}^{B\ A} = V_{D\ C}^{B\ A}$ has been used, which is not true in general. By the same token, we can collect the other two terms in Eq. 3.52 and arrive at.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}} a_1^\dagger a_2 =\ & \frac{i}{\hbar} (\varepsilon_1 - \varepsilon_2) a_1^\dagger a_2 - \frac{i}{\hbar} \mu_k \left(a_2^\dagger a_2 - a_1^\dagger a_1 \right) E(z,t) \\
& + \frac{i}{\hbar} \sum_{A,B,D} V_{1\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_2 - \frac{i}{\hbar} \sum_{A,B,D} V_{B\ D}^{2\ A}\ a_1^\dagger a_A^\dagger a_D a_B
\end{aligned} \tag{3.57}$$

Now we prove this is equivalent to the more prevalent electron-hole picture[79], in the most simple case, i.e. the two band limit, $A, B, D = c, v|\forall k$, we have 2^3 combinations out of each four-operators term before performing Hartree-Fock(HF) factorization, this leads to the following explicit form(the resulting summation indices reduce to k', q, and a, b, d are the band indices associated with the corresponding composite indices A, B, D),

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}} a_{vk}^\dagger a_{ck}|_{H_{cc}} =\ & \frac{i}{\hbar} \sum_{A,B,D} V_{vk\ D}^{A\ B}\ a_A^\dagger a_B^\dagger a_D a_{ck} - \frac{i}{\hbar} \sum_{A,B,D} V_{B\ D}^{ck\ A}\ a_{vk}^\dagger a_A^\dagger a_D a_B \\
=\ & \frac{i}{\hbar} \sum_{a,b,d}^{k',q} \left\{ V_{vk\ dk'}^{ak+q\ bk'-q}\ a_{ak+q}^\dagger a_{bk'-q}^\dagger a_{dk'} a_{ck} - V_{bk-q\ dk'}^{ck\ \ \ ak'-q}\ a_{vk}^\dagger a_{ak'-q}^\dagger a_{dk'} a_{bk-q} \right\} \\
=\ & \frac{i}{\hbar} \sum_{k',q} \left\{ V_{vk\ ck'}^{ck+q\ ck'-q}\ a_{ck+q}^\dagger a_{ck'-q}^\dagger a_{ck'} a_{ck} - V_{ck-q\ ck'}^{ck\ \ \ ck'-q}\ a_{vk}^\dagger a_{ck'-q}^\dagger a_{ck'} a_{ck-q} \right\} \\
& + \frac{i}{\hbar} \sum_{k',q} \left\{ V_{vk\ vk'}^{vk+q\ vk'-q}\ a_{vk+q}^\dagger a_{vk'-q}^\dagger a_{vk'} a_{ck} - V_{vk-q\ vk'}^{ck\ \ \ vk'-q}\ a_{vk}^\dagger a_{vk'-q}^\dagger a_{vk'} a_{vk-q} \right\} \\
& + \frac{i}{\hbar} \sum_{k',q} \left\{ V_{vk\ vk'}^{ck+q\ ck'-q}\ a_{ck+q}^\dagger a_{ck'-q}^\dagger a_{vk'} a_{ck} - V_{ck-q\ vk'}^{ck\ \ \ ck'-q}\ a_{vk}^\dagger a_{ck'-q}^\dagger a_{vk'} a_{ck-q} \right\}
\end{aligned}$$

$$+\frac{i}{\hbar}\sum_{k',q}\left\{V^{ck+q\ vk'-q}_{vk\ \ \ vk'}a^{\dagger}_{ck+q}a^{\dagger}_{vk'-q}a_{vk'}a_{ck} - V^{ck\ \ \ ck'-q}_{vk-q\ vk'}a^{\dagger}_{vk}a^{\dagger}_{ck'-q}a_{vk'}a_{vk-q}\right\}$$

$$+\frac{i}{\hbar}\sum_{k',q}\left\{V^{ck+q\ vk'-q}_{vk\ \ \ ck'}a^{\dagger}_{ck+q}a^{\dagger}_{vk'-q}a_{ck'}a_{ck} - V^{ck\ \ \ ck'-q}_{vk-q\ ck'}a^{\dagger}_{vk}a^{\dagger}_{ck'-q}a_{ck'}a_{vk-q}\right\}$$

$$+\frac{i}{\hbar}\sum_{k',q}\left\{V^{vk+q\ vk'-q}_{vk\ \ \ ck'}a^{\dagger}_{vk+q}a^{\dagger}_{vk'-q}a_{ck'}a_{ck} - V^{ck\ \ \ vk'-q}_{vk-q\ ck'}a^{\dagger}_{vk}a^{\dagger}_{vk'-q}a_{ck'}a_{vk-q}\right\}$$

$$+\frac{i}{\hbar}\sum_{k',q}\left\{V^{vk+q\ ck'-q}_{vk\ \ \ ck'}a^{\dagger}_{vk+q}a^{\dagger}_{ck'-q}a_{ck'}a_{ck} - V^{ck\ \ \ vk'-q}_{ck-q\ ck'}a^{\dagger}_{vk}a^{\dagger}_{vk'-q}a_{ck'}a_{ck-q}\right\}$$

$$+\frac{i}{\hbar}\sum_{k',q}\left\{V^{vk+q\ ck'-q}_{vk\ \ \ vk'}a^{\dagger}_{vk+q}a^{\dagger}_{ck'-q}a_{vk'}a_{ck} - V^{ck\ \ \ vk'-q}_{ck-q\ vk'}a^{\dagger}_{vk}a^{\dagger}_{vk'-q}a_{vk'}a_{ck-q}\right\}$$

$$(3.58)$$

Then we make the diagonal approximation, i.e. drop all the coulomb terms that do not conserve particle numbers,

$$i\hbar\cdot\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{vk}a_{ck}\big|_{H_{cc}} = -\sum_{k',q}V^{vk+q\ vk'-q}_{vk\ \ \ vk'}a^{\dagger}_{vk+q}a^{\dagger}_{vk'-q}a_{vk'}a_{ck}$$

$$-\sum_{k',q}V^{ck+q\ vk'-q}_{vk\ \ \ ck'}a^{\dagger}_{ck+q}a^{\dagger}_{vk'-q}a_{ck'}a_{ck}$$

$$-\sum_{k',q}V^{vk+q\ ck'-q}_{vk\ \ \ ck'}a^{\dagger}_{vk+q}a^{\dagger}_{ck'-q}a_{ck'}a_{ck}$$

$$+\sum_{k',q}V^{ck\ \ \ ck'-q}_{ck-q\ ck'}a^{\dagger}_{vk}a^{\dagger}_{ck'-q}a_{ck'}a_{ck-q}$$

$$+\sum_{k',q}V^{ck\ \ \ vk'-q}_{vk-q\ ck'}a^{\dagger}_{vk}a^{\dagger}_{vk'-q}a_{ck'}a_{vk-q}$$

$$+\sum_{k',q}V^{ck\ \ \ vk'-q}_{ck-q\ vk'}a^{\dagger}_{vk}a^{\dagger}_{vk'-q}a_{vk'}a_{ck-q} \qquad (3.59)$$

As usual, we perform the HF factorization and make the random phase approximation(RPA),

$$i\hbar\cdot\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{vk}a_{ck}\big|_{H_{cc}} = +\sum_{q}V^{vk+q\ vk}_{vk\ \ \ vk+q}n_v(k+q)p(k)$$

$$+\sum_{q}V^{ck+q\ vk}_{vk\ \ \ ck+q}n_c(k+q)p(k)$$

$$+\sum_{q}V^{vk+q\ ck}_{vk\ \ \ ck+q}n_c(k)p(k+q)$$

$$-\sum_{q}V^{ck\ \ \ ck-q}_{ck-q\ ck}n_c(k-q)p(k)$$

$$-\sum_{q}V^{ck\ \ \ vk-q}_{vk-q\ ck}n_v(k-q)p(k)$$

$$-\sum_q V^{ck}_{ck-q} {}^{vk-q}_{vk} n_v(k)p(k-q) \tag{3.60}$$

Now we can group terms and compare with the Semiconductor-Bloch equations equations in electron-hole picture,

$$
\begin{aligned}
i\hbar \cdot \frac{\mathrm{d}}{\mathrm{d}t} a^\dagger_{vk} a_{ck}|_{H_{cc}} =\ & \sum_q V^{vk+q}_{vk} {}^{vk}_{vk+q} n_v(k+q)p(k) - \sum_q V^{ck}_{ck+q} {}^{ck+q}_{ck} n_c(k+q)p(k) \\
& + \sum_q V^{vk+q}_{vk} {}^{ck}_{ck+q} n_c(k)p(k+q) - \sum_q V^{ck}_{ck+q} {}^{vk+q}_{vk} n_v(k)p(k+q) \\
& + \sum_q V^{ck+q}_{vk} {}^{vk}_{ck+q} n_c(k+q)p(k) - \sum_q V^{ck}_{vk+q} {}^{vk+q}_{ck} n_v(k+q)p(k) \tag{3.61}
\end{aligned}
$$

In the above equation, the first row indicates the diagonal band gap renormalization, the second corresponds to the nondiagonal term, the third is omitted since it represent a interband transition.

$$
\begin{aligned}
i\hbar \cdot \frac{\mathrm{d}}{\mathrm{d}t} a^\dagger_{vk} a_{ck}|_{H_{cc}} =\ & \left(\sum_q V^{vk+q}_{vk} {}^{vk}_{vk+q} n_v(k+q) - \sum_q V^{ck}_{ck+q} {}^{ck+q}_{ck} n_c(k+q) \right) p(k) \\
& - \sum_q V^{ck}_{ck+q} {}^{vk+q}_{vk} (n_v(k) - n_c(k))p(k+q) \\
=\ & \left(\sum_{k'} V^{vk'}_{vk} {}^{vk}_{vk'} n_v(k') - \sum_{k'} V^{ck}_{ck'} {}^{ck'}_{ck} n_c(k') \right) p(k) \\
& - \left(\sum_{k'} V^{ck}_{ck'} {}^{vk'}_{vk} p(k') \right) (n_v(k) - n_c(k)) \tag{3.62}
\end{aligned}
$$

Finally it is demonstrated that the Semiconductor-Bloch equations derived in all-electron picture is equivalent to the one in the electron-hole picture, with $n_v(k)$ substituted by $1 - n_h(k)$.

$$\frac{\mathrm{d}}{\mathrm{d}t} p_k = -i\tilde{\omega}_k p_k - \frac{i}{\hbar} \left\{ E(z,t)\mu_k + \sum_{k'} V^{ck}_{ck'} {}^{vk'}_{vk} p_{k'} \right\} (n_{ek} + n_{hk} - 1) + \frac{\mathrm{d}}{\mathrm{d}t} p_k|_{col} \tag{3.63}$$

$$
\begin{aligned}
\hbar\tilde{\omega}_k &= (\varepsilon_{ck} - \varepsilon_{vk}) - \sum_{k'} \left\{ V^{ck}_{ck'} {}^{ck'}_{ck} n_c(k') - V^{vk'}_{vk} {}^{vk}_{vk'} n_v(k') \right\} \\
&= \left(\varepsilon_{ck} - (\varepsilon_{vk} - \sum_{k'} V^{vk'}_{vk} {}^{vk}_{vk'}) \right) - \sum_{k'} \left\{ V^{ck}_{ck'} {}^{ck'}_{ck} n_c(k') + V^{vk'}_{vk} {}^{vk}_{vk'} (1 - n_v(k')) \right\} \\
&= (\varepsilon_{ck} + \varepsilon_{hk}) - \sum_{k'} \left\{ V^{ck}_{ck'} {}^{ck'}_{ck} n_c(k') + V^{vk'}_{vk} {}^{vk}_{vk'} n_h(k') \right\} \tag{3.64}
\end{aligned}
$$

Notice that in the renormalized energy $\tilde{\omega}_k$, relation $\varepsilon_{hk} = -(\varepsilon_{vk} - \sum_{k'} V^{vk'}_{vk} {}^{vk}_{vk'})$ has been used.

3.4 Semiconductor-Bloch equations: second order

The second order terms derived in this section will lead to the collision terms in Eg. 3.50. Again, the equation of motion of the polarization operator at HF level can be written as,

$$
\begin{aligned}
\frac{d}{dt}a_1^\dagger a_2 =\ & \frac{i}{\hbar}(\varepsilon_1 - \varepsilon_2)\, a_1^\dagger a_2 - \frac{i}{\hbar}\mu_k\left(a_2^\dagger a_2 - a_1^\dagger a_1\right)E\left(z,t\right) \\
& + \frac{i}{\hbar}\sum_{A,B,C} V_{1\ C}^{A\ B}\, a_A^\dagger a_B^\dagger a_C a_2 - \frac{i}{\hbar}\sum_{A,B,C} V_{B\ C}^{2\ A}\, a_1^\dagger a_A^\dagger a_C a_B
\end{aligned}
\tag{3.65}
$$

The next step in the hierarchy is obtained by including the correlation terms that describe the deviations from the corresponding HF factorization, i.e., the terms of the form

$$
\tilde{C}_{3\ 4}^{1\ 2} = C_{3\ 4}^{1\ 2} - (\sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24})
\tag{3.66}
$$

where,

$$
C_{3\ 4}^{1\ 2} = \langle a_1^\dagger a_2^\dagger a_3 a_4 \rangle
$$

The dynamics of such four-operator terms will again couple to six-operator terms and yields

$$
\begin{aligned}
i\hbar\frac{d}{dt}C_{3\ 4}^{1\ 2}\big|_{H_0} =\ & -(\varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4)C_{3\ 4}^{1\ 2} \\
i\hbar\frac{d}{dt}C_{3\ 4}^{1\ 2}\big|_{H_{cc}} =\ & \sum_{abc} V_{1\ c}^{a\ b}\left(\langle a_a^\dagger a_b^\dagger a_2^\dagger a_c a_3 a_4\rangle - C_{3\ 4}^{a\ b}\delta_{2c}\right) \\
& -\sum_{abc} V_{2\ c}^{a\ b}\langle a_1^\dagger a_a^\dagger a_b^\dagger a_c a_3 a_4\rangle + \sum_{abc} V_{b\ c}^{3\ a}\langle a_1^\dagger a_2^\dagger a_a^\dagger a_4 a_c a_b\rangle \\
& -\sum_{abc} V_{b\ c}^{4\ a}\left(\langle a_1^\dagger a_2^\dagger a_a^\dagger a_3 a_c a_b\rangle - C_{c\ b}^{1\ 2}\delta_{3a}\right)
\end{aligned}
\tag{3.67}
$$

where anti-commutation relation has been used for various times,

$$
\begin{aligned}
i\hbar\frac{d}{dt}(\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23})\big|_{H_0} =\ & -(\varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4)(\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23}) \\
i\hbar\frac{d}{dt}(\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23})\big|_{H_{cc}} =\ & \sum_{abc} V_{1\ c}^{a\ b}(C_{c\ 4}^{a\ b}\sigma_{23} - C_{c\ 3}^{a\ b}\sigma_{24}) + \sum_{abc} V_{2\ c}^{2\ b}(C_{c\ 3}^{a\ b}\sigma_{14} - C_{c\ 4}^{a\ b}\sigma_{13}) \\
& \sum_{abc} V_{b\ c}^{3\ a}(C_{c\ b}^{1\ a}\sigma_{24} - C_{c\ b}^{2\ a}\sigma_{14}) + \sum_{abc} V_{b\ c}^{4\ a}(C_{c\ b}^{2\ a}\sigma_{13} - C_{c\ b}^{1\ a}\sigma_{23})
\end{aligned}
\tag{3.68}
$$

Neglecting all the high order correlation terms arising from Eq. 3.67 and Eq. 3.68 leads to

$$
\frac{d}{dt}\tilde{C}_{3\ 4}^{1\ 2}\big|_{H_0 + H_{cc}} = (\frac{i}{\hbar}\Delta\varepsilon - \gamma)\tilde{C}_{3\ 4}^{1\ 2} - \frac{i}{\hbar}K(t)
\tag{3.69}
$$

where γ is a phenomenological parameter taking into account all the high order contributions, and the scattering kernel K is,

$$K = \sum_{abcd}(\mathrm{V}^{a\ b}_{d\ c} - \mathrm{V}^{b\ a}_{d\ c})(\sigma_{1d}\sigma_{2c}\sigma^{\dagger}_{b3}\sigma^{\dagger}_{a4} - \sigma^{\dagger}_{1d}\sigma^{\dagger}_{2c}\sigma_{b3}\sigma_{a4}) \tag{3.70}$$

Derivation of the scattering kernel K is a lengthy process, during which the essential ingredients are

$$\hat{\mathrm{V}}^{a\ b}_{c\ d} = \mathrm{V}^{a\ b}_{c\ d} - \mathrm{V}^{b\ a}_{c\ d} = \mathrm{V}^{a\ b}_{c\ d} - \mathrm{V}^{a\ b}_{d\ c} = \hat{\mathrm{V}}^{b\ a}_{d\ c},$$

$$\mathrm{C}^{a\ b}_{c\ d} = \sigma_{ad}\sigma_{bc} - \sigma_{ac}\sigma_{bd} + \tilde{\mathrm{C}}^{a\ b}_{c\ d},$$

$$\mathrm{C}^{a\ b\ c}_{d\ e\ f} = \sigma_{ad}(\sigma_{bf}\sigma_{ce} - \sigma_{be}\sigma_{cf}) + \sigma_{ae}(\sigma_{bd}\sigma_{cf} - \sigma_{bf}\sigma_{cd}) + \sigma_{af}(\sigma_{be}\sigma_{cd} - \sigma_{bd}\sigma_{ce})$$
$$+ \tilde{\mathrm{C}}^{a\ b}_{d\ e}\sigma_{cf} - \tilde{\mathrm{C}}^{a\ b}_{d\ f}\sigma_{ce} + \tilde{\mathrm{C}}^{a\ b}_{e\ f}\sigma_{cd} - \tilde{\mathrm{C}}^{a\ c}_{d\ e}\sigma_{bf} + \tilde{\mathrm{C}}^{a\ c}_{d\ f}\sigma_{be} - \tilde{\mathrm{C}}^{a\ c}_{f\ e}\sigma_{bd}$$
$$+ \tilde{\mathrm{C}}^{b\ c}_{d\ e}\sigma_{af} - \tilde{\mathrm{C}}^{b\ c}_{d\ f}\sigma_{af} + \tilde{\mathrm{C}}^{b\ c}_{e\ f}\sigma_{ad} + \tilde{\mathrm{C}}^{a\ b\ c}_{d\ e\ f} \tag{3.71}$$

Since we are interested in the terms of the order up to $\tilde{\mathrm{C}}^{a\ b}_{c\ d}$, the correlation expansion are truncated here and higher order terms $\tilde{\mathrm{C}}^{a\ b}_{c\ d}\sigma_{ef}, \tilde{\mathrm{C}}^{a\ b\ c}_{d\ e\ f}$ are all neglected. After factorizing the expression in Eq. 3.67 and Eq. 3.68 and summing them up, it turns out that all the terms arising from Eq. 3.68 can be canceled out, leaving,

$$\begin{aligned}
K =\ & \sum_{abcd}\mathrm{V}^{a\ b}_{d\ c}\delta_{1d}\left[(-\sigma_{a3}\sigma_{b4}\sigma_{2c} + \sigma_{a4}\sigma_{b3}\sigma_{2c}) - (\sigma_{a4}\sigma_{b3} - \sigma_{a3}\sigma_{b4})\delta_{2c}\right] \\
& + \sum_{abcd}\mathrm{V}^{a\ b}_{d\ c}\delta_{2d}(-\sigma_{1c}\sigma_{a4}\sigma_{b3} + \sigma_{1c}\sigma_{a3}\sigma_{b4}) \\
& + \sum_{abcd}\mathrm{V}^{d\ a}_{b\ c}\delta_{3d}(-\sigma_{1c}\sigma_{2b}\sigma_{a4} + \sigma_{1b}\sigma_{2c}\sigma_{a4}) \\
& + \sum_{abcd}\mathrm{V}^{d\ a}_{b\ c}\delta_{4d}\left[(\sigma_{1c}\sigma_{2b}\sigma_{a3} - \sigma_{1b}\sigma_{2c}\sigma_{a3}) + (\sigma_{1b}\sigma_{2c} - \sigma_{1c}\sigma_{2b})\delta_{a3}\right] \\
=\ & \sum_{abcd}\mathrm{V}^{a\ b}_{d\ c}\delta_{1d}\left[\sigma_{a3}\sigma_{b4}(\delta_{2c} - \sigma_{2c}) - \sigma_{a4}\sigma_{b3}(\delta_{2c} - \sigma_{2c})\right] \\
& + \sum_{abcd}\mathrm{V}^{a\ b}_{c\ d}\delta_{2c}(-\sigma_{1d}\sigma_{a4}\sigma_{b3} + \sigma_{1d}\sigma_{a3}\sigma_{b4}) \quad\quad d \leftrightarrow c \\
& + \sum_{abcd}\mathrm{V}^{b\ a}_{d\ c}\delta_{3b}(-\sigma_{1c}\sigma_{2d}\sigma_{a4} + \sigma_{1d}\sigma_{2c}\sigma_{a4}) \quad\quad d \leftrightarrow b \\
& + \sum_{abcd}\mathrm{V}^{a\ d}_{b\ c}\delta_{4a}\left[\sigma_{1b}\sigma_{2c}(\delta_{d3} - \sigma_{d3}) - \sigma_{1c}\sigma_{2b}(\delta_{d3} - \sigma_{d3})\right] \quad\quad d \leftrightarrow a \\
=\ & \sum_{abcd}\hat{\mathrm{V}}^{a\ b}_{d\ c}\delta_{1d}\left[\sigma_{a4}\sigma_{b3}(\sigma_{2c} - \delta_{2c})\right] \\
& + \sum_{abcd}\hat{\mathrm{V}}^{b\ a}_{c\ d}\delta_{2c}(\sigma_{1d}\sigma_{a4}\sigma_{b3}) \quad\quad a \leftrightarrow b \\
& - \sum_{abcd}\hat{\mathrm{V}}^{b\ a}_{c\ d}\delta_{3b}(\sigma_{a4}\sigma_{1d}\sigma_{2c}) \quad\quad d \leftrightarrow c
\end{aligned}$$

45

$$+ \sum_{abcd} \hat{V}^a_d{}^b_c \delta_{4a} \left[\sigma_{1d}\sigma_{2c}(\delta_{b3} - \sigma_{b3}) \right] \qquad \textcolor{red}{d \leftrightarrow b} \tag{3.72}$$

At this step, a useful zero addition(highlighted with blue color) is used to formally symmetrize the scatter "in" and "out" parts,

$$
\begin{aligned}
K = \; & \sum_{abcd} \hat{V}^a_d{}^b_c \sigma_{b3}\sigma_{a4} \left[-\delta_{1d}(\delta_{2c} - \sigma_{2c}) + \delta_{2c}\sigma_{1d} - \textcolor{red}{\sigma_{1d}\sigma_{2c}} \right] \\
& + \sum_{abcd} \hat{V}^a_d{}^b_c \sigma_{1d}\sigma_{2c} \left[\textcolor{blue}{\sigma_{b3}\sigma_{a4}} - \delta_{b3}\sigma_{a4} + \delta_{a4}(\delta_{b3} - \sigma_{b3}) \right]
\end{aligned}
\tag{3.73}
$$

Finally, using the relation $\sigma^\dagger_{ab} = \delta_{ab} - \sigma_{ab}$,

$$
\begin{aligned}
K = \; & -\sum_{abcd} \hat{V}^a_d{}^b_c \sigma_{b3}\sigma_{a4} \left[(\delta_{1d} - \sigma_{1d})(\delta_{2c} - \sigma_{2c})\sigma_{b3}\sigma_{a4} \right] \\
& \sum_{abcd} \hat{V}^a_d{}^b_c \sigma_{b3}\sigma_{a4} \left[\sigma_{1d}\sigma_{2c}(\delta_{b3} - \sigma_{b3})(\delta_{a4} - \sigma_{a4}) \right] \\
= \; & \sum_{abcd} \hat{V}^a_d{}^b_c (\sigma_{1d}\sigma_{2c}\sigma^\dagger_{b3}\sigma^\dagger_{a4} - \sigma^\dagger_{1d}\sigma^\dagger_{2c}\sigma_{b3}\sigma_{a4})
\end{aligned}
\tag{3.74}
$$

The first order differential equation characterized by Eq. 3.69 is of the form $y' + Py = Q$, which can be solved by the method of integrating factor, resulting in the standard solution,

$$y = e^{-I} \int Q e^I dx + c e^{-I} \tag{3.75}$$

where $I = \int P\,dx$ is the so-called integrating factor, then by the markov approximation we can take Q outside the integral,

$$y(t) \approx Q(t) \int_{-\infty}^{t} e^{I(t') - I(t)} dt' \tag{3.76}$$

the formal solution of Eq. 3.69 turns out to be

$$
\begin{aligned}
\tilde{C}^1_3{}^2_4(t) = \; & -\frac{i}{h} K(t) \int_{-\infty}^{t} exp\left[(\frac{i}{\hbar}\Delta\varepsilon - \gamma)(t - t') \right] dt' \\
= \; & -\frac{i}{h} K(t) \int_{-\infty}^{0} exp\left[(\frac{i}{\hbar}\Delta\varepsilon - \gamma)(t - t') \right] d(t' - t) \\
= \; & -\frac{i}{\hbar} K(t) \lim_{\gamma \to 0} \frac{1}{-(\frac{i}{\hbar}\Delta\varepsilon - \gamma)} = -\frac{i}{\hbar} K(t) \lim_{\gamma \to 0} \frac{i}{\Delta\varepsilon/\hbar + i\gamma} \\
= \; & -K(t) \left[i\pi\delta(\Delta\varepsilon) - \mathcal{P}(\frac{1}{\Delta\varepsilon/\hbar})\frac{1}{\hbar} \right]
\end{aligned}
\tag{3.77}
$$

where $\delta(\alpha x) = \frac{1}{|\alpha|}\delta(x)$ has been used, $\mathcal{P}$ stands for the Cauchy principal value and γ can be an arbitrarily small but positive number. Neglecting the Cauchy principal value and plugging the result for $\tilde{C}^1_3{}^2_4(t)$ into the HF equation(Eq. 3.65), we arrive at,

$$\frac{d}{dt}\sigma_{12}\Big|_{2nd} = \frac{\pi}{\hbar} \sum_{ABC} V^2_B{}^A_C \left[\sum_{abcd} \hat{V}^a_d{}^b_c (\sigma^\dagger_{1d}\sigma^\dagger_{Ac}\sigma_{bC}\sigma_{aB} - \sigma_{1d}\sigma_{Ac}\sigma^\dagger_{bC}\sigma^\dagger_{aB})\delta(\varepsilon_1 + \varepsilon_A - \varepsilon_C - \varepsilon_B) \right]$$

$$-\frac{\pi}{\hbar} \sum_{ABC} V_1^{A\ B}{}_C \left[\sum_{abcd} \hat{V}_d^{a\ b}{}_c (\sigma_{Ad}^\dagger \sigma_{Bc}^\dagger \sigma_{bC} \sigma_{a2} - \sigma_{Ad} \sigma_{Bc} \sigma_{bC}^\dagger \sigma_{a2}^\dagger) \delta(\varepsilon_A + \varepsilon_B - \varepsilon_C - \varepsilon_2) \right]$$

(3.78)

The above equation is written in a general form, a straightforward application would be the time derivative for diagonal density matrix $\dot{\rho}_1 = \frac{\mathrm{d}}{\mathrm{d}t} \sigma_{11}$,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \sigma_{11}\big|_{2nd} = \quad & \frac{\pi}{\hbar} \sum_{abc} V_b^{1\ a}{}_c \left[\hat{V}_1^{b\ c}{}_a ((1-\rho_1)(1-\rho_a)\rho_c\rho_b - \rho_1\rho_a(1-\rho_c)(1-\rho_b))\delta(\varepsilon_1 + \varepsilon_a - \varepsilon_c - \varepsilon_b) \right] \\
& -\frac{\pi}{\hbar} \sum_{abc} V_1^{a\ b}{}_c \left[\hat{V}_a^{1\ c}{}_b ((1-\rho_a)(1-\rho_b)\rho_c\rho_1 - \rho_a\rho_b(1-\rho_c)(1-\rho_1))\delta(\varepsilon_a + \varepsilon_b - \varepsilon_c - \varepsilon_1) \right]
\end{aligned}
$$

(3.79)

applying the change $a \leftrightarrow c$ to the second row,

$$
\begin{aligned}
= \quad & \frac{\pi}{\hbar} \sum_{abc} V_b^{1\ a}{}_c \left[\hat{V}_1^{b\ c}{}_a ((1-\rho_1)(1-\rho_a)\rho_c\rho_b - \rho_1\rho_a(1-\rho_c)(1-\rho_b))\delta(\varepsilon_1 + \varepsilon_a - \varepsilon_c - \varepsilon_b) \right] \\
& -\frac{\pi}{\hbar} \sum_{abc} V_1^{c\ b}{}_a \left[\hat{V}_c^{1\ a}{}_b ((1-\rho_c)(1-\rho_b)\rho_a\rho_1 - \rho_c\rho_b(1-\rho_a)(1-\rho_1))\delta(\varepsilon_c + \varepsilon_b - \varepsilon_a - \varepsilon_1) \right] \\
= \quad & \frac{\pi}{\hbar} \sum_{abc} (V_b^{1\ a}{}_c \hat{V}_1^{b\ c}{}_a + V_1^{b\ c}{}_a \hat{V}_b^{1\ a}{}_c)\delta(\varepsilon_1 + \varepsilon_a - \varepsilon_c - \varepsilon_b) \times \\
& [(1-\rho_1)(1-\rho_a)\rho_c\rho_b - \rho_1\rho_a(1-\rho_c)(1-\rho_b)] \qquad {\color{red} b \leftrightarrow c} \\
= \quad & \frac{2\pi}{\hbar} \sum_{abc} V_b^{1\ a}{}_c \hat{V}_1^{b\ c}{}_a [(1-\rho_1)(1-\rho_a)\rho_c\rho_b - \rho_1\rho_a(1-\rho_c)(1-\rho_b)] \, \delta(\varepsilon_1 + \varepsilon_a - \varepsilon_b - \varepsilon_c)
\end{aligned}
$$

(3.80)

The results for $\dot{\rho}_1$ can be written in a more illustrative way,

$$\dot{\rho}_1\big|_{2nd} = \Gamma_1^{in(cc)}(1-\rho_1) - \Gamma_1^{out(cc)}\rho_1 \tag{3.81}$$

with the in- and out-scattering rates:

$$\Gamma_1^{in(cc)} = \frac{2\pi}{\hbar} \sum_{abc} V_b^{1\ a}{}_c \hat{V}_1^{b\ c}{}_a (1-\rho_a)\rho_b\rho_c\delta(\varepsilon_1 + \varepsilon_a - \varepsilon_b - \varepsilon_c) \tag{3.82}$$

$$\Gamma_1^{out(cc)} = \frac{2\pi}{\hbar} \sum_{abc} V_b^{1\ a}{}_c \hat{V}_1^{b\ c}{}_a \rho_a(1-\rho_b)(1-\rho_c)\delta(\varepsilon_1 + \varepsilon_a - \varepsilon_b - \varepsilon_c) \tag{3.83}$$

The result for the dynamics of non-diagonal density matrix $\dot{p}_k$ can be cast in the following form,

$$\frac{\mathrm{d}}{\mathrm{d}t} p_k\big|_{2nd} = \frac{\mathrm{d}}{\mathrm{d}t} \sigma_{12}\big|_{2=(c,k)}^{1=(v,k)} = -\gamma_k^\lambda p_k + \sum_{k_a} \left[\mathcal{U}_{kk_a}^{vc} p_{k_a} + \mathcal{U}_{kk_a}^{cv} p_{k_a}^* \right] \tag{3.84}$$

where the k-diagonal part can be trivially derived from Eq. 3.78,

$$\gamma_k^\lambda = \frac{1}{2} \sum_{\lambda=1,2} \left[\Gamma_{\lambda k}^{in(cc)} + \Gamma_{\lambda k}^{out(cc)} \right] \tag{3.85}$$

nevertheless deriving non-diagonal in k part(denoted by p_{k_a}) from Eq. 3.78 is more involved, here we elaborate $\mathcal{U}^{vc}_{kk_a}$ in detail,

$$\mathcal{U}^{vc}_{kk_a} p_{k_a} \cdot \frac{\hbar}{\pi} =$$

$$\sum_{bc} V^{ck}_{b}{}^{vk_a}_{c} \hat{V}^{b}_{vk}{}^{c}_{ck_a} [(1-\rho_1)(-p_{k_a})\rho_b\rho_c - \rho_1 p_{k_a}(1-\rho_c)(1-\rho_b)]\delta(\varepsilon_{vk} + \varepsilon_{vk_a} - \varepsilon_c - \varepsilon_b)$$

$$+\sum_{ab} V^{ck}_{b}{}^{a}_{ck_a} \hat{V}^{b}_{vk}{}^{vk_a}_{a} [(1-\rho_1)(1-\rho_a)p_{k_a}\rho_b + \rho_1\rho_a p_{k_a}(1-\rho_b)]\delta(\varepsilon_{vk} + \varepsilon_a - \varepsilon_{ck_a} - \varepsilon_b)$$

$$+\sum_{ac} V^{ck}_{ck_a}{}^{a}_{c} \hat{V}^{vk_a}_{vk}{}^{c}_{a} [(1-\rho_1)(1-\rho_a)\rho_c p_{k_a} + \rho_1\rho_a(1-\rho_c)p_{k_a}]\delta(\varepsilon_{vk} + \varepsilon_a - \varepsilon_c - \varepsilon_{ck_a})$$

$$-\sum_{bc} V^{vk_a}_{vk}{}^{b}_{c} \hat{V}^{ck}_{ck_a}{}^{c}_{b} [(-p_{k_a})(1-\rho_b)\rho_c\rho_2 - p_{k_a}\rho_b(1-\rho_c)(1-\rho_2)]\delta(\varepsilon_{vk_a} + \varepsilon_b - \varepsilon_c - \varepsilon_{ck})$$

$$-\sum_{ac} V^{a}_{vk}{}^{vk_a}_{c} \hat{V}^{ck}_{a}{}^{c}_{ck_a} [(1-\rho_a)(-p_{k_a})\rho_c\rho_2 - \rho_a p_{k_a}(1-\rho_c)(1-\rho_2)]\delta(\varepsilon_a + \varepsilon_{vk_a} - \varepsilon_c - \varepsilon_{ck})$$

$$-\sum_{ab} V^{a}_{vk}{}^{b}_{ck_a} \hat{V}^{ck}_{a}{}^{vk_a}_{b} [(1-\rho_a)(1-\rho_b)p_{k_a}\rho_2 + \rho_a\rho_b p_{k_a}(1-\rho_2)]\delta(\varepsilon_a + \varepsilon_b - \varepsilon_{ck_a} - \varepsilon_{ck})$$

rename all the free indices using b and c,

$$\mathcal{U}^{vc}_{kk_a} p_{k_a} \cdot \frac{\hbar}{\pi} =$$

$$-\sum_{bc} V^{ck}_{b}{}^{vk_a}_{c} \hat{V}^{b}_{vk}{}^{c}_{ck_a} [(1-\rho_1)\rho_b\rho_c p_{k_a} + \rho_1(1-\rho_c)(1-\rho_b)p_{k_a}]\delta(\varepsilon_{vk} + \varepsilon_{vk_a} - \varepsilon_c - \varepsilon_b)$$

$$+\sum_{bc} V^{ck}_{b}{}^{c}_{ck_a} \hat{V}^{b}_{vk}{}^{vk_a}_{c} [(1-\rho_1)(1-\rho_c)p_{k_a}\rho_b + \rho_1\rho_c p_{k_a}(1-\rho_b)]\delta(\varepsilon_{vk} + \varepsilon_c - \varepsilon_{ck_a} - \varepsilon_b)$$

$$+\sum_{bc} V^{ck}_{ck_a}{}^{c}_{b} \hat{V}^{vk_a}_{vk}{}^{b}_{c} [(1-\rho_1)(1-\rho_c)\rho_b p_{k_a} + \rho_1\rho_c(1-\rho_b)p_{k_a}]\delta(\varepsilon_{vk} + \varepsilon_c - \varepsilon_b - \varepsilon_{ck_a}) \qquad \textcolor{red}{b \leftrightarrow c}$$

$$+\sum_{bc} V^{vk_a}_{vk}{}^{b}_{c} \hat{V}^{ck}_{ck_a}{}^{c}_{b} [(1-\rho_b)\rho_c\rho_2 p_{k_a} + \rho_b(1-\rho_c)(1-\rho_2)p_{k_a}]\delta(\varepsilon_{vk_a} + \varepsilon_b - \varepsilon_c - \varepsilon_{ck})$$

$$+\sum_{bc} V^{b}_{vk}{}^{vk_a}_{c} \hat{V}^{ck}_{b}{}^{c}_{ck_a} [(1-\rho_b)\rho_c\rho_2 p_{k_a} + \rho_b(1-\rho_c)(1-\rho_2)p_{k_a}]\delta(\varepsilon_b + \varepsilon_{vk_a} - \varepsilon_c - \varepsilon_{ck})$$

$$-\sum_{bc} V^{b}_{vk}{}^{c}_{ck_a} \hat{V}^{ck}_{b}{}^{vk_a}_{c} [(1-\rho_b)(1-\rho_c)\rho_2 p_{k_a} + \rho_b\rho_c(1-\rho_2)p_{k_a}]\delta(\varepsilon_b + \varepsilon_c - \varepsilon_{ck_a} - \varepsilon_{ck}) \qquad \textcolor{red}{b \leftrightarrow c}$$

$$\tag{3.86}$$

now it is possible to collect the $2 \sim 5^{th}$ rows in the above equation by observing the fact that,

$$\hat{V}^{ck}_{ck_a}{}^{c}_{b} \hat{V}^{vk_a}_{vk}{}^{b}_{c} = V^{ck}_{b}{}^{c}_{ck_a} \hat{V}^{b}_{vk}{}^{vk_a}_{c} + V^{ck}_{ck_a}{}^{c}_{b} \hat{V}^{vk_a}_{vk}{}^{b}_{c}$$

$$\hat{V}^{vk_a}_{vk}{}^{b}_{c} \hat{V}^{ck}_{ck_a}{}^{c}_{b} = V^{vk_a}_{vk}{}^{b}_{c} \hat{V}^{ck}_{ck_a}{}^{c}_{b} + V^{b}_{vk}{}^{vk_a}_{c} \hat{V}^{ck}_{b}{}^{c}_{ck_a} \tag{3.87}$$

therefore,

$$\mathcal{U}^{vc}_{kk_a} p_{k_a} \cdot \frac{\hbar}{\pi} =$$

48

$$- \sum_{bc} \mathrm{V}^{ck\ vk_a}_{b\ \ c} \hat{\mathrm{V}}^{b\ \ c}_{vk\ ck_a} [(1-\rho_1)\rho_b\rho_c p_{k_a} + \rho_1(1-\rho_c)(1-\rho_b)p_{k_a}]\delta(\varepsilon_{vk} + \varepsilon_{vk_a} - \varepsilon_c - \varepsilon_b)$$

$$- \sum_{bc} \mathrm{V}^{b\ \ c}_{vk\ ck_a} \hat{\mathrm{V}}^{ck\ vk_a}_{b\ \ c} [(1-\rho_b)(1-\rho_c)\rho_2 p_{k_a} + \rho_b\rho_c(1-\rho_2)p_{k_a}]\delta(\varepsilon_b + \varepsilon_c - \varepsilon_{ck_a} - \varepsilon_{ck})$$

$$+ \sum_{bc} \hat{\mathrm{V}}^{ck\ \ c}_{ck_a\ b} \hat{\mathrm{V}}^{vk_a\ b}_{vk\ \ c} [(1-\rho_1)(1-\rho_c)\rho_b p_{k_a} + \rho_1\rho_c(1-\rho_b)p_{k_a}]\delta(\varepsilon_{vk} + \varepsilon_c - \varepsilon_b - \varepsilon_{ck_a})$$

$$+ \sum_{bc} \hat{\mathrm{V}}^{vk_a\ b}_{vk\ \ c} \hat{\mathrm{V}}^{ck\ \ c}_{ck_a\ b} [(1-\rho_b)\rho_c\rho_2 p_{k_a} + \rho_b(1-\rho_c)(1-\rho_2)p_{k_a}]\delta(\varepsilon_{ck} + \varepsilon_c - \varepsilon_b - \varepsilon_{vk_a}) \quad (3.88)$$

Finally, $\mathcal{U}^{vc}_{kk_a}$ can be written in a compact way, as well as the counter part $\mathcal{U}^{cv}_{kk_a}$,

$$\mathcal{U}^{vc}_{kk_a} = \sum_{bc} \sum_{\lambda=1,2} \frac{\pi}{\hbar} \Bigg[\quad \hat{\mathrm{V}}^{ck\ \ c}_{ck_a\ b} \hat{\mathrm{V}}^{vk_a\ b}_{vk\ \ c} (\rho_\lambda(1-\rho_b)\rho_c + (1-\rho_\lambda)\rho_b(1-\rho_c))\delta^+$$

$$- \mathrm{V}^{ck\ vk_a}_{b\ \ c} \hat{\mathrm{V}}^{b\ \ c}_{vk\ ck_a} ((1-\rho_\lambda)\rho_b\rho_c + \rho_\lambda(1-\rho_b)(1-\rho_c))\delta^- \Bigg] \quad (3.89)$$

The delta functions read $\delta^+ = \delta(\varepsilon_\lambda - \varepsilon^{\tilde{\lambda}}_{k_a} - \varepsilon_b + \varepsilon_c)$, $\delta^- = \delta(\varepsilon_\lambda + \varepsilon^{\lambda}_{k_a} - \varepsilon_b - \varepsilon_c)$, where $\tilde{\lambda}$ indicates selecting band $\tilde{\lambda} = 2(1)$ if $\lambda = 1(2)$, while $\mathcal{U}^{cv}_{kk_a}$ can be derived by exchanging the band indices c and v.

3.5 Semiconductor-Bloch equations: carrier-phonon

Carrier-phonon interaction is also an efficient scattering channel, when taken into account, the Hamiltonian needs to be modified accordingly,

$$H = H_0 + H_{cf} + H_{cc} + H_{cp} \quad (3.90)$$

where,

$$H_0 = \sum_l \varepsilon_l a^\dagger_l a_l + \sum_n \hbar\omega_n (b^\dagger_n b_n + \frac{1}{2})$$

$$H_{cp} = \sum_{\lambda_1\lambda_2 k} \sum_{jq} \left[g^{\lambda_1\lambda_2 j}_{k,q} a^\dagger_{\lambda_1 k} a_{\lambda_2 k-q} b_{jq} + g^{\lambda_1\lambda_2 j*}_{k,q} a^\dagger_{\lambda_2 k-q} a_{\lambda_1 k} b^\dagger_{jq} \right] \quad (3.91)$$

Turning our attention to the carrier-phonon interaction, we apply the correlation expansion to the expectation value of the phonon-assisted quantities, e.g.

$$\langle a^\dagger_1 a_2 b_u \rangle = \langle a^\dagger_1 a_2 \rangle \langle b_u \rangle \delta_{q,0} + \langle a^\dagger_1 a_2 b_u \rangle^c \quad (3.92)$$

At the lowest order, this expansion gives the product of a carrier density and the expectation value of a single phonon operator $\langle b_u \rangle$ which describes coherent phonons [80, 81, 82]. The condition of spatial homogeneity in the free directions has been used leading to the Kronecker delta. As coherent phonons effects will be neglected in the present work

(their effect being relevant only in the case of a local charge imbalance between electrons and holes[80]), we have to consider the next order by setting up the equations of motion of phonon-assisted correlations, involving expectation values of four operators (the electron-phonon and electron-electron two-particle density matrices). In the spirit of the correlation expansion, these quantities have to be decomposed into all possible lower-order factorizations, leading to the following equations

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u\rangle|_{H_0+H_{cp}} = -\Delta\varepsilon_{12}^-\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u\rangle - \sum_{45}g_{45}^*Q^{1452} \tag{3.93}$$

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u^\dagger\rangle|_{H_0+H_{cp}} = -\Delta\varepsilon_{12}^+\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u^\dagger\rangle + \sum_{45}g_{45}Q^{4215} \tag{3.94}$$

where $\Delta\varepsilon_{12}^\pm = (\varepsilon_1 - \varepsilon_2 \pm \varepsilon_Q)$ and

$$Q^{1234} = \sigma_{12}^\dagger\sigma_{34}n_q^j - \sigma_{12}\sigma_{34}^\dagger(n_q^j+1) \tag{3.95}$$

Formally integrating (3.94) we have

$$\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u\rangle = \frac{i}{\hbar}\sum_{45}g_{45}^*\int_{-\infty}^t Q^{1452}(t')e^{\frac{i}{\hbar}\Delta\varepsilon_{12}^-(t-t')}dt' \tag{3.96}$$

$$\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u^\dagger\rangle = -\frac{i}{\hbar}\sum_{45}g_{45}\int_{-\infty}^t Q^{4215}(t')e^{\frac{i}{\hbar}\Delta\varepsilon_{12}^+(t-t')}dt' \tag{3.97}$$

As usual we separate the fast oscillations contained in Q from the slowly varying part using the ansatz

$$Q^{1234}(t) = \tilde{Q}^{1234}(t)e^{i(\omega_{12}+\omega_{34})t} \tag{3.98}$$

and we apply the Markov limit

$$\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u\rangle = i\pi\sum_{45}g_{45}^*Q^{1452}\delta\left(\Delta\varepsilon_{45}^-\right) \tag{3.99}$$

$$\langle\mathrm{a}_1^\dagger\mathrm{a}_2\mathrm{b}_u^\dagger\rangle = -i\pi\sum_{45}g_{45}Q^{4215}\delta\left(\Delta\varepsilon_{54}^+\right). \tag{3.100}$$

Substituting in (3.49) we get

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\sigma_{12} = \frac{\pi}{\hbar}\sum_{345u}[\,g_{23}g_{45}^*Q^{1453}\delta\left(\Delta\varepsilon_{45}^-\right) - g_{31}g_{45}^*Q^{3452}\delta\left(\Delta\varepsilon_{45}^-\right)$$
$$-g_{32}^*g_{45}Q^{4315}\delta\left(\Delta\varepsilon_{54}^+\right) + g_{13}^*g_{45}Q^{4235}\delta\left(\Delta\varepsilon_{54}^+\right)\,] \tag{3.101}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_\lambda|_{corr}^{(cp)} = \Gamma_\lambda^{in(cp)}(1-\rho_\lambda) - \Gamma_\lambda^{out(cp)}\rho_\lambda \tag{3.102}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_k|_{corr}^{(cp)} = -\gamma_k^{(cp)} p_k + \sum_{k'} \mathcal{U}_{kk'}^{(cp)} p_{k'} \tag{3.103}$$

where the in- and out-scattering rates are defined as

$$\Gamma_\lambda^{in(cp)} = \frac{2\pi}{\hbar} \sum_{\lambda',u} |g_u^{\lambda\lambda'}|^2 \left[\rho_{\lambda'}(n_q^j + 1)\delta(\Delta\varepsilon_{\lambda\lambda'}^+) + \rho_{\lambda'} n_q^j \delta(\Delta\varepsilon_{\lambda\lambda'}^-) \right] \tag{3.104}$$

$$\Gamma_\lambda^{out(cp)} = \frac{2\pi}{\hbar} \sum_{\lambda',u} |g_u^{\lambda\lambda'}|^2 \left[(1 - \rho_{\lambda'})n_q^j \delta(\Delta\varepsilon_{\lambda\lambda'}^+) + (1 - \rho_{\lambda'})(n_q^j + 1)\delta(\Delta\varepsilon_{\lambda\lambda'}^-) \right], \tag{3.105}$$

and the diagonal and off-diagonal dephasing are given by

$$\gamma_k^{(cp)} = \frac{1}{2} \sum_{\lambda=1,2} \left(\Gamma_\lambda^{in(cp)} + \Gamma_\lambda^{out(cp)} \right) \tag{3.106}$$

$$\begin{aligned}
\mathcal{U}_{kk'}^{(cp)} = &\frac{\pi}{\hbar} g_u^{ck,ck'} \left(g_u^{vk,vk'} \right)^* \left[(1 - \rho_{vk})n_q^j + \rho_{vk}(n_q^j + 1) \right] \delta(\Delta\varepsilon_{vk,vk'}^-) \\
&\frac{\pi}{\hbar} g_u^{vk',vk} \left(g_u^{ck',ck} \right)^* \left[\rho_{ck}n_q^j + (1 - \rho_{ck})(n_q^j + 1) \right] \delta(\Delta\varepsilon_{ck',ck}^-) \\
&\frac{\pi}{\hbar} \left(g_u^{ck',ck} \right)^* g_u^{vk',vk} \left[\rho_{vk}n_q^j + (1 - \rho_{vk})(n_q^j + 1) \right] \delta(\Delta\varepsilon_{vk,vk'}^+) \\
&\frac{\pi}{\hbar} \left(g_u^{vk,vk'} \right)^* g_u^{ck,ck'} \left[(1 - \rho_{ck})n_q^j + \rho_{ck}(n_q^j + 1) \right] \delta(\Delta\varepsilon_{ck',ck}^+)
\end{aligned} \tag{3.107}$$

The off-diagonal terms reported above for carrier-phonon scattering are not fully consistent with Section 2.2.3 of Ref.[78] where the phonon matrix elements appear squared. Our formulation is in agreement with Ref.[83].

3.6 Numerical implementation and results

Computing carrier-carrier scattering matrix elements involves high dimensional integrals, we take advantage of the delta function of energy conservation to ease the computational demand.

For example, Eq.(3.89) may lead to a term proportional to $\sum_{abc} V_{vk_b}^{vk} {}_{ck_c}^{ck_a} V_{ck_c}^{ck} {}_{vk_b}^{vk_a} \delta(k)\delta(E)$, which represents a cross product between a direct term and an exchange term(i.e. $\propto V_{q_D}^{cv} V_{q_X}^{cv}$). In order to exploit some geometric properties, we can define intermediate vectors,

$$P = k_1 + k_2$$

$$q_D = k_1 - k_4$$

$$q_X = k_1 - k_3 \tag{3.108}$$

In this case $k_1 = k, k_2 = k_a, k_3 = k_b, k_4 = k_c$, first we integrate over $|k_a|$ and $\angle k_a = \angle(k_2, k_1)$, now we can identify a triangle composed of k_1, k_a, P. Then we choose $|k_3|$ as

the third integration variable, and our aim is to eliminate the integration over $\theta = \angle(P, k_3)$ using the energy delta function, it can be realized through a change of variable,

$$\int \delta[E(\theta)]d\theta = \frac{1}{dE/d\theta} = \frac{dk_4}{dE_4} \cdot \frac{d\theta}{dk_4} \tag{3.109}$$

where dk_4/dE_4 can be computed numerically, and $d\theta/dk_4$ can be computed semi-analytically employing law of cosine. Finally we are left with three dimensional integrals(integration over $|k_a|$, $\angle k_a$, $|k_3|$). Simply put, we first identify the triangle with k_1 as one side and then exploit the delta function to fix the other triangle in the 2D wave vector plane, where the key point is to choose a unique vector P(the common side that the two triangles share).

Regarding carrier-phonon scattering, we proceed with the integration using similar strategy as for the carrier-carrier part. For instance, a typical term arising from Eq.(3.103) reads

$$\frac{\pi}{\hbar} \sum_{q,j} g^{c,c}_{k,q} \cdot g^{v,v*}_{k,q} [(1 - \rho^v_k)p_{k-q}n^j_q + \rho^v_k p_{k-q}(n^j_q + 1)] \cdot \delta(\varepsilon^v_k - \varepsilon^v_{k-q} - h\omega^j_q) \tag{3.110}$$

where p_{k-q} indicates that it is a term non-diagonal in k, the integration over $|q|$ can be eliminated by the energy delta function, at the expense of computing dk/dE.
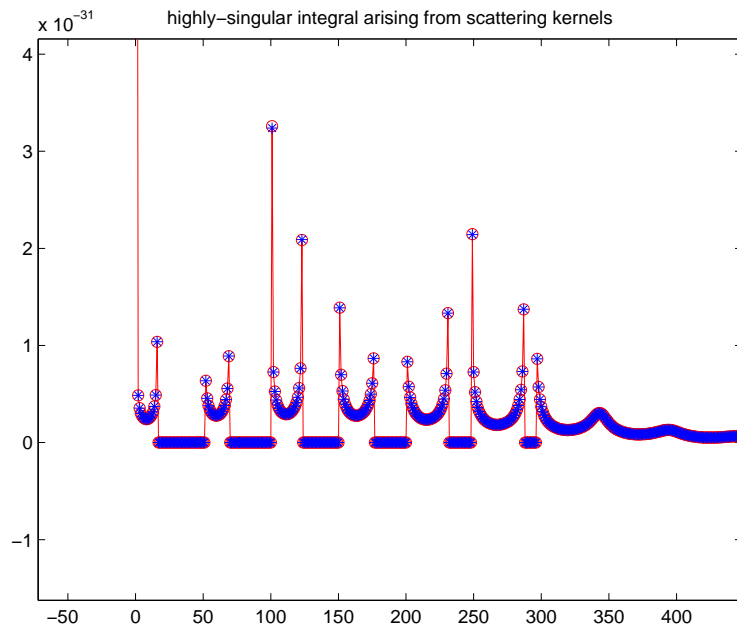

Figure 3.4: Integrals that are strongly singular arise from the electron-electron scattering kernels.

However, the resulting integral turns out to be strongly singular(see Fig. 3.4), which brings substantial difficulty for numerical quadrature. Our solution is to combine different strategies(e.g. Clenshaw-Curtis and Gaussian-Kronrod quadrature), blending them

through homegrown algorithms, the code is implemented in C++ with OpenMP and MPI parallelization.

The original Semiconductor-Bloch equations(Eq. 3.50) consist of a set of time-dependent differential equations, indeed they can be numerically integrated in time domain, such that we can simulate the relaxation dynamics of an excited system. Here we consider the diagonal part of the density matrix, i.e. the carrier density. We excited an electron density distribution with Gaussian profile(the green line in Fig. 3.5), then integrated in time with Runge-Kutta method for 1 pico-second, the excited carrier distribution finally relaxed to the fermi-dirac distribution(indicated by the red dots in Fig. 3.5). Fig. 3.6 shows the time evolution of the momentum-resolved carrier distribution, the carrier will get close to its steady-state in 250 femtoseconds, due to the fact that electron-electron and electron-phonon scattering allow for efficient cooling and redistribution of the hot carriers.
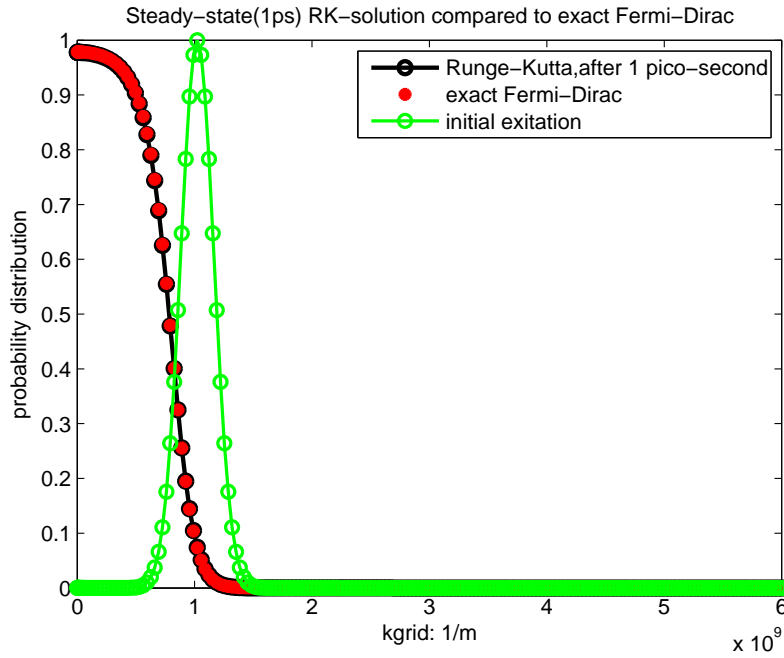
Figure 3.5: Initial and final state of relaxation dynamics.

Now going to the frequency domain, using the slowly varying envelopes $s_{nm,k}$ instead of the fast oscillating polarization,

$$s_{nm,k} = p_{nm,k}e^{-i(k_0z-\omega t-\phi(z))} \tag{3.111}$$

at steady state, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}s_{nm,k} = 0$, Eq. 3.63 can be transformed to

$$s_{nm,k} = \frac{1}{i(\tilde{\omega}_{mn,k}-\omega)+\gamma}\frac{i}{\hbar}\left\{M_{nm}^k\frac{E(z)}{2}+\sum_{k'}V_{mk'\ nk}^{mk\ nk'}p_{k'}\right\}(n_{mk}+n_{nk}-1) \tag{3.112}$$

53
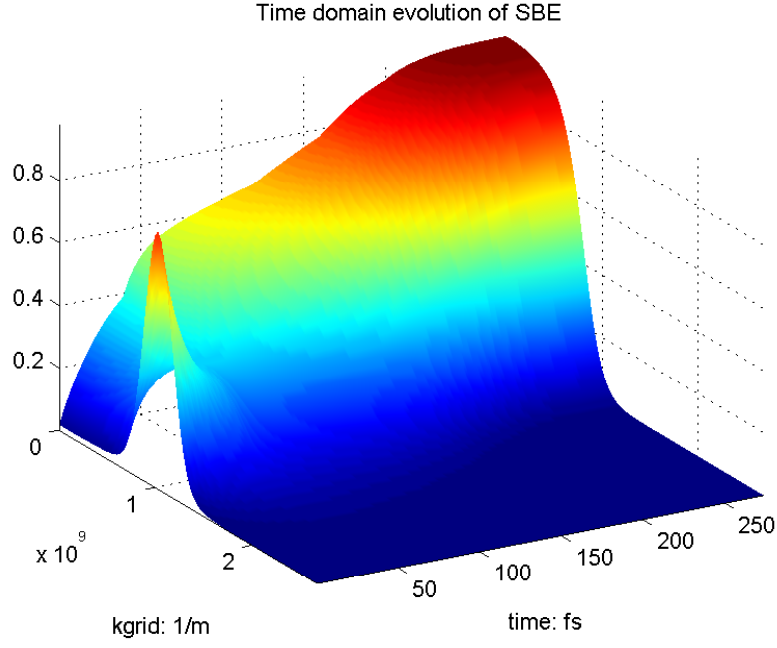
Figure 3.6: Time evolution of the carrier distribution towards steady state.

By defining the field-independent quantity[73]

$$\lambda_{nm,k} = \frac{2s_{nm,k}}{E(z)} \tag{3.113}$$

and

$$\Omega_{nm,k} = \frac{i}{\hbar} \frac{(n_{mk} + n_{nk} - 1)}{i(\tilde{\omega}_{mn,k} - \omega) + \gamma} \tag{3.114}$$

we finally reach a linear system equation for $\lambda_{nm,k}$

$$\Omega_{nm,k} \sum_{k'} \mathrm{V}^{mk}_{mk'} {}^{nk'}_{nk} \lambda_{nm,k'} + \lambda_{nm,k} = -M^k_{nm} \Omega_{nm,k} \tag{3.115}$$

where the scattering term is approximated as a constant decaying rate γ. Notice in the free carrier picture, the term $\mathrm{V}^{mk}_{mk'} {}^{nk'}_{nk}$ will disappear and we simply have

$$\lambda_{nm,k} = -M^k_{nm} \Omega_{nm,k} \tag{3.116}$$

and the optical susceptibility can be calculated by

$$\chi(\omega) = \frac{1}{n^2 \epsilon_0} \frac{P(z)}{E(z)} = \frac{1}{n^2 \epsilon_0} \sum_{nm,k} M^{k*}_{nm} \lambda_{nm,k} \tag{3.117}$$

$$= -\frac{1}{n^2 \epsilon_0} \sum_{nm,k} |M^k_{nm}|^2 \Omega_{nm,k} \tag{3.118}$$

the relation between macroscopic polarization $P(z)$ and microscopic polarization is given in Ref. [77]

$$P(z) = 2e^{-i(k_0 z - \omega t - \phi(z))} \frac{1}{V} \sum_{cv,k} M^{k*}_{cv} p_{vc,k} \tag{3.119}$$

54

Under the slowly varying envelope approximation[73], the classical electric field E satisfies the following equation,

$$\partial_z E(z) - iE(z)\partial_z\phi(z) = i\frac{\mu_0\omega^2}{2k_0}\chi(z)E(z) \tag{3.120}$$

where $k_0 = \frac{2\pi}{\lambda} = \omega n/c$ is the photon wave-number, $\phi(z)$ is the phase shift such that a monochromatic field $E(z,t) = \frac{1}{2}\hat{e}_t E(z)e^{i(k_0z-\omega t-\phi(z))} + c.c$ satisfies the inhomogeneous Helmholtz equation, and ω denotes the field frequency. The real and imaginary parts of Eq. 3.120 can be separated, recalling the macroscopic polarization field $P(z) = \epsilon\chi(z)E(z)$, we now have

$$\partial_z E(z) = \frac{\omega}{2\epsilon_0 nc}\,\text{Im}\{P(z)\} = -\frac{k_0}{2}\chi''(z)E(z) \tag{3.121}$$

$$\partial_z\phi(z) = -\frac{1}{E(z)}\frac{\omega}{2\epsilon_0 nc}\,\text{Re}\{P(z)\} = -\frac{k_0}{2}\chi'(z) \tag{3.122}$$

where optical susceptibility $\chi = \chi' + i\chi''$. Therefore, the intensity gain that characterizes how the amplitude of electric field can be amplified per unit length can be represented by substituting Eq. 3.118

$$G = -k_0\chi'' = \frac{k_0}{n^2\epsilon_0}\,\text{Im}\{\sum_{nm,k}|M_{nm}^k|^2\Omega_{nm,k}\}$$

$$= \frac{\omega}{\epsilon_0 n_b c}\frac{1}{\hbar\mathcal{L}}\sum_{nm,k}|M_{nm}^k|^2(n_{mk}+n_{nk}-1)\frac{\gamma}{(\tilde{\omega}_{mn,k}-\omega)^2+\gamma^2} \tag{3.123}$$

which is equivalent to the original Chuang's result(see Eq. 3.14), as confirmed by the numerical result(denoted as 'FCT' and 'FCT-Chuang', respectively) show in Fig. 3.7. In the above equation, the sum can be replaced with a integration

$$\sum_k = \frac{A}{(2\pi)^2}\int k\,dk\int d\phi \tag{3.124}$$

At Hartree-Fock level, we first have to solve the linear system designated by Eq. 3.115, unphysical absorption(negative gain) can be observed under the gap, this is due to the crude approximation of scattering terms as a constant value γ in Eq. 3.63. This spurious absorption can be corrected in Semiconductor-Bloch equations at second level by substituting γ with a full matrix, where the matrix elements can be computed according to Eq. 3.89. In Fig. 3.7, 'HF' and 'cc' correspond to Hartree-Fock and carrier-carrier scattering inclusions, 'diag' means that only the diagonal part of the Hartree-Fock or carrier-carrier scattering matrix is included. In the case of 'cc-full', both the carrier-carrier scattering and the carrier-phonon scattering are taken into account, the resulting gain spectrum is free from pathological absorption(in Hartree-Fock case) under the band gap.
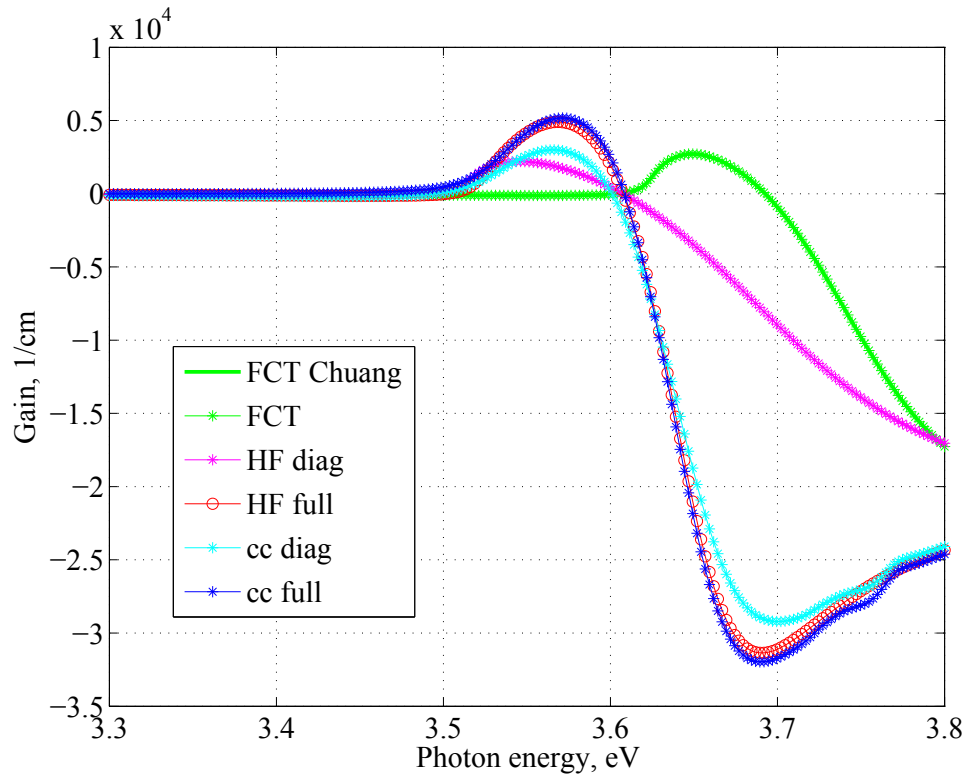
Figure 3.7: 3nm $Al_{0.25}GaN$ Quantum well: TE gain spectra for a sheet charge density of $N_{2D} = 1.0 \times 10^{13} \, cm^{-2}$, free carrier and cases involving scattering mechanism are reported.

In summary, we have solved the semiconductor-Bloch equations in time domain and frequency domain respectively. Solving the equations in time domain amounts to numerically integrate a nonlinear differential equation, in such a way we are able to instrument the ultra-fast relaxation dynamics quantitatively. While in frequency domain, the semiconductor-Bloch equations can be transformed into a linear system, we can obtain the optical gain by numerically solving the matrix equations.

Chapter 4

Towards genuine quantum transport: NEGF

4.1 Overview of classical and quantum transport

Transport of electrons and holes through electronic and optoelectronic devices can be modeled at a number of different levels of sophistication. Numerical simulation and design of GaN-based light-emitting diodes (LEDs) is, at present, largely based on a drift-diffusion (DD) description of carrier transport[84]. However, for nanostructures, DD has some inherent problems, e.g. DD requires a mobility model(see Appendix. D) which is ill-defined in the active region of laser diodes or LEDs.

The intrinsic inability of DD models to properly describe hot-carrier effects could be overcome by directly solving the BTE through MC transport simulation. However, MC shares with all other BTE-derived models the need for semi-empirical quantum corrections for the treatment of quantum interference effects. Moreover, all models based on Boltzmann transport equation (BTE), including not only DD, but also hydrodynamic [85] and particle-based Monte Carlo (MC) descriptions[86, 87], must introduce an artificial separation between "unconfined" (3D) and "bound" (2D) carrier populations, which have to be coupled through the definition of capture/escape rates. In spite of these limitations, "quantum-corrected" DD simulators have been sucessfully applied to the analysis of LEDs and lasers based on conventional III-V materials systems[88].

To fully describe the far-from-equilibrium transport on which technologically relevant optoelectronic devices depend, an accurate model of vertical carrier transport across the

active region of GaN-based LEDs would imply to replace the BTE framework with genuine quantum approaches based, e.g., on the density-matrix (DM) formalism, the nonequilibrium Green's function (NEGF) theory, or the Wigner function picture [89].

Our motivation for heading towards non-equilibrium Green's functions(NEGF) is mainly due to the fact that NEGF considers both (quasi-)bound and current-carrying states on equal footing and appears to be more device-oriented. Approaching the problem from an NEGF perspective means addressing the staggering computational cost required by the calculation of Green's functions, as conventional recursive techniques are not viable when the fully nonlocal carrier-photon interaction is included among the scattering self-energies. Notable DM and NEGF contributions along these lines in the context of transistors[90, 91], solar cells[92] and LEDs[93] must be mentioned alongside with promising projects such as ANGEL[94] and NEMO5[95].

4.2   Scattering states and QTBM

When a quantum device is connected to the outside through contacts or other terminals, a current-carrying states description is preferable to physically depict the device operation. We have primarily dealt with bound states in the previous chapters, now we turn our attention to the current-carrying states that are more relevant from a device-oriented point of view.

After assembling the finite element matrices by summing over the individual mesh element(Eq. 2.65), in general(without the assumption of natural boundary condition), one obtains the matrix equation

$$[A] \cdot \{\zeta\} - E[B] \cdot \{\zeta\} = \{s\} \tag{4.1}$$

where $[A]$ and $[B]$ are $MN \times MN$ sparse finite-element matrices, the column vector $\{\zeta\} = \{\zeta_1\zeta_2...\zeta_M\}^T$ is the finite-element representation of the nanostructure envelope $\zeta(z)$, N is the number of points in real space, and $\{s\}$ is the source term arising from the boundary conditions. Let us assume that the one-dimensional nanostructure is in contact with left and right reservoirs consisting of semi-infinite extensions of the same structure held at chemical potentials V_L and V_R, respectively. Without loss of generality, we consider the case of electrons flowing from left to right, i.e. $-eV_L > -eV_R$. The finite-element equation (Eq. 2.65) will provide the values $\{\zeta_l\}$ of the envelopes $\zeta(z)$ at z_l inside the nanostructure, but the open nature of the problem at hand requires the knowledge of the wavefunction just outside the device, that is, $\{\zeta_1\}$ and $\{\zeta_N\}$ at z_1 and z_N, respectively.

We follow the standard approach[96] to define open boundary conditions by starting from the complex band structure of the reservoirs. It is convenient to write the bulk $M \times M$ $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian in the form,

$$H(k, k_z) = H_0(k) + H_1(k)k_z + H_2(k)k_z^2 \tag{4.2}$$

which explicitly displays its dependence on k_z. In the conventional band structure problem, one fixes (k, k_z) and finds possible values of E by solving the eigenvalue problem $H(k)\phi = E\phi$. In the complex band structure problem, one fixes the in-plane wave-vector k and the energy E, and finds possible values for k_z. This leads to a quadratic eigenvalue problem that can be cast into a linear one

$$\begin{bmatrix} 0 & I \\ H_0 - E \cdot I & H_1 \end{bmatrix} \begin{pmatrix} \phi \\ k_z\phi \end{pmatrix} = k_z \begin{bmatrix} I & 0 \\ 0 & -H_2 \end{bmatrix} \begin{pmatrix} \phi \\ k_z\phi \end{pmatrix} \tag{4.3}$$

Complex eigenvalues occur in pairs, i.e. running (evanescent) states traverse the nanostructure with the same real (imaginary) part and opposite signs. In Fig. 4.1, complex band structure for GaN is reported, the real and imaginary part of the eigenvalues are calculated respectively.
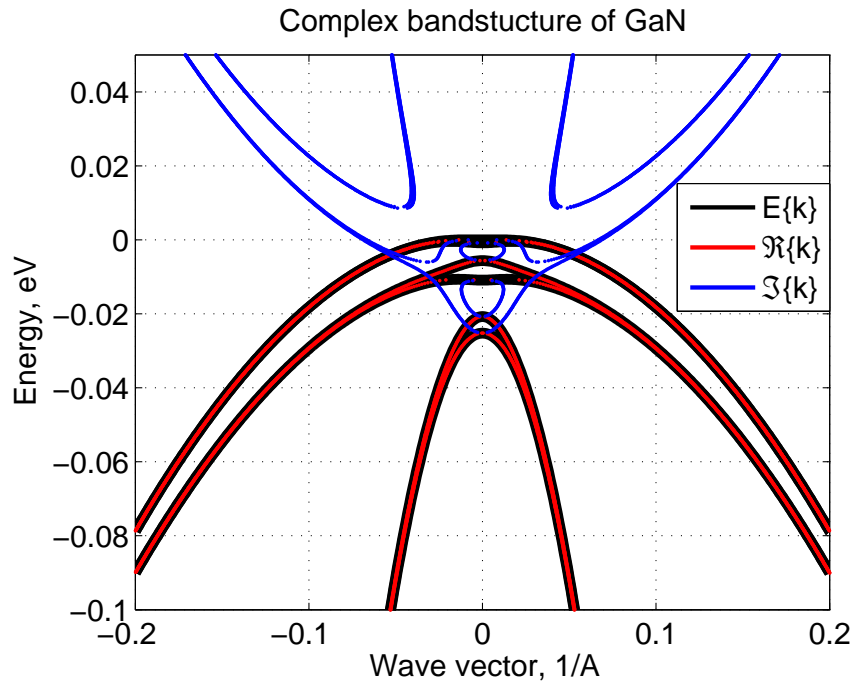

Figure 4.1: Complex band structure for bulk GaN with $k_\perp = 2 \cdot 10^{-8} m^{-1}$. Conventional band structure for valence band is plotted with black line for comparison.

At each lead, we classify the $2M$ solutions of the complex band structure problem according to their direction of motion labelled by the index $\sigma = \pm$. We now consider a

wave injected from the left reservoir in band n, at energy $E_n(k_{Li}^+) = E - eV_L$, propagating from the left with wavenumber k_{Li}^+, and partially reflected into the left contact. The reflected wave must be a linear superposition of all propagating and evanescent waves towards the left at energy $E - eV_L$. These waves are found by selecting only real k_{Ln} such that the group velocity $v_g(k_{Ln}) < 0$ (reflected wave propagating to the left) and these complex k_{Ln} for which $\text{Im}(k_{Ln}) < 0$ (reflected wave decaying to the left). The group velocity $v_g = 1/\hbar \nabla_k E$ can be obtained directly from

$$v_g(k_{Ln}) = 1/\hbar \left(\{\phi\}^\dagger ([H_1] + 2[H_2]k_{Ln})\{\phi\} \right). \tag{4.4}$$

If we inject a bulk wave from the left contact, the continuity of the wavefunction within the first element of the mesh implies

$$\{\zeta_1\} = I\{\phi_{Li}^+\}e^{jk_{Li}^+x_1} + [\phi_{Ln}^-][e^{jk_{Ln}^-x_1}]\{\alpha_{Ln}\} \tag{4.5}$$

$$\{\zeta_2\} = I\{\phi_{Li}^+\}e^{jk_{Li}^+x_2} + [\phi_{Ln}^-][e^{jk_{Ln}^-x_2}]\{\alpha_{Ln}\} \tag{4.6}$$

where $[e^{jk_{Ln}^-x_1}] = \text{diag}(e^{jk_{Ln}^-x_1})$, and $[\phi_{Ln}^-]$ is a $M \times M$ matrix whose columns are the the complex band eigensolutions of the left reservoir propagating or decaying to the right computed at energy $E - eV_L$. Solving (4.6) for the reflection coefficients $\{\alpha_{Ln}\}$ we get

$$\{\alpha_{Ln}\} = [e^{-jk_{Ln}^-x_2}][\phi_{Ln}^-]^{-1}\left(\{\zeta_2\} - I\{\phi_{Li}^+\}e^{jk_{Li}^+x_2} \right) \tag{4.7}$$

Substituting in (4.5) we obtain the boundary condition at the left contact

$$\{\zeta_1\} - [T_L]\{\zeta_2\} = \left([I]e^{jk_{Li}^+x_1} - [T_L]e^{jk_{Li}^+x_2} \right)\{\phi_{Li}^+\} \tag{4.8}$$

where $[I]$ is the identity matrix and $[T_L] = [\phi_{Ln}^-][e^{jk_{Ln}^-(x_1-x_2)}][\phi_{Ln}^-]^{-1}$. Similarly, at the right reservoir, we express the transmitted wave in terms of waves propagating or decaying to the right at energy $E_p(k_{Rn}^+) = E - eV_R$

$$\{\zeta_{N-1}\} = [\phi_{Rn}^+][e^{jk_{Rn}^+x_{N-1}}]\{\beta_{Rn}\} \tag{4.9}$$

$$\{\zeta_N\} = [\phi_{Rn}^+][e^{jk_{Rn}^+x_N}]\{\beta_{Rn}\} \tag{4.10}$$

leading to the boundary condition

$$\{\zeta_N\} - [T_R]\{\zeta_{N-1}\} = 0 \tag{4.11}$$

with $[T_R] = [\phi_{Rn}^+][e^{jk_{Rn}^+(x_N-x_{N-1})}][\phi_{Rn}^+]^{-1}$. The liner system (4.1) augmented by the boundary conditions (4.8) and (4.11) yields the scattering states of the open system. Similarly injection from the right reservoir leads to

$$\{\zeta_1\} = [\phi_{Ln}^-][e^{jk_{Ln}^-x_1}]\{\beta_{Ln}\} \tag{4.12}$$

60

$$\{\zeta_2\} = [\phi_{Ln}^-][e^{jk_{Ln}^- x_2}]\{\beta_{Ln}\} \tag{4.13}$$

$$\{\zeta_{N-1}\} = [\phi_{Rn}^+][e^{jk_{Rn}^+ x_{N-1}}]\{\alpha_{Rn}\} + I\{\phi_{Ri}^-\}e^{jk_{Ri}^- x_{N-1}} \tag{4.14}$$

$$\{\zeta_N\} = [\phi_{Rn}^+][e^{jk_{Rn}^+ x_N}]\{\alpha_{Rn}\} + I\{\phi_{Ri}^-\}e^{jk_{Ri}^+ x_N} \tag{4.15}$$

and the resulting boundary conditions are

$$\{\zeta_1\} - [T_L]\{\zeta_2\} = 0 \tag{4.16}$$

$$\{\zeta_N\} - [T_R]\{\zeta_{N-1}\} = \left([I]e^{jk_{Ri}^- x_N} - [T_R]e^{jk_{Ri}^- x_{N-1}}\right)\{\phi_{Ri}^-\}. \tag{4.17}$$

The above two equations constitute the open boundary condition for a discretized system that supports scattering states. Fig. 4.2 shows the amplitudes of scattering states associated with different valence components when the injected valence states hit a double barrier structure.
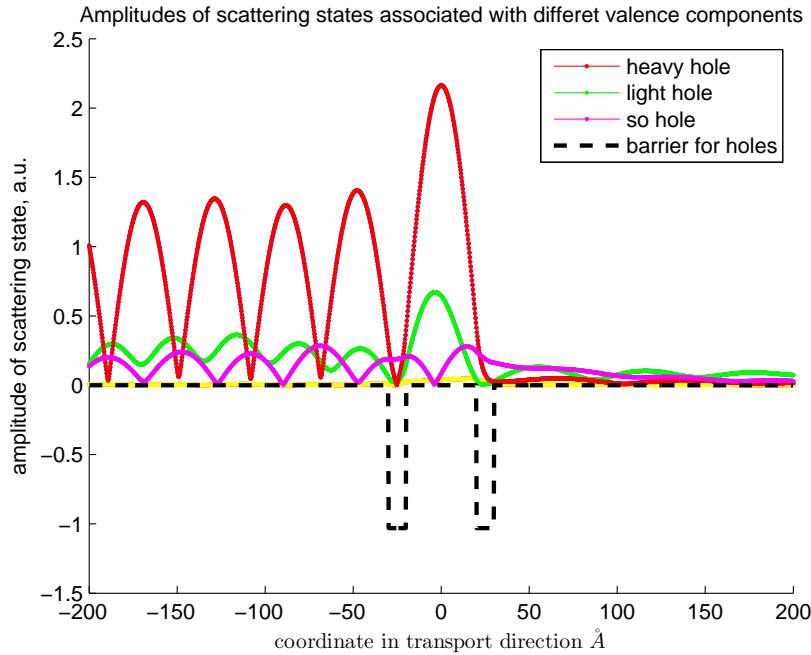
Figure 4.2: Valence states injected from the left reservoir with normal incidence($k_\perp = 0$). Amplitudes of scattering states associated with different valence components(heavy-hole, light-hole, split-off hole) are plotted at certain incident energy.

Besides current-carrying scattering states (i.e. states whose wavefunctions are superpositions of incident and scattered waves), the structure may also support bound states for $E < -eV_R$ (i.e. states with exponentially decaying asymptotic tails at both ends) and resonant states for $E > -eV_R$ (quasi-bound states, i.e wave functions that have only outgoing waves at large distances)[97]. Resonant states corresponding to complex eigenenergies are similar to the leaky modes in optical waveguides. Bound states have real

eigen-energies due to the exponentially decaying wavefunctions in the leads. Both bound and resonant states may be obtained by setting to zero the source term in (4.1) and searching for the roots of the determinant of the resulting homogeneous system. However, it is more convenient to reformulate the problem as an eigenvalue one. The unknown vectors $\{\zeta_1\}$ and $\{\zeta_N\}$ are eliminated by incorporating, in the finite element matrices $[A]$ and $[B]$, the boundary conditions $\{\zeta_1\} = [T_L]\{\zeta_2\}$ and $\{\zeta_N\} = [T_R]\{\zeta_{N-1}\}$ at each lead-domain interface. As $[T_L]$ and $[T_R]$ (i.e. the complex band structure of the leads) depend on the eigenvalue E, the resulting eigenvalue problem is nonlinear.

The idea of scattering states is generalized in the quantum transmitting boundary method(QTBM)[98, 99], which is numerically equivalent to the open boundary condition we have discussed in this section.

4.3   A leap from QTBM to NEGF

In practice, the QTBM represents an intermediate step towards the idea of NEGF. We demonstrate by symmetrizing the QTBM Hamiltonian using effective mass approximation for brevity ($A\zeta = (\hat{H} - E \cdot I)\zeta = s$, s being the source term),

$$
\begin{pmatrix}
1 & -e^{ik_L a} & 0 & \cdots & \cdots & & \cdots & 0 \\
A_{21} & A_{22} & A_{23} & \ddots & & & & \vdots \\
0 & A_{32} & A_{33} & \ddots & \ddots & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & & \ddots & \vdots \\
\vdots & & \ddots & \ddots & A_{N-2,N-2} & A_{N-2,N-1} & 0 \\
\vdots & & & \ddots & A_{N-1,N-2} & A_{N-1,N-1} & A_{N-2,N} \\
0 & \cdots & \cdots & \cdots & 0 & -e^{ik_R a} & 1
\end{pmatrix}
\begin{pmatrix}
\zeta_1 \\ \zeta_2 \\ \vdots \\ \vdots \\ \vdots \\ \zeta_{N-1} \\ \zeta_N
\end{pmatrix}
=
\begin{pmatrix}
I(1 - e^{i2k_L a}) \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 0
\end{pmatrix}
$$

(4.18)

then by multiplying the first row by a constant$(-H_{12} \cdot e^{-ik_L a})$, we have

$$
\begin{pmatrix}
-A_{12} \cdot e^{-ik_L a} & A_{12} & 0 & 0 & \cdots \\
A_{21} & A_{22} & A_{23} & 0 & \cdots \\
0 & A_{32} & A_{33} & A_{34} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\begin{pmatrix}
\zeta_1 \\ \zeta_2 \\ \vdots \\ \vdots
\end{pmatrix}
=
\begin{pmatrix}
-A_{12} \cdot e^{-ik_L a} \cdot (1 - e^{i2k_L a}) \\ 0 \\ \vdots \\ \vdots
\end{pmatrix}
$$

(4.19)

At this point it is possible to perform a Schur transformation[100] of the form,

$$\begin{pmatrix} \alpha & b^T \\ b & \hat{A} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l & I \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \hat{A} - bb^T/\alpha \end{pmatrix} \begin{pmatrix} 1 & l^T \\ 0 & I \end{pmatrix} \qquad (4.20)$$

where, $\hat{A} - bb^T/\alpha$ is the so-called Schur complement, α is often referred to as a pivot and $l = b/\alpha$, in the effective mass case, is a column vector $[-e^{ik_L a} \quad 0, \quad 0, \quad ...]^T$. After we have factorized Eq. 4.19 according to Eq. 4.20, we can multiply each side with the inverse accompanying matrix in the Schur transformation, notice that

$$\begin{pmatrix} 1 & 0 \\ l & I \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -l & I \end{pmatrix} \qquad (4.21)$$

Finally it is interesting to observe that the first node is decoupled from the remaining system of equations, and Eq.(4.18) reduces to

$$\begin{pmatrix} A_{22} + \Sigma^R & A_{23} & 0 & \cdots \\ A_{32} & A_{33} & A_{34} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \zeta_2 \\ \zeta_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} -A_{12} \cdot (1 - e^{i2k_L a}) \\ 0 \\ \vdots \end{pmatrix} \qquad (4.22)$$

where an additional term $\Sigma^R = -[bb^T/\alpha]_{11}$ arises at the diagonal part of the LHS, this is nothing but the so-called retarded boundary self-energy in the NEGF language. In a discretized algebraic system, the boundary self-energy is in analogy to the implementation of the Robin boundary condition in finite element or finite volume analysis, where additional terms emerge at the diagonal in addition to the natural boundary condition. In essence, it constitutes the matching(DtN) between the planar orbital states in the device active region and the extended reservoir modes at the interface of the contacts.

With the above argument, it can be shown that QTBM is exactly equivalent to the NEGF theory in the so-called wave-function(WF) form[90]. In fact, under coherent assumption(ballistic condition), both QTBM and NEGF(WF) are equivalent to the Landauer Bttiker Formalism[101, 102]. In Fig. 4.3, the transmission coefficients are plotted for both the QTBM calculations and analytical theory[103].

In the NEGF theory, the corresponding retarded self-energy can be represented as a matrix consisting of only two elements at the diagonal, assuming 1D structure with effective mass Hamiltonian,
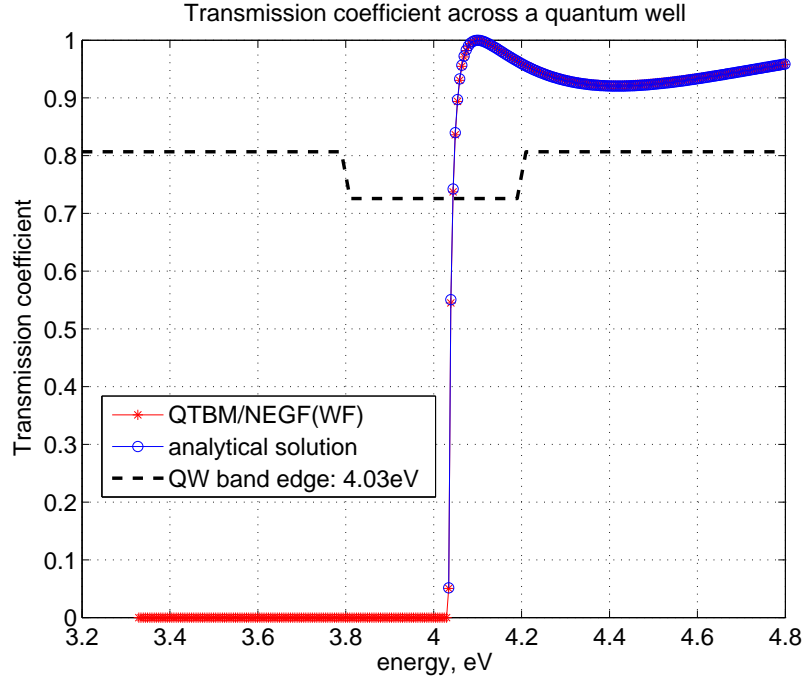
Figure 4.3: Transmission coefficient across a quantum well, numerical results for QTB-M/NEGF(WF) are given, compared to the analytical formulas in [103].

$$\Sigma^R(E) = \begin{pmatrix} \Sigma^R_{11} & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \Sigma^R_{NN} \end{pmatrix} \tag{4.23}$$

where the nonzero elements are

$$\Sigma^R_{11} = [bb^T/\alpha]_{11} = -A_{1,2}exp(ik_L a)$$
$$\Sigma^R_{NN} = [bb^T/\alpha]_{NN} = -A_{N,N-1}exp(ik_R a) \tag{4.24}$$

Therefore the original Schrödinger equation can be written as

$$\{E \cdot I - H - \Sigma^R(E)\} \cdot \{\zeta\} = \{s\} \tag{4.25}$$

where s represents the source term, it is not important in the sense that we are interested in how the system responses to an excitation rather than the strength of the excitation itself.

From a mathematical point of view, the self energies characterize how far the system deviate from its hermitian counterpart, this can be sensed if we split the real and imaginary

part of the self energy

$$H + \Sigma^R(E) = H + \frac{\Sigma^R(E) + \Sigma^R(E)^\dagger}{2} + \frac{\Sigma^R(E) - \Sigma^R(E)^\dagger}{2}$$

$$= \hat{H} + \frac{\Sigma^R(E) - \Sigma^R(E)^\dagger}{2} = \hat{H} - \frac{i}{2}\Gamma \qquad (4.26)$$

where Γ is the so-called broadening matrix, hence the mathematical view reveals the physical sense that Γ determines the carrier exchange rate between the contacts and the device, whereas $\hat{H}$ retains the hermitian property

$$\Gamma = i(\Sigma^R(E) - \Sigma^R(E)^\dagger) \qquad (4.27)$$

$$\hat{H} = H + \text{Re}\{\Sigma^R(E)\} \qquad (4.28)$$

The retarded Green's function is defined as

$$G^R = \{E \cdot I - H - \Sigma^R(E)\}^{-1} \qquad (4.29)$$

where G^R is a matrix of the same size as the Hamiltonian H.

4.4   NEGF in the ballistic picture

We consider a device described by Hamiltonian H connected to two contacts with chemical potential μ_1 and μ_2, as depicted in Fig. 4.4.
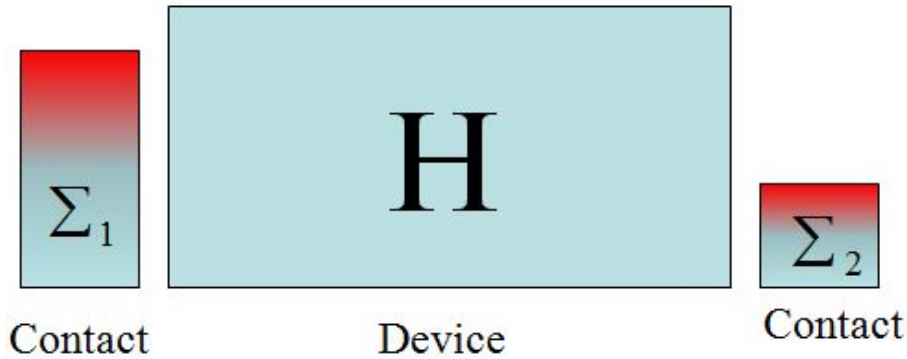

Figure 4.4: A device characterized by Hamiltonian H is in contact with two contacts, with boundary self-energy Σ_1 and Σ_2, respectively .

Within the framework of NEGF, the carrier density can be calculated as

$$n = diag\{\sum_k (\Phi_k \cdot \Phi_k^\dagger)f_k\} \qquad (4.30)$$

notice that $\Phi_k \cdot \Phi_k^\dagger$ is a density matrix while $\Phi_k^\dagger \cdot \Phi_k$ is a scalar, using the definition for Green's function(G^R and G^A stand for the retarded and advanced Green's functions), we have

$$
\begin{aligned}
n &= diag\{\sum_k (G_k^R \cdot s \cdot s^\dagger \cdot G_k^A) f_k\} \\
&= diag\{\int_E \frac{dE}{2\pi}(G^R(E) \cdot \Gamma \cdot G^A(E))F(E - \mu)\} \\
&= -i \cdot diag\{\int_E \frac{dE}{2\pi}G^R(E) \cdot \Sigma^<(E) \cdot G^A(E)\} \\
&= -i \cdot diag\{\int_E \frac{dE}{2\pi}G^<(E)\}
\end{aligned}
\tag{4.31}
$$

The lesser Green's function satisfied the so-called Keldysh equation

$$
G^<(E) = G^R(E) \cdot \Sigma^<(E) \cdot G^A(E)
\tag{4.32}
$$

while

$$
ss^\dagger = (G^R)^{-1} \cdot (-i \cdot G^<) \cdot (G^A)^{-1} = -i\Sigma^< = \Sigma^{in}
\tag{4.33}
$$

It can be observed that $G^R(E) \cdot \Gamma \cdot G^A(E)$ resembles the classical definition of effective density of states(DOS) that contains information about the available states. While in NEGF language, the local density of states(LDOS) is redefined as related to the spectral function A[94],

$$
A = i \cdot \{G^R - G^A\} = G^R(E) \cdot \Gamma \cdot G^A(E)
\tag{4.34}
$$

$$
LDOS = \frac{1}{2\pi} \cdot A
\tag{4.35}
$$

this can be proved by multiplying $(G^R)^{-1}$ and $(G^A)^{-1}$ on the left and right at each side of the equation and realizing that,

$$
(G^A)^{-1} - (G^R)^{-1} = \Sigma - \Sigma^\dagger
\tag{4.36}
$$

where Σ is a shorthand for the retarded boundary self-energy Σ_R, in case of a two terminal device the spectral function has contributions from each contact,

$$
A = A_1 + A_2 = G^R(E) \cdot \Gamma_1 \cdot G^A(E) + G^R(E) \cdot \Gamma_2 \cdot G^A(E)
\tag{4.37}
$$

The function $F(E - \mu)$ in Eq. 4.31 is the Fermi integrals at each contact, under effective mass approximation it can be calculated as

$$
F(E - \mu) = \sum_k f(E - \mu) = \frac{m_{eff}k_B T}{\pi\hbar^2}log\left(1 + exp(-\frac{E - \mu}{k_B T})\right)
\tag{4.38}
$$

Once we have computed the lesser Green's function from Eq. 4.47, the current can be extracted,

$$I = q \cdot Tr\{\frac{d}{dt}\Phi\Phi^\dagger\}$$
$$= q \cdot Tr\{\frac{d\Phi}{dt}\Phi^\dagger + \Phi\frac{d\Phi^\dagger}{dt}\} \tag{4.39}$$

with the time-dependent version of Eq. 4.25

$$I = \frac{q}{i\hbar}Tr\{(H\Phi + \Sigma\Phi + s)\Phi^\dagger - \Phi(\Phi^\dagger H + \Phi^\dagger\Sigma^\dagger + s^\dagger)\}$$
$$= \frac{q}{i\hbar}Tr\{(H\Phi\Phi^\dagger - \Phi\Phi^\dagger H) + (\Sigma\Phi\Phi^\dagger - \Phi\Phi^\dagger\Sigma^\dagger) + (s\Phi^\dagger - \Phi s^\dagger)\}$$
$$= \frac{q}{\hbar}Tr\{(HG^< - G^<H) + (\Sigma G^< - G^<\Sigma^\dagger) - i \cdot (\Sigma^{in}G^A - G^R\Sigma^{in})\} \tag{4.40}$$

taking into account the trace invariant operation $Tr\{A \cdot B\} = Tr\{B \cdot A\}$, the definition for broadening matrix Γ and for spectral function A, as well as Eq. 4.33 the current can be simplified as,

$$I = \frac{q}{\hbar}Tr\{\Sigma^{in}A - \Gamma G^n\}, \quad G^n = i \cdot G^< \tag{4.41}$$

This expression gives information about current flow at a specific terminal m, since each contact in general possesses a different component of Σ^{in} and Γ, at each contact,

$$I_m(E) = \frac{q}{\hbar}Tr\{\Sigma_m^{in}A - \Gamma_m G^n\}, \quad \Sigma_m^{in} = \Gamma_m \cdot f_m(E, \mu_m) \tag{4.42}$$

the current flow through terminal m to terminal n at a specific energy can be computed as

$$I_{mn}(E) = \frac{q}{\hbar}Tr\{\Sigma_m^{in}G^R\Gamma_n G^A - \Gamma_m G^R\Sigma_n^{in}G^A\}$$
$$= \frac{q}{\hbar}Tr\{\Gamma_m G^R\Gamma_n G^A\} \cdot [f_m(E, \mu_m) - f_n(E, \mu_n)]$$
$$= \frac{q}{\hbar}Tr\{G^R\Gamma_n G^A\Gamma_m\} \cdot [f_m(E, \mu_m) - f_n(E, \mu_n)] \tag{4.43}$$

which is equivalent to the Landauer-Bttiker Formalism. Notice that in the effective mass approximation, Γ_n and Γ_m are matrices consisting of only two matrix elements at the corner, in practice, this can ease the computation since the matrix elements will select the corresponding vectors in G^R and G^A, therefore the problem reduces to vector operations.

In the ballistic picture, a self-consistent solution of electrostatics is similar to the case of Poisson-Schrödinger solver, instead of solving the Schrödinger now we are solving for the Green's function(Eq. 4.29) and Keldysh equation(Eq. 4.47), computing the carrier density used as input for Poisson equation in the next iteration.

67

In order to assess the model validity, we first apply it to a nin junction, the material parameters are chosen to be the same as in [94]. Fig. 4.5(a) shows the local density of states in the structure, quantum interference pattern that is not available from classical DD calculations can be clearly observed.

(a) LDOS at 0.2V bias

(b) IV characteristics
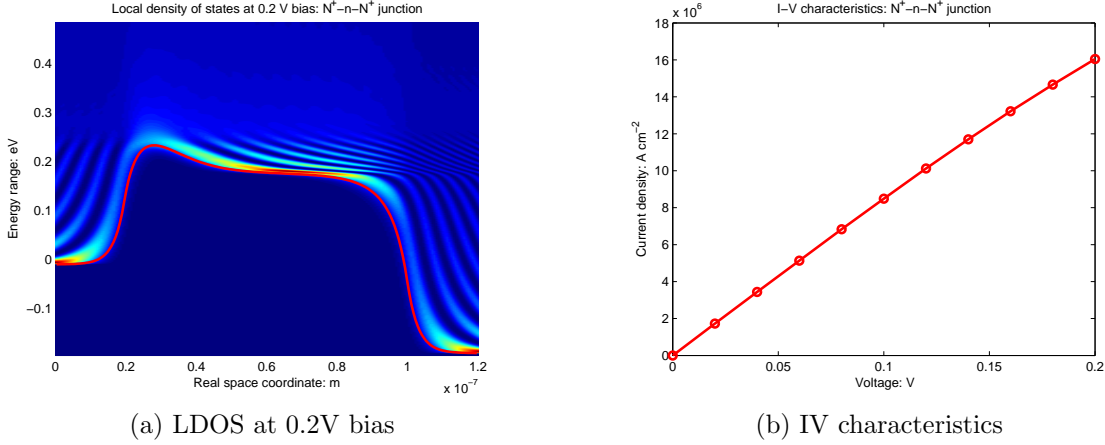
Figure 4.5: LDOS and IV characteristic of a nin junction

In ballistic picture, the carrier density and current are calculated by summing up the contributions from all the energy levels at which electrons could potentially flow through,

$$n = \int \frac{dE}{2\pi} [F(E - \mu_1)A_1 + F(E - \mu_2)A_2] \tag{4.44}$$

where A_1 and A_2 are the left and right LDOS defined in Eq. 4.37, the current can be calculated by further simplifying Eq. 4.43 to the transmission formalism,

$$I = -\frac{q}{h} \int dE T(E)(F(E - \mu_1) - F(E - \mu_2)), \quad T(E) = Tr\{G^R \Gamma_1 G^A \Gamma_2\} \tag{4.45}$$

We have obtained the IV characteristics by scanning the voltage up to 0.2V, the results in Fig. 4.5(b) agree well with [94]. During the ramp-up of the bias, the contacts are allowed to float numerically through the Newman boundary condition. With a modest convergence criteria, the self-consistent calculation is able to converge in less than 10 iterations at each bias point.

We have also applied the coherent NEGF calculation to the resonant tunneling diode, the LDOS and IV characteristics are reported in Fig. 4.6(a) and Fig. 4.6(b), respectively. The resonant states can be visualized and allow the current to be enhanced(at applied bias around $0.2 \sim 0.3$V), this quantum mechanical effect can hardly be captured by classical DD models.

(a) LDOS at $\sim 0.6\text{V}$ bias

(b) IV characteristics

Figure 4.6: LDOS and IV characteristic of a resonant tunneling diode

4.5 NEGF in the scattering picture

In the coherent picture, the available states that carriers could potentially flow through are independent, while in the scattering picture theses channels are coupled through carr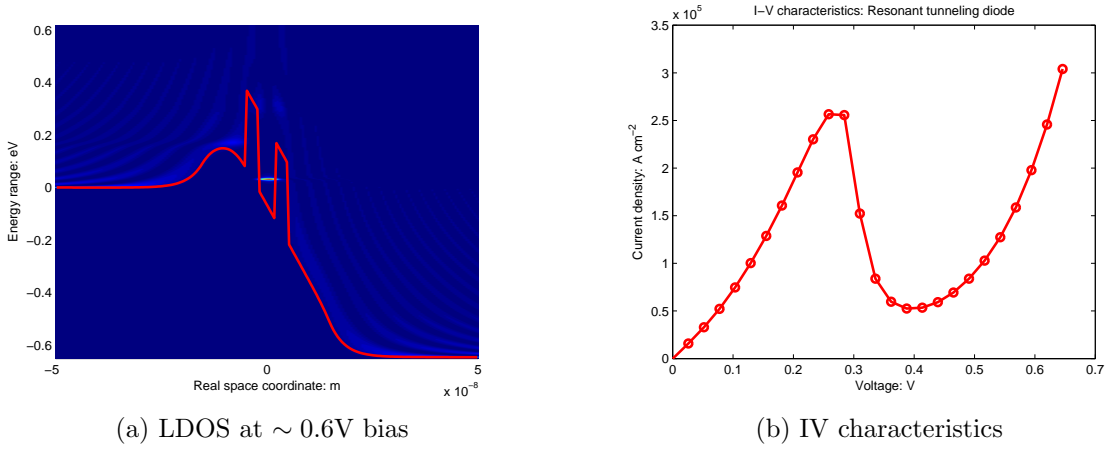ier-phonon and carrier-photon interactions and therefore the energy and momentum can be redistributed among different states. In practice, this means that in(out)-scattering events at (k, E) depend on the Green's functions at other (k', E'). To account for the scattering, formally a scattering self-Energy Σ_S is introduced in the NEGF formalism to augment the boundary self-energy Σ_B in the Kelydysh equation(Eq. 4.47),

$$G^{\gtrless}(E) = G^R(E) \cdot (\Sigma_B^{\gtrless}(E) + \Sigma_S^{\gtrless}(E)) \cdot G^A(E) \tag{4.46}$$

the broadening matrix due to scattering becomes

$$\Gamma_S = i(\Sigma_S^>(E) - \Sigma_S^<(E)) \tag{4.47}$$

which is not directly related to the Fermi functions in the reservoir any more.

In the previous sections, NEGF is introduced in a somewhat heuristic manner, rigorous derivations for self energies require perturbation expansion of the Green's functions and steady state analysis of the quantum kinetic equations[104, 105, 106]. The complete procedure will not be pursued in this work, however, we will explain how the electron-phonon self-energy is introduced to account for dissipative transport in quantum structures. In III-nitride devices, the Polar-optical phonon is a critical relaxation channel, and the self energy can be computed as[94],

$$\Sigma(k, E)_{ph}^{\gtrless} = \frac{q^2 \hbar \omega}{4\pi^2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0}\right) \int_0^{q_t,max} dq_t q_t F(q_t, \triangle, k, q_0)$$

69

$$\cdot \left(N_{ph} G^{\gtrless}(q_t, E \pm \hbar\omega_{ph}) + (N_{ph} + 1)G^{\gtrless}(q_t, E \mp \hbar\omega_{ph}) \right) \qquad (4.48)$$

where $\triangle \equiv x_i - x_j$, F can be precomputed as a toeplitz matrix

$$F(q_t, \triangle, k, q_0) = \int_0^{\pi/a} dq_z cos(q_z\triangle) \left(\frac{1}{\sqrt{(q_z^2 + q_t^2 + q_0^2 + k^2)^2 - 4k^2q_t^2}} \right.$$
$$\left. - q_0^2 \frac{q_z^2 + q_t^2 + q_0^2 + k^2}{\left((q_z^2 + q_t^2 + q_0^2 + k^2)^2 - 4k^2q_t^2\right)^{3/2}} \right) \qquad (4.49)$$
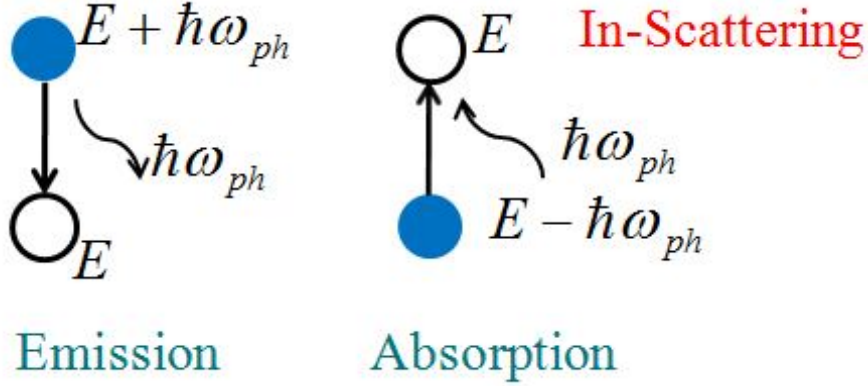
Figure 4.7: Phonon emission and absorption in scatter-in events.
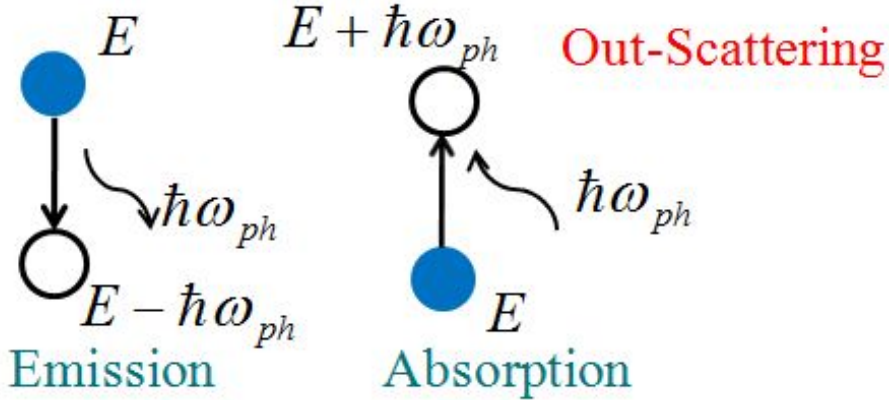
Figure 4.8: Phonon emission and absorption in scatter-out events.

Some simplified electron-phonon scattering models are also proposed[107, 108], they varies in the explicit form of electron-phonon interaction strength(denoted as D_{ph}) as well as it's coupling with energy and momentum. In analogy to the classical Boltzmann transport, the electron-phonon scattering self-energy can be interpreted as composed of scatter-in(Fig. 4.7) and scatter-out(Fig. 4.8) components, both of which consist of an absorption and an emission process. In order to approach the scattering self energy in a physically sensible way, the scatter-in self and scatter-out energies in Eq. 4.48 can be

formally separated,

$$\Sigma_{ph}^<(k, E) = \Sigma_{emi}^<(k, E) + \Sigma_{abs}^<(k, E)$$
$$= D_{ph}(N_{ph} + 1)G^<(E + \hbar\omega_{ph}) + D_{ph}N_{ph}G^<(E - \hbar\omega_{ph}) \qquad (4.50)$$

$\Sigma_{ph}^<$ represents the total scatter-in self-energy, which has contributions from the emission part $\Sigma_{emi}^<$ and the absorption counterpart $\Sigma_{abs}^<$, similarly, the scatter-out self-energy has the form,

$$\Sigma_{ph}^>(k, E) = \Sigma_{emi}^>(k, E) + \Sigma_{abs}^>(k, E)$$
$$= D_{ph}(N_{ph} + 1)G^>(E - \hbar\omega_{ph}) + D_{ph}N_{ph}G^>(E + \hbar\omega_{ph}) \qquad (4.51)$$

At this point, it is meaningful to revisit the derivation of current within the NEGF framework, an alternative expression for the divergence of the electron current is proposed in [109],

$$\frac{\partial J}{\partial x}(x_i, k, E) = \frac{1}{\hbar\triangle_i}(\Sigma^<(k, E)G^>(k, E) - G^<(k, E)\Sigma^>(k, E)) \qquad (4.52)$$

this leads to a natural definition of scattering current, more specifically, the electron-phonon scattering current can be computed as,

$$\frac{\partial J_{ph}}{\partial x}(x_i, k, E) = \frac{1}{\hbar\triangle_i}(\Sigma_{ph}^<(k, E)G^>(k, E) - G^<(k, E)\Sigma_{ph}^>(k, E)) \qquad (4.53)$$

The physical interpretation of Eq. 4.51 according to Fig. 4.8 could be, when an electron at state (k, E) emit a phonon of energy $\hbar\omega_{ph}$, it requires that an un-occupied state $G^>(E - \hbar\omega_{ph})$ is available. Similarly, when an electron at state (k, E) undergoes an absorption process, a state $G^>(E + \hbar\omega_{ph})$ must be available. In case of a phonon emission(scatter-out), the initial state corresponds to the Green's function $G^<(k, E)$ in Eq. 4.53, while the final state in a phonon emission(scatter-in) process would correspond to $G^>(k, E)$.

It is interesting to consider the problem from the point view of conservation, by plugging Eq. 4.50 and Eq. 4.51 into Eq. 4.53, we can obtain a physically ordered combination of Green's functions $G^<G^>$, i.e. from initial state to the final state. In addition, the LHS of Eq. 4.53 would be ideally zero in steady state, therefore indicating that the in/out-scattering processes described by the RHS would be exactly balanced(similar to the idea of detailed balance within the framework of Semiconductor Bloch equation).

A triple-barriers structure is conceived in Fig. 4.9 to interpret NEGF in the scattering picture, where the spatially and energetically resolved current is calculated. The triple-barrier structure is deliberately engineered such that the difference of two primary
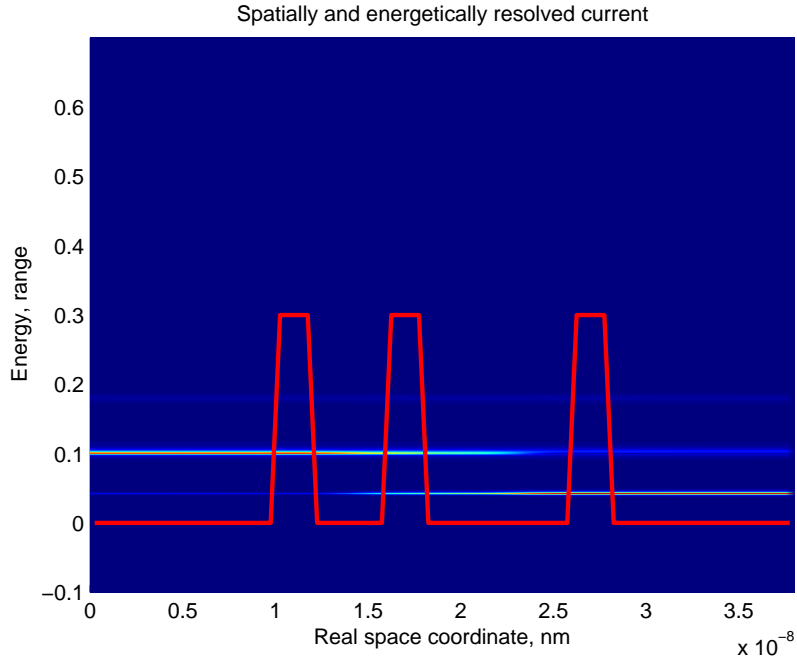
Figure 4.9: A triple-barriers structure similar to the one in[110] is considered, in which the quantized energy levels are engineered to match the optical phonon energy(in GaAs is $\sim$ 60 meV) in order to favor the phonon-assisted tunneling process.

quantized energy levels is close to the optical phonon energy. The carrier-phonon scattering model is included, and the Dyson equation is iterated with the Keldysh equation(See Fig. 4.10) to steady state. The electron current is injected from the left contact(Fermi level $\sim$ 0.1 V), after tunneling through the second barrier, the transport channel is switched to a lower energy one, finally drained at the right contact(Fermi level $\sim$ 0.0 V). This switching of current transport channel indicate that phonon emissions have taken place when the electrons flow from the left to the right contact, representing a concrete example of phonon-assisted tunneling.

The procedure of NEGF calculations in the scattering picture is summarized in Fig. 4.10. It starts with the initial calculation of the boundary self-energy as input to the Dyson's equation(Eq. 4.29). Subsequently the Keldysh's equation(Eq. 4.47) is solved to compute the Green's functions, which are used in turn to determine the self-energies(e.g. Eq. 4.48. The procedure in which self-energies and Green's functions are iterated to self -consistency is called the inner loop, the convergence criteria could be chosen as numerical combination of currents and carrier densities. Once the self-consistency for the inner loop has been achieved, the carrier densities are used for the input of Poisson outer loop, therefore the whole procedure amounts to replacing the Schrödinger part in the Poisson-Schrödinger solver with the inner loop.
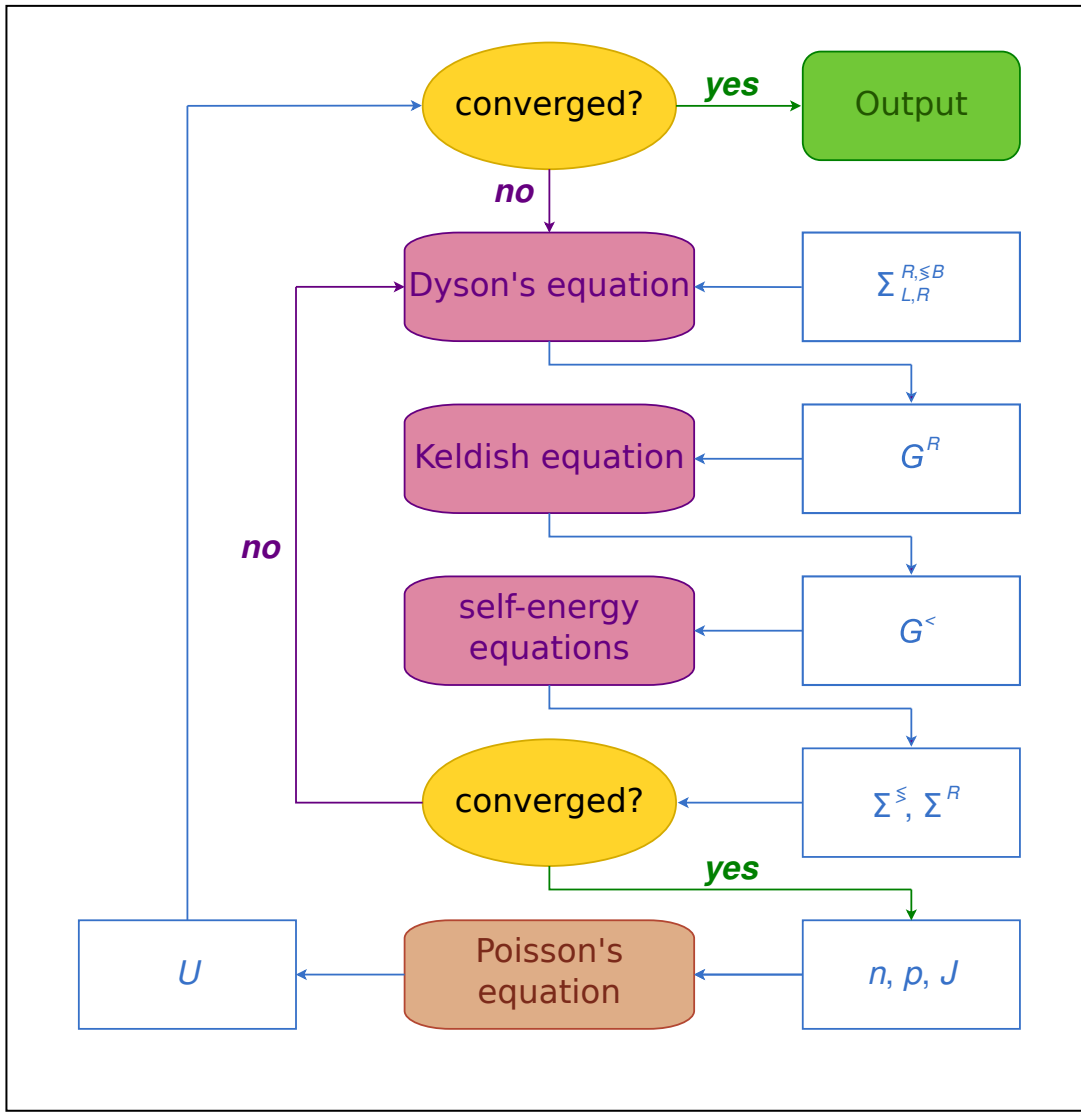
Figure 4.10: Flow Chart of NEGF calculations with both inner and outer loops.

Fig. 4.11 depicts the current across a AlGaN QW at high injection level, electron-optical phonon scattering is included to account for possible vertical transport(carrier-capture and escape from the quantum well). The electron-photon scattering has not been included yet, however since it happens at a very different time scale, this simplified NEGF model is still able to provide some useful insight. For instance, it is possible to extract empirical parameters for the behavioral modeling of excitation-level dependent cross-over current[111].

It is worth noticing that the current in the scattering picture cannot be computed according to Eq. 4.43, which is valid only for the ballistic case. In practice, it's more convenient to use an alternative expression[94, 90],

$$I_{x_n \to x_{n+1}} = \frac{q}{\hbar} \int \frac{dE}{2\pi} Tr\left(H_{n,n+1} G^<_{n+1,n} - G^<_{n,n+1} H_{n+1,n}\right) \tag{4.54}$$
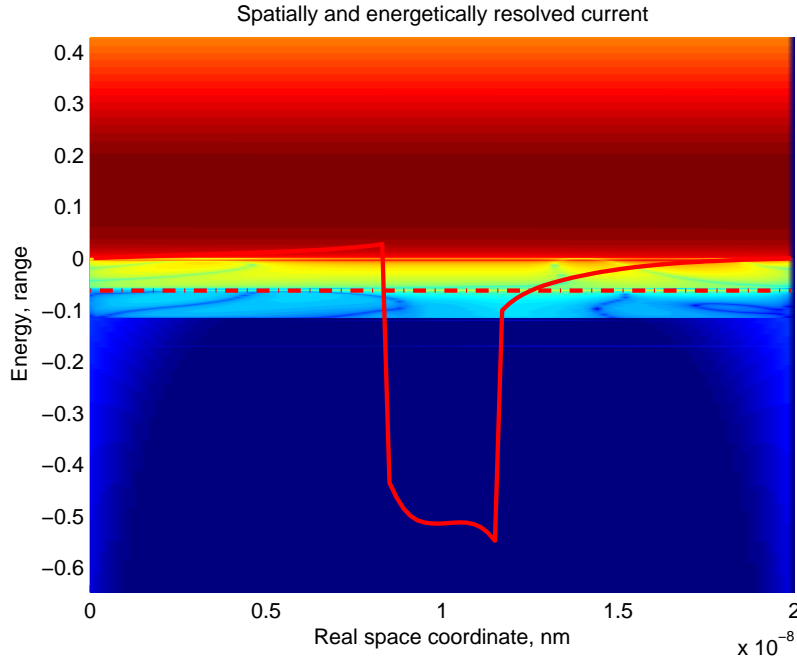
Figure 4.11: Currents transport through AlGaN QW, piezoelectric field is included self-consistently through Poisson-$\mathbf{k} \cdot \mathbf{p}$ solver, optical-phonon scattering is considered. The solid red line represents the conduction band edge, while the dashed line stands for the quasi-fermi level in the quantum well.

This current expression can be demonstrated to be equivalent to several other forms[109], the essence of current computation within the NEGF framework can be revisited from the perspective of conservation law. In traditional transport modeling, the drift-diffusion equations account for carrier conservation and current conservation simultaneously. However, at genuine quantum level, we don't have semi-empirical quantities such as carrier mobility, therefore the current continuity equation is missing. Indeed, in order to reintroduce the current conservation, the price we have to pay for NEGF(with respect to a Poisson-Schrödinger approach) is the self-consistent inner loop which take care of the current and scattering balance at steady state(conceptually, the Schrödinger in a typical Poisson-Schrödinger loop is replaced with the additional inner loop). In practice, an NEGF approach for realistic device simulation would still incorporate a drift-diffusion part that allows to reduce the computational cost, while the NEGF will be used for the more critical region in the device[93, 112].

Chapter 5

Conclusion and future work

We have developed a physics-based multiscale modeling hierarchy to understand photoluminescence and electroluminescence phenomena in III-nitride light emitters. First of all a reliable $\mathbf{k} \cdot \mathbf{p}$ band/subband model derived from NL-EPM full band structure is proposed for the transport calculations in the following parts. Then we try to model the photoluminescence with both Poisson-$\mathbf{k} \cdot \mathbf{p}$ solver and the Semiconductor-Bloch equation solver. In the last part, we address the modeling of electroluminescence from a device oriented point of view, in which case NEGF is chosen instead of quantum corrected semi-classical models. The main advantage of NEGF is its natural incorporation of the open boundary condition, and its ability to describe (quasi-)bound and unbound scattering states on equal footing.

The NEGF model needs to be further extended to account for realistic electron-photon interaction and also electron-electron interaction, therefore increasing the model and computation complexity. However, it is imperative to construct such a complete NEGF tool in order to truly understand efficiency droop[14] that cannot be predicted with traditional simulation techniques.

From a practical point of view, electron-photon and electron-electron interactions could span a large range of energy, this renders the inter-processor network traffic increasing drastically in practice. In the future instead of traditional HPC techniques(e.g. OpenMP, MPI), alternative computation paradigms like Spark(possibly built upon CPU+GPU hybrid architectures) deserve to be considered.

Appendix A

Line element in Finite Element Method

In FEM theory, within each element we have approximated

$$\Psi(x) = \sum_i \psi_i N_i(x) \tag{A.1}$$

N_i here is also called shape functions[64], while the number of nodes in each element coinsides with the number of polynomial expansion, hence eq.(A.1)can be written in matrix form

$$\Psi = \{N\}^T \{\psi\}_e \tag{A.2}$$

Then for a 2-node([0,1],[1,0]) linear element

$$\int_e \{N\}\{N\}^T dx = \frac{l_e}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{A.3}$$

$$\int_e \frac{\partial\{N\}}{\partial x} \frac{\partial\{N\}^T}{\partial x} dx = \frac{1}{l_e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{A.4}$$

$$\int_e \frac{\partial\{N\}}{\partial x}\{N\}^T dx = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \tag{A.5}$$

for a 3-node([0,1],[0.5,0.5],[1,0]) quadratic element

$$\int_e \{N\}\{N\}^T dx = \frac{l_e}{30} \begin{bmatrix} 4 & -1 & 2 \\ -1 & 4 & 2 \\ 2 & 2 & 16 \end{bmatrix}$$

$$\int_e \frac{\partial \{N\}}{\partial x} \frac{\partial \{N\}^T}{\partial x} dx = \frac{1}{3l_e} \begin{bmatrix} 7 & 1 & -8 \\ 1 & 7 & -8 \\ -8 & -8 & 16 \end{bmatrix} \tag{A.6}$$

$$\int_e \frac{\partial \{N\}}{\partial x} \{N\}^T dx = \frac{1}{6} \begin{bmatrix} -3 & 1 & -4 \\ -1 & 3 & 4 \\ 4 & -4 & 0 \end{bmatrix} \tag{A.7}$$

where $\int_e \frac{\partial \{N\}}{\partial x} \{N\}^T dx$ is the transpose of $\int_e \{N\} \frac{\partial \{N\}^T}{\partial x} dx$

Appendix B

Density matrix: Schrödinger and Heisenberg picture

Before discussing density matrix, we would first introduce some properties of the trace operation.

$$
\begin{cases}
d\ tr(X) = tr(dX) \\
tr(A + B) = tr(A) + tr(B) \\
tr(AB) = tr(BA),\ tr(A) = tr(A^T) \\
tr(ABC) = tr(BCA) = tr(CAB)
\end{cases}
\tag{B.1}
$$

Density operator is usually defined in the Schrödinger picture, hence it's dynamics obey the Liouville von Neumann equation

$$
i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\rho = [H, \rho]
\tag{B.2}
$$

the time derivative of the observable $\langle O \rangle = \langle \mathrm{a}_1^\dagger \mathrm{a}_2 \rangle$ reads

$$
\begin{aligned}
\frac{\mathrm{d}\langle O \rangle}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t} tr(\rho O) = \frac{\mathrm{d}}{\mathrm{d}t} tr(\rho\ \mathrm{a}_1^\dagger \mathrm{a}_2) \\
&= tr(\frac{\mathrm{d}\rho}{\mathrm{d}t}\ \mathrm{a}_1^\dagger \mathrm{a}_2 + \rho \cancelto{0}{\frac{\mathrm{d}\mathrm{a}_1^\dagger \mathrm{a}_2}{\mathrm{d}t}})
\end{aligned}
\tag{B.3}
$$

the latter term is canceled because the observable is constant in the Schrödinger picture,

$$
\begin{aligned}
i\hbar\ tr(\frac{\mathrm{d}\rho}{\mathrm{d}t}\ \mathrm{a}_1^\dagger \mathrm{a}_2) &= tr([H, \rho]\ \mathrm{a}_1^\dagger \mathrm{a}_2) \\
&= tr((H\rho - \rho H)\ \mathrm{a}_1^\dagger \mathrm{a}_2) \\
&= tr((H\rho \mathrm{a}_1^\dagger \mathrm{a}_2 - \rho H \mathrm{a}_1^\dagger \mathrm{a}_2) \\
&= tr(\mathrm{a}_1^\dagger \mathrm{a}_2 H\rho - H \mathrm{a}_1^\dagger \mathrm{a}_2 \rho)
\end{aligned}
$$

$$=tr([a_1^\dagger a_2, H]\rho)$$

$$=tr(\rho[a_1^\dagger a_2, H]) = \left\langle [a_1^\dagger a_2, H] \right\rangle \tag{B.4}$$

therefore the equation of motion for $\left\langle a_1^\dagger a_2 \right\rangle$ is,

$$i\hbar \frac{d}{dt} \left\langle a_1^\dagger a_2 \right\rangle = - \left\langle [H, a_1^\dagger a_2] \right\rangle \tag{B.5}$$

Notice that there's a minus sign w.r.t Eq. B.2. While in Heisenberg picture,

$$\frac{d\langle O \rangle}{dt} = \frac{d}{dt} tr(\rho O) = \frac{d}{dt} tr(\rho\, a_1^\dagger a_2)$$

$$= tr(\frac{d\rho}{dt} a_1^\dagger a_2 + \rho \frac{da_1^\dagger a_2}{dt}) = \frac{d}{dt} \left\langle a_1^\dagger a_2 \right\rangle \tag{B.6}$$

here $d\rho/dt = 0$ since ρ is not an observable(in the author's opinion, it's a constant "state") in the Heisenberg picuture, then Eq. B.5 can be derived by directly substituting the Heisenberg equation of motion, leading to the usual equation of motion in operator form,

$$i\hbar \frac{d}{dt} a_1^\dagger a_2 = - [H, a_1^\dagger a_2] \tag{B.7}$$

Therefore it is proved that either Schrödinger or Heisenberg picture gives the same expectation values. Throughout this work, we'll use the Heisenberg picture because it turns out to be more convenient in the many-body treatment.

Appendix C

Broyden iteration

In newton's method, the residual vector of the subsequent iteration could be written as,

$$\left| F^{(l+1)} \right\rangle \approx \left| F^{(l)} \right\rangle + J^{(l)} \left(\left| x^{(l+1)} \right\rangle - \left| x^{(l)} \right\rangle \right) \tag{C.1}$$

where $J^{(l)}$ is the Jacobian, the computational cost of calculating the Jacobian is prohibitive for high-dimensional problems, e.g. density functional calculations. Instead in the Broyden method, one do not have to compute explicitly the Jacobian matrix. In the present case the Broyden's second method is employed, which is based on updating the approximate inverse Jacobian,

$$\left| \Delta x^{(l)} \right\rangle = G^{(l+1)} \left| \Delta F^{(l)} \right\rangle . \tag{C.2}$$

the Broyden criteria leads to the following update formula for the inverse Jacobian,

$$G^{(l+1)} - G^{(l)} = \frac{1}{a_{ll}} \left(\left| \Delta x^{(l)} \right\rangle \left\langle \Delta F^{(l)} \right| - G^{(l)} \left| \Delta F^{(l)} \right\rangle \left\langle \Delta F^{(l)} \right| \right) \tag{C.3}$$

here the matrix a is the overlap matrix of the difference residual vectors, i.e.,

$$a_{nm} = \left\langle \Delta F^{(m)} \mid \Delta F^{(n)} \right\rangle \tag{C.4}$$

Again for the first iteration we adopt the simple mixing according to $G^{(1)} = -\beta \cdot I$, usually $\beta = 0.5$ and I is the identity matrix, notice that Broyden update leads to a dense matrix which requires to be stored at each iteration.

More sophisticated variants of Broyden's method have been developed in recent years, which also help advance the quantum chemistry community in various fronts of density functional computations[113, 114].

Appendix D

Drift-diffusion implementation: an 1D code

Drift-diffusion theory has been around for more than 30-years[84], however, a straightforward and self-contained implementation of the Scharfetter-Gummel[115] method appears to be unavailable publicly as far as the author has noticed. Therefore we think it's valuable to provide an open-sourced implementation, which may be of help to understand classical transport theory and the motivation for quantum transport. In this case, Finite element method(FEM) and finite volume method(FVM) have been employed to discretize a PN junction.

```
1  %1D Drift-diffusion solver, by Xiangyu Zhou@PoliTo, 01.2015
2  %
3  %Run the code for a biased PN junction
4  %mesh=pn_1d(1.5e16,1.5e16,0.6);
5  function mesh = pn_1d(dop_n,dop_p,bias)
6  q=1.602e-19;T=300;
7  Kb=1.3807*1e-23;
8  vt=Kb*T/q;                                      %KbT
9  epsi=11.7*8.854*1e-14;                          %Si
10 Eg=1.08;                                        %Si_300K_1.12eV
11 ni=1.45*1e10;                                   %Si,intrinsic
12 tau=1e-10;                                      %->recombination const
13 N_V=1.04e19;
14 %
15 mesh.N=501;                                     %discretization
16 mesh.L=2e-6*100;                                %2um->2e^-6m->cm
17 mesh.le=mesh.L/(mesh.N-1);
18 mesh.Le=mesh.le*ones(1,mesh.N-1);
19 V_ref=vt*log(dop_n/ni);                         %V_ref
20 mesh.EF=vt*log(N_V/dop_p);
21 N_A=dop_p*[ones(1,(mesh.N-1)/2)  zeros(1,(mesh.N+1)/2)]';
22 N_D=dop_n*[zeros(1,(mesh.N-1)/2) ones(1,(mesh.N+1)/2) ]';
23 %--------------------------------------------->Begin Equilibrium case
```

```matlab
24 phi_0=[(V_ref-vt*log(dop_p/ni))*ones((mesh.N-1)/2,1);...
25      (V_ref+vt*log(dop_n/ni))*ones((mesh.N+1)/2,1)];  phi=phi_0;
26 %
27 in1=1:(mesh.N-1);in2=2:mesh.N;
28 inr12=[in1 in1 in2 in2];
29 inc12=[in1 in2 in1 in2];
30 NxNx11=1./mesh.Le; NxNx12=-1./mesh.Le; NxNx22=1./mesh.Le;
31 Amat=sparse(inr12,inc12,epsi*[NxNx11 NxNx12 NxNx12 NxNx22],mesh.N,mesh.N);
32 Amat(1,:)=0; Amat(mesh.N,:)=0;
33     %
34     conver_eq=0;iter=0;
35     while(~conver_eq)
36     iter=iter+1;
37     rho=mesh.le*q*(N_D-N_A-2*ni*sinh((phi-V_ref)/vt));
38     E=sparse(1:mesh.N,1:mesh.N, 2*mesh.le*q*ni/vt*cosh((phi-V_ref)/vt) );
39     E(1,1)=1;E(mesh.N,mesh.N)=1;
40     R_re=-(Amat*phi-rho);
41     Jacob=Amat+E;
42     [LL,UU,PP,QQ]=lu(Jacob);
43     delta_phi=QQ*(UU\(LL\(PP*R_re)));
44     %
45     if(norm(delta_phi)/norm(phi)<=1e-9)
46     conver_eq=1; end
47     %
48     phi=phi+delta_phi;
49     end
50 mesh.phi=phi';
51 mesh.elecf=-diff(mesh.phi)/mesh.le;
52 mesh.nn=ni*exp(+(phi-V_ref)/vt);
53 mesh.np=ni*exp(-(phi-V_ref)/vt);
54 rhot=(N_D'-N_A'-2*ni*sinh((mesh.phi-V_ref)/vt));
55 plot_equi(rhot,Eg,mesh);
56 %--------------------------------------------------->Begin non-Equilibrium
57 vstep=0.5*vt;ivm=ceil(bias/vstep);vbias=ones(1,ivm);
58 for iv = 1:ivm
59     conver_neq=0;iter=0;
60     mesh.phi(end)=mesh.phi(end)-vstep;
61     %
62   while(~conver_neq)
63     iter=iter+1;
64     fprintf('ibias: %d | iter: %d\n',iv,iter)
65     mesh.elecf=-diff(mesh.phi)/mesh.le;
66     mesh=comput_mobility(mesh,vt);
67
68     %%%begin Continuity,FVM
69     in0=1:mesh.N;in1=1:(mesh.N-1);in2=2:mesh.N;
70     inr12=[in0 in1 in2];
71     inc12=[in0 in2 in1];
72     %
73     nnw=q*mesh.Dn.*BER1(-diff(mesh.phi)/vt)/mesh.le/mesh.le;
74     nne=q*mesh.Dn.*BER1( diff(mesh.phi)/vt)/mesh.le/mesh.le;
75     Vw=[nnw(1:end-1) 0]; Ve=[0 nne(2:end)];
76     nnc=-(nnw(2:end)+nne(1:end-1)); Vc=[1 nnc 1];
```

```matlab
77      Nmat=sparse(inr12,inc12,[Vc Ve Vw],mesh.N,mesh.N);
78      %
79      npw=q*mesh.Dp.*BER1( diff(mesh.phi)/vt)/mesh.le/mesh.le;
80      npe=q*mesh.Dp.*BER1(-diff(mesh.phi)/vt)/mesh.le/mesh.le;
81      Vw=[npw(1:end-1) 0]; Ve=[0 npe(2:end)];
82      npc=-(npw(2:end)+npe(1:end-1)); Vc=[1 npc 1];
83      Pmat=sparse(inr12,inc12,[Vc Ve Vw],mesh.N,mesh.N);
84      %
85      Gn= q*(mesh.nn.*mesh.np-ni*ni)./(tau*(mesh.nn+ni)+tau*(mesh.np+ni));
86      Gp= q*(mesh.nn.*mesh.np-ni*ni)./(tau*(mesh.nn+ni)+tau*(mesh.np+ni));
87      Gn(1)=mesh.nn(1);Gn(end)=mesh.nn(end);
88      Gp(1)=mesh.np(1);Gp(end)=mesh.np(end);
89      mesh.nn=Nmat\Gn;
90      mesh.np=Pmat\Gp;
91
92      %%%begin Poisson,FEM/FDM
93      Apmat=Amat+ sparse(1:mesh.N,1:mesh.N,...
94              q*mesh.le/vt*(mesh.nn+mesh.np) );
95      bp=mesh.le*q*(N_D-N_A+mesh.np-mesh.nn)+...
96              q*mesh.le/vt*(mesh.nn+mesh.np).*mesh.phi';
97      mesh.oldphi=mesh.phi;
98      Apmat(1,1)=1;Apmat(mesh.N,mesh.N)=1;
99      bp(1)=mesh.phi(1);bp(end)=mesh.phi(end);
100     phi=Apmat\bp;delta_phi=phi'-mesh.oldphi;
101     if(norm(delta_phi)/norm(mesh.phi)<=1e-6)
102     conver_neq=1;vbias(iv)=vstep*iv;
103     fprintf('converged! bias=%f V\n',vbias(iv));
104     end
105     mesh.phi=phi';
106   end
107
108     %%%elemental current
109     Jn1=q/mesh.le*mesh.Dn.*(BER1( diff(mesh.phi)/vt).*mesh.nn(2:end)'-...
110                             BER1(-diff(mesh.phi)/vt).*mesh.nn(1:end-1)');
111     Jp1=q/mesh.le*mesh.Dp.*(BER1(-diff(mesh.phi)/vt).*mesh.np(2:end)'-...
112                             BER1( diff(mesh.phi)/vt).*mesh.np(1:end-1)');
113     mesh.Jnx=Jn1;
114     mesh.Jpx=Jp1;
115     mesh.Jtot=mesh.Jnx-mesh.Jpx;
116     mesh.Jv(iv)=mesh.Jtot(1);
117     mesh.Jn(iv)=mesh.Jnx(1);
118     mesh.Jp(iv)=mesh.Jpx(1);
119
120 end   %--------------------------------------------->End Non-Equilibrium
121
122 mesh.EFn=-mesh.phi+vt*log(mesh.nn'/ni);
123 mesh.EFp=-mesh.phi-vt*log(mesh.np'/ni);
124 rhot=N_D-N_A+mesh.np-mesh.nn;
125 plot_nonequi(vbias,rhot,Eg,mesh);
126 figure(3),hold on
127 plot(mesh.Jnx,'-r') %I-x
128 plot(-mesh.Jpx,'-b')
129 plot(mesh.Jtot,'-.g')
```

```matlab
130  legend('Jn','Jp','Jtot')
131  title('Converged current,A/cm^2','fontweight','bold')
132
133  %---------------------------------------------------------------------%
134
135  function mesh = comput_mobility(mesh,vt)
136  % Mobility: Thomas model | ATLAS: VSAT = (2.4*10^7)/(1+0.8*exp(T/600))
137  vsatn=1.07e7;vsatp=8.37e6;betan=2; betap=1; %cm/s
138  mobn0=960;mobp0=435;% ~2e16 doping level dependent,cm^2/V/sec
139  mesh.mobn=mobn0./((1+(mobn0.*abs(mesh.elecf)/vsatn).^betan).^(1/betan));
140  mesh.mobp=mobp0./((1+(mobp0.*abs(mesh.elecf)/vsatp).^betap).^(1/betap));
141  mesh.Dn=mesh.mobn*vt;
142  mesh.Dp=mesh.mobp*vt;
143
144  %---------------------------------------------------------------------%
145
146  function B=BER1(x)
147  % according to the formula given in 1984Selberherr, p.169.
148  % The parameters x1,...,x5 are evaluated for MATLAB.
149  % FB October 03, 2007
150
151  % Defines the nodes for the approximation
152  x1=-36.25; x2=-7.63e-6; x3=-x2; x4=32.92; x5=36.5;
153  B=zeros(size(x));
154  B1 = (x<=x1);
155  B(B1) = -x(B1);
156  B2= (x>x1)&(x<x2);
157  B(B2) = +  x(B2)./(exp(x(B2))-1+1.e-99);
158  B3 = (x>=x2)&(x<=x3);
159  B(B3) = 1-x(B3)/2;
160  B4 = (x>x3)&(x<x4);
161  B(B4) = x(B4).*exp(-x(B4))./(1-exp(-x(B4))+1.e-99);
162  B5= (x>=x4)&(x<x5);
163  B(B5) =   x(B5).*exp(-x(B5));
164  B6= (x>=x5);
165  B(B6) = 0.0;
166
167  %---------------------------------------------------------------------%
168
169  function [] = plot_nonequi(vbias,rhot,Eg,mesh)
170  x_ratio=1e-6/(mesh.le/100);    %1e-6->um|le/100:cm->um
171  N=mesh.N;
172  %
173  figure(2),
174  subplot(2,2,1),
175  plot(vbias,mesh.Jv,'b','linewidth',1.5);
176  set(gca,'yscale','log')
177  xlabel('bias voltage, V','fontweight','bold')
178  ylabel('total current, A/cm^2','fontweight','bold')
179  title('Forward biased PN: IV curve','fontweight','bold')
180  %
181  subplot(2,2,2)
182  plot(mesh.elecf,'b','linewidth',1.5)
```

```matlab
183  xlabel('position in 1D,um','fontweight','bold')
184  ylabel('intensity V/cm','fontweight','bold')
185  title('electric field,V/cm','fontweight','bold')
186  set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
187  set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
188  %
189  subplot(2,2,3),hold on
190  plot(mesh.nn,'g','linewidth',1.5)
191  plot(mesh.np,'b','linewidth',1.5)
192  plot(rhot,'--r','linewidth',1.5)
193  legend('electron','hole','total charge')
194  xlabel('position in 1D,um','fontweight','bold')
195  ylabel('concerntration cm_-3 (log)','fontweight','bold')
196  title('carrier density','fontweight','bold')
197  set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
198  set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
199  %
200  subplot(2,2,4),hold on
201  p1=plot(-mesh.phi+Eg,'m','linewidth',1.5);
202  p2=plot(-mesh.phi    ,'m','linewidth',1.5);
203  p3=plot(-mesh.phi+Eg/2,'--k','linewidth',1.5);%approximately meffe!=meffh
204  p4=plot( mesh.EFn+Eg/2,'r','linewidth',1.5);
205  p5=plot( mesh.EFp+Eg/2,'b','linewidth',1.5);
206  legend([p1 p2 p3 p4 p5],'Ec','Ev','EFi','EFn','EFp')
207  xlabel('position in 1D,um','fontweight','bold')
208  ylabel('Energy level eV','fontweight','bold')
209  title('band diagram','fontweight','bold')
210  ylim([-.5 1.5]);
211  set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
212  set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
213
214  %----------------------------------------------------------------%
215
216  function [] = plot_equi(rhot,Eg,mesh)
217  x_ratio=1e-6/(mesh.le/100);    %1e-6->um|le/100:cm->um
218  N=mesh.N;
219  %
220  figure(1),
221  subplot(2,2,1)
222  plot(mesh.phi,'b','linewidth',1.5)
223  xlabel('position in 1D,um','fontweight','bold')
224  ylabel('potential V','fontweight','bold')
225  title('Equilibrium PN: Potential','fontweight','bold')
226  set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
227  set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
228  %
229  subplot(2,2,2)
230  plot(mesh.elecf,'b','linewidth',1.5)
231  xlabel('position in 1D,um','fontweight','bold')
232  ylabel('intensity V/cm','fontweight','bold')
233  title('electric field,V/cm','fontweight','bold')
234  set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
235  set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
```

```matlab
236 %
237 subplot(2,2,3),hold on
238 %plot(rhoabs,'b','linewidth',1.5)
239 plot(mesh.nn,'g','linewidth',1.5)
240 plot(mesh.np,'b','linewidth',1.5)
241 plot(rhot,'--r','linewidth',1)
242 legend('electron','hole','total charge')
243 xlabel('position in 1D,um','fontweight','bold')
244 ylabel('concerntration cm_-3 (log)','fontweight','bold')
245 title('carrier density','fontweight','bold')
246 set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
247 set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
248 %
249 subplot(2,2,4),hold on
250 p1=plot(-mesh.phi,'m','linewidth',1.5);
251 p2=plot(-mesh.phi+Eg,'b','linewidth',1.5);
252 p3=plot([0 N],[mesh.EF mesh.EF],'--r','linewidth',1.5);
253 legend([p1 p2 p3],'Ec','Ev','EF')
254 xlabel('position in 1D,um','fontweight','bold')
255 ylabel('Energy level eV','fontweight','bold')
256 title('band diagram','fontweight','bold')
257 ylim([-1 1.5]);
258 set(gca,'xlim',[0,N],'Xtick',0:(N-1)/4:N-1);
259 set(gca,'xticklabel',str2double(get(gca,'XTickLabel'))/x_ratio);
```

Bibliography

[1] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Philadelphia: Saunders, 1976.

[2] J. H. Ryou, P. D. Yoder, J. Liu, Z. Lochner, H. Kim, S. Choi, H. J. Kim, and R. D. Dupuis, "Control of quantum-confined stark effect in InGaN-based quantum wells," *IEEE J. Select. Topics Quantum Electron.*, vol. 15, no. 4, p. 1080, 2009.

[3] M. Sumiya and S. Fuke, "Review of polarity determination and control in gan," *MRS Internet J. Nitride Semicond. Res.*, vol. 9, no. 1, 2004.

[4] J. C. Phillips, "Ionicity of the chemical bond in crystals," *Rev. Modern Phys.*, vol. 42, no. 3, pp. 317–356, 1970.

[5] M. L. Cohen and V. Heine, "The fitting of pseudopotentials to experimental data and their subsequent application," in *Solid State Physics. Advances in Research and Applications* (H. Ehrenreich, F. Seitz, and D. Turnbull, eds.), vol. 24, pp. 37–248, New York: Academic Press, 1970.

[6] M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*. Berlin: Springer-Verlag, 1988.

[7] A.-B. Chen and P. Srichaikul, "Shallow donor levels and the conduction band edge structures in polytypes of SiC," *Phys. Status Solidi B*, vol. 202, pp. 81–106, 1997.

[8] P. Friedel, M. S. Hybertsen, and M. Schlüter, "Local empirical pseudopotential approach to the properties of Si/Ge superlattices," *Phys. Rev. B*, vol. 39, pp. 7974–7977, Apr. 1989.

[9] S. K. Pugh, D. J. Dugdale, S. Brand, and R. A. Abram, "Electronic structure calculations on nitride semiconductors," *Semiconductor Sci. Tech.*, vol. 14, pp. 23–31, 1999.

[10] M. Goano, E. Bellotti, E. Ghillino, G. Ghione, and K. F. Brennan, "Band structure nonlocal pseudopotential calculation of the III-nitride wurtzite phase materials system. Part I. Binary compounds GaN, AlN, and InN," *J. Appl. Phys.*, vol. 88, pp. 6467–6475, Dec. 2000.

[11] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B*, vol. 14, pp. 556–582, July 1976.

[12] X. Zhou, F. Bertazzi, M. Goano, G. Ghione, and E. Bellotti, "Deriving $\mathbf{k}\cdot\mathbf{p}$ parameters from full-brillouin-zone descriptions: A finite-element envelope function model for quantum-confined wurtzite nanostructures," *J. Appl. Phys.*, vol. 116, no. 3, p. 033709, 2014.

[13] J. Piprek, ed., *Nitride Semiconductor Devices: Principles and Simulation.* Weinheim: Wiley-VCH Verlag, 2007.

[14] G. Verzellesi, D. Saguatti, M. Meneghini, F. Bertazzi, M. Goano, G. Meneghesso, and E. Zanoni, "Efficiency droop in InGaN/GaN blue light-emitting diodes: physical mechanisms and remedies," *J. Appl. Phys.*, submitted for publication.

[15] M. G. Burt, "The justification for applying the effective-mass approximation to microstructures," *J. Phys. Condens. Matter*, vol. 4, no. 32, pp. 6651–6690, 1992.

[16] B. A. Foreman, "Effective mass Hamiltonian and boundary conditions for the valence bands of semiconductor microstructures," *Phys. Rev. B*, vol. 48, no. 7, pp. 4964–4967, 1993.

[17] F. Mireles and S. E. Ulloa, "Ordered Hamiltonian and matching conditions for heterojunctions with wurtzite symmetry: $GaN/Al_xGa_{1-x}N$ quantum wells," *Phys. Rev. B*, vol. 60, no. 19, pp. 13659–13667, 1999.

[18] F. Mireles and S. E. Ulloa, "Strain and crystallographic orientation effects on the valence subbands of wurtzite quantum wells," *Phys. Rev. B*, vol. 62, no. 4, pp. 2562–2572, 2000.

[19] R. G. Veprek, S. Steiger, and B. Witzigmann, "Ellipticity and the spurious solution problem of $\mathbf{k}\cdot\mathbf{p}$ envelope equations," *Phys. Rev. B*, vol. 76, no. 16, pp. 165320–1–165320–9, 2007.

[20] I. Vurgaftman and J. R. Meyer, "Band parameters for nitrogen-containing semiconductors," *J. Appl. Phys.*, vol. 94, pp. 3675–3696, Sept. 2003.

[21] I. Vurgaftman and J. R. Meyer, "Electron bandstructure parameters," in Piprek [13], ch. 2, pp. 13–48.

[22] P. Rinke, M. Winkelnkemper, A. Qteish, D. Bimberg, J. Neugebauer, and M. Scheffler, "Consistent set of band parameters for the group-III nitrides AlN, GaN, and InN," *Phys. Rev. B*, vol. 77, no. 7, p. 075202, 2008.

[23] M. M. Rieger and P. Vogl, "Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates," *Phys. Rev. B*, vol. 48, pp. 14276–14287, Nov. 1993.

[24] D. J. Dugdale, S. Brand, and R. A. Abram, "Direct calculation of $k \cdot p$ parameters for wurtzite AlN, GaN, and InN," *Phys. Rev. B*, vol. 61, pp. 12933–12938, May 2000.

[25] J. Hader, J. V. Moloney, B. Pasenow, S. W. Koch, M. Sabathil, N. Linder, and S. Lutgen, "On the importance of radiative and Auger losses in GaN-based quantum wells," *Appl. Phys. Lett.*, vol. 92, no. 26, p. 261103, 2008.

[26] F. Bertazzi, M. Goano, and E. Bellotti, "A numerical study of Auger recombination in bulk InGaN," *Appl. Phys. Lett.*, vol. 97, p. 231118, Dec. 2010.

[27] F. Bertazzi, M. Goano, and E. Bellotti, "Numerical analysis of indirect Auger transitions in InGaN," *Appl. Phys. Lett.*, vol. 101, p. 011111, July 2012.

[28] F. Bertazzi, X. Zhou, M. Goano, G. Ghione, and E. Bellotti, "Auger recombination in InGaN/GaN quantum wells. A full-Brillouin-zone study," *Appl. Phys. Lett.*, vol. 103, p. 081106, Aug. 2013.

[29] F. Bertazzi, M. Moresco, and E. Bellotti, "Theory of high field carrier transport and impact ionization in wurtzite GaN. Part I: A full band Monte Carlo model," *J. Appl. Phys.*, vol. 106, p. 063718, Sept. 2009.

[30] J. Hader, V. Moloney, A. Thränhardt, and S. Koch, "Interband transitions in InGaN quantum wells," in Piprek [13], ch. 7, pp. 145–167.

[31] A. D. Andreev and E. P. O'Reilly, "Theory of the electronic structure of GaN/AlN hexagonal quantum dots," *Phys. Rev. B*, vol. 62, pp. 15851–15870, Dec. 2000.

[32] B. A. Foreman, "Elimination of spurious solutions from eight band $\mathbf{k} \cdot \mathbf{p}$ theory," *Phys. Rev. B*, vol. 56, no. 20, pp. R12 746–R12 751, 1997.

[33] A. V. Rodina, A. Y. Alekseev, A. L. Efros, M. Rosen, and B. K. Meyer, "General boundary conditions for the envelope function in the multiband $\mathbf{k \cdot p}$ model," *Phys. Rev. B*, vol. 65, p. 125302, Feb 2002.

[34] M. Holm, M.-E. Pistol, and C. Pryor, "Calculations of the electronic structure of strained inas quantum dots in InP," *J. Appl. Phys.*, vol. 92, no. 2, p. 932, 2002.

[35] K. I. Kolokolov, J. Li, and C. Z. Ning, "$\mathbf{k} \cdot \mathbf{p}$ hamiltonian without spurious-state solutions," *Phys. Rev. B*, vol. 68, p. 161308, Oct 2003.

[36] F. Szmulowicz, "Solution to spurious bands and spurious real solutions in the envelope-function approximation," *Phys. Rev. B*, vol. 71, p. 245117, Jun 2005.

[37] R. G. Veprek, S. Steiger, and B. Witzigmann, "Reliable $\mathbf{k \cdot p}$ band structure calculation for nanostructure using finite elements," *J. Comp. Electron.*, vol. 7, pp. 521–529, 2008.

[38] R. G. Veprek, S. Steiger, and B. Witzigmann, "Operator ordering, ellipticity and

spurious solutions in $\mathbf{k}\cdot\mathbf{p}$ calculations of III-nitride nanostructures," *Opt. Quantum Electron.*, vol. 40, pp. 1169–1174, 2008.

[39] T. Eissfeller and P. Vogl, "Real-space multiband envelope-function approach without spurious solutions," *Phys. Rev. B*, vol. 84, p. 195122, Nov 2011.

[40] S. V. Patankar, *Numerical Heat Transfer and Fluid Flow.* New York: McGraw-Hill, 1980.

[41] A. N. Brooks and T. J. Hughes, "Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 32, no. 13, pp. 199 – 259, 1982.

[42] V. John, P. Knobloch, and S. B. Savescu, "A posteriori optimization of parameters in stabilized methods for convectiondiffusion problems part i," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 4144, pp. 2916 – 2929, 2011.

[43] M. S. Hybertsen and S. G. Louie, "*Ab initio* static dielectric matrices from the density functional approach. I. Formulation and application to semiconductors and insulators," *Phys. Rev. B*, vol. 35, pp. 5585–5601, Apr. 1987.

[44] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, "Linear optical properties in the projector-augmented wave methodology," *Phys. Rev. B*, vol. 73, p. 045112, Jan 2006.

[45] C. Motta, M. Giantomassi, M. Cazzaniga, K. Gal-Nagy, and X. Gonze, "Implementation of techniques for computing optical properties in 03 dimensions, including a real-space cutoff, in {ABINIT}," *Computational Materials Science*, vol. 50, no. 2, pp. 698 – 703, 2010.

[46] H. Bross and H. Weilacher, "Properties describing the dynamics of crystal electrons for the case of a non-local potential," *Z. Phys.*, vol. 262, no. 2, pp. 95–104, 1973.

[47] C. J. Pickard and M. C. Payne, "Second-order $\mathbf{k}\cdot\mathbf{p}$ perturbation theory with vanderbilt pseudopotentials and plane waves," *Phys. Rev. B*, vol. 62, pp. 4383–4388, Aug 2000.

[48] S. L. Chuang and C. S. Chang, "$\mathbf{k}\cdot\mathbf{p}$ method for strained wurtzite semiconductors," *Phys. Rev. B*, vol. 54, no. 4, pp. 2491–2504, 1996.

[49] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems.* Springer Tracts in Modern Physics, Berlin: Springer-Verlag, 2003.

[50] G. L. Bir and G. E. Pikus, *Symmetry and Strain Induced Effects in Semiconductors.* New York: John Wiley & Sons, 1972.

[51] Y. C. Yeo, T. C. Chong, and M. F. Li, "Electronic band structures and effective-mass parameters of wurtzite GaN and InN," *J. Appl. Phys.*, vol. 83, pp. 1429–1436, Feb. 1998.

[52] G. B. Ren, Y. M. Liu, and P. Blood, "Valence band structure of wurtzite GaN including the spin-orbit interaction," *Appl. Phys. Lett.*, vol. 74, no. 8, pp. 1117–1119, 1999.

[53] K. Kim, W. R. L. Lambrecht, B. Segall, and M. van Schilfgaarde, "Effective masses and valence-band splittings in GaN and AlN," *Phys. Rev. B*, vol. 56, pp. 7363–7375, Sept. 1997.

[54] K. T. Delaney, P. Rinke, and C. G. Van de Walle, "Auger recombination rates in nitrides from first principles," *Appl. Phys. Lett.*, vol. 94, p. 191109, 2009.

[55] A. Svane, N. E. Christensen, I. Gorczyca, M. van Schilfgaarde, A. N. Chantis, and T. Kotani, "Quasiparticle self-consistent GW theory of III-V nitride semiconductors: Bands, gap bowing, and effective masses," *Phys. Rev. B*, vol. 82, p. 115102, Sept. 2010.

[56] L. C. de Carvalho, A. Schleife, and F. Bechstedt, "Influence of exchange and correlation on structural and electronic properties of AlN, GaN, and InN polytypes," *Phys. Rev. B*, vol. 84, p. 195105, Nov. 2011.

[57] M. Moresco, F. Bertazzi, and E. Bellotti, "Theory of high field carrier transport and impact ionization in wurtzite GaN. Part II: Application to avalanche photodetectors," *J. Appl. Phys.*, vol. 106, p. 063719, Sept. 2009.

[58] M. Goano, E. Bellotti, E. Ghillino, C. Garetto, G. Ghione, and K. F. Brennan, "Band structure nonlocal pseudopotential calculation of the III-nitride wurtzite phase materials system. Part II. Ternary alloys $Al_xGa_{1-x}N$, $In_xGa_{1-x}N$, and $Al_xIn_{1-x}N$," *J. Appl. Phys.*, vol. 88, pp. 6476–6482, Dec. 2000.

[59] E. Bellotti, F. Bertazzi, and M. Goano, "Alloy scattering in AlGaN and InGaN: A numerical study," *J. Appl. Phys.*, vol. 101, no. 12, p. 123706, 2007.

[60] A. Rubio, J. L. Corkill, M. L. Cohen, E. L. Shirley, and S. G. Louie, "Quasiparticle band structure of AlN and GaN," *Phys. Rev. B*, vol. 48, pp. 11810–11816, Oct. 1993.

[61] K. Miwa and A. Fukumoto, "First-principle calculation of the structural, electronic, and vibrational properties of gallium nitride and aluminum nitride," *Phys. Rev. B*, vol. 48, pp. 7897–7902, Sept. 1993.

[62] M.-Z. Huang and W. Y. Ching, "A minimal basis semi-*ab initio* approach to the

band structures of semiconductors," *J. Phys. Chem. Solids*, vol. 46, no. 8, pp. 977–995, 1985.

[63] S.-H. Wei and A. Zunger, "Valence band splittings and band offsets of AlN, GaN and InN," *Appl. Phys. Lett.*, vol. 69, pp. 2719–2721, Oct. 1996.

[64] M. Koshiba, *Optical Waveguide Theory by the Finite Elements Method.* Tokyo: KTK Scientific Publishers, 1992.

[65] G. Pelosi, R. Coccioli, and S. Selleri, *Quick Finite Elements for Electromagnetic Waves.* Norwood, MA: Artech House, 1998.

[66] L. W. Wang, A. Franceschetti, and A. Zunger, "Million-atom pseudopotential calculation of $\Gamma - X$ mixing in GaAs/AlAs superlattices and quantum dots," *Phys. Rev. Lett.*, vol. 78, no. 14, pp. 2819–2822, 1997.

[67] L.-W. Wang and A. Zunger, "Linear combination of bulk bands method for large-scale electronic structure calculations on strained nanostructures," *Phys. Rev. B*, vol. 59, pp. 15806–15818, June 1999.

[68] D. Esseni and P. Palestri, "Linear combination of bulk bands method for investigating the low-dimentional electrons gas in nanostructured devices," *Phys. Rev. B*, vol. 72, no. 16, pp. 165342–1–165342–14, 2005.

[69] D. Esseni and P. Palestri, "Fullband quantization analysis reveals a third valley in (001) silicon inversion layers," *IEEE Electron Device Lett.*, vol. 26, pp. 413–415, June 2005.

[70] S. L. Chuang, "Optical gain of strained wurtzite GaN quantum-well lasers," *IEEE J. Quantum Electron.*, vol. 32, no. 10, pp. 1791–1800, 1996.

[71] H. Zhao, R. A. Arif, Y.-K. Ee, and N. Tansu, "Self-consistent analysis of strain-compensated inganalgan quantum wells for lasers and light-emitting diodes," *IEEE J. Quantum Electron.*, vol. 45, no. 1, pp. 66–78, 2009.

[72] W. G. Scheibenzuber, U. T. Schwarz, R. G. Veprek, B. Witzigmann, and A. Hangleiter, "Calculation of optical eigenmodes and gain in semipolar and nonpolar InGaN/GaN laser diodes," *Phys. Rev. B*, vol. 80, pp. 115320–1–115320–16, 2009.

[73] R. G. Veprek, *Computational modeling of semiconductor nanostructures for opto-electronics.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2009.

[74] S. L. Chuang, *Physics of Photonic Devices.* Wiley, 2009.

[75] M. Kira and S. Koch, "Many-body correlations and excitonic effects in semiconductor spectroscopy," *Progress Quantum Electron.*, pp. 155–296, 2006.

[76] D. A. B. Miller, *Quantum Mechanics for Scientists and Engineers.* New York:

Cambridge University Press, 2009.

[77] W. W. Chow and S. W. Koch, *Semiconductor-Laser Fundamentals. Physics of the Gain Materials.* Berlin: Springer-Verlag, 1999.

[78] E. Malic and A. Knorr, *Graphene and Carbon Nanotubes. Ultrafast Relaxation Dynamics and Optics.* Weinheim: Wiley-VCH Verlag, 2013.

[79] H. Haug and S. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors.* World Scientific, 2004.

[80] J. Schilp, T. Kuhn, and G. Mahler, "Electron-phonon quantum kinetics in pulse-excited semiconductors: Memory and renormalization effects," *Phys. Rev. B*, vol. 50, no. 8, pp. 5435–5447, 1994.

[81] E. Schöll, *Theory of Transport Properties of Semiconductor Nanostructures.* Dordrecht: Springer-Verlag, 1998.

[82] K.-T. Tsen, *Ultrafast Phenomena in Semiconductors.* New York: Springer-Verlag, 2001.

[83] I. Waldmuller, J. Forstner, S. Lee, A. Knorr, M. Woerner, K. Reimann, R. A. Kaindl, T. Elsaesser, R. Hey, and K. H. Ploog, "Optical dephasing of coherent intersubband transitions in a quasi-two-dimensional electron gas," *Phys. Rev. B*, vol. 69, no. 20, pp. 205307–1–205307–9, 2004.

[84] S. Selberherr, *Analysis and Simulation of Semiconductor Devices.* Wien: Springer-Verlag, 1984.

[85] E. M. Azoff, "Energy transport numerical simulation of graded algaas/gaas heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 36, no. 4, pp. 609–616, 1989.

[86] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation.* Computational Microelectronics, Wien: Springer-Verlag, 1989.

[87] C. Jacoboni, *Theory of Electron Transport in Semiconductors. A Pathway from Elementary Physics to Nonequilibrium Green Functions.* Berlin: Springer-Verlag, 2010.

[88] M. Grupen and K. Hess, "Simulation of carrier transport and nonlinearities in quantum-well laser diodes," *IEEE J. Quantum Electron.*, vol. 34, pp. 120–140, Jan. 1998.

[89] F. Rossi, *Theory of Semiconductor Quantum Devices. Microscopic Modeling and Simulation Strategies.* Berlin: Springer-Verlag, 2011.

[90] M. Luisier, "Atomistic simulation of transport phenomena in nanoelectronic devices," *Chemical Society Reviews*, vol. 43, pp. 115421–1–115421–7, 2014.

[91] M. Moussavou, N. Cavassilas, E. Dib, and M. Bescond, "Influence of uniaxial strain in si and ge p-type double-gate metal-oxide-semiconductor field effect transistors," *J. Appl. Phys.*, vol. 118, pp. 114503–1–114503–5, 2015.

[92] U. Aeberhard, "Quantum-kinetic theory of steady-state photocurrent generation in thin films: Coherent versus incoherent coupling," *Phys. Rev. B*, vol. 89, pp. 115303–1–115303–10, 2014.

[93] A. Shedbalkar, Z. Andreev, and B. Witzigmann, "Simulation of an indium gallium nitride quantum well light-emitting diode with the non-equilibrium green's function method," *Phys. Status Solidi B*, vol. 253, no. 1, pp. 158–163, 2016.

[94] S. Steiger, *Modelling Nano-LEDs.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2009.

[95] S. Steiger, M. Povolotskyi, H. Park, T. Kubis, and G.Klimeck, "Nemo5: A parallel multiscale nanoelectronics modeling tool," *IEEE Trans. Nanotech.*, vol. 10, no. 6, pp. 1464–1474, 2011.

[96] D. L. Smith and C. Mailhiot, "Theory of semiconductor superlattice electronic structure," *Rev. Modern Phys.*, vol. 62, pp. 173–234, Jan. 1990.

[97] S. A. Rakityansky, "Unified treatment of bound, scattering, and resonant states in one-dimensional semiconductor nanostructures," *Phys. Rev. B*, vol. 68, no. 19, pp. 195320–1–195320–17, 2003.

[98] W. R. Frensley, "Numerical evaluation of resonant states," *Superlatt. Microstruct.*, vol. 11, no. 3, pp. 347–350, 1992.

[99] Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, "Efficient, numerically stable multiband $\mathbf{k}\cdot\mathbf{p}$ treatment of quantum transport in semiconductor heterostructures," *Phys. Rev. B*, vol. 54, no. 8, pp. 5675–5683, 1996.

[100] L. Lin, "Fast algorithm for extracting the diagonal of the inverse matrix with application to the electronic structure analysis of metallic systems," *Comp. Mater. Sci.*, vol. 7, no. 3, pp. 755–777, 2009.

[101] S. Datta, *Electronic Transport in Mesoscopic systems.* Cambridge University Press, 1995.

[102] S. Datta, "Nanoscale device modeling: the greens function method," *Superlatt. Microstruct.*, vol. 28, pp. 253–278, 2000.

[103] O. Manasreh, *Semiconductor Heterojunctions and Nanostructures.* New York: McGraw-Hill, 2005.

[104] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics.* Westview Press, 1962.

[105] C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, "Direct calculation of the tunneling current," *J. Phys. C*, vol. 4, pp. 9166–929, 1971.

[106] P. Danielewicz, "Direct calculation of the tunneling current," *Ann. Phys.*, vol. 152, pp. 239–304, 1984.

[107] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, "Nonequilibrium greens function treatment of phonon scattering in carbon-nanotube transistors," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2339–2351, 2007.

[108] R. Golizadeh-Mojarad and S. Datta, "Nonequilibrium greens function based models for dephasing in quantum transport," *Phys. Rev. B*, vol. 75, no. 8, pp. 081301–1–081301–4, 2007.

[109] R. K. Lake and R. R. Pandey, "Non-equilibrium green functions in electronic device modeling, cond-mat/0607219v1," *arXiv*, pp. 1–38, 2007.

[110] F. Dolcini, R. C. Iotti, and F. Rossi, "Interplay between energy dissipation and reservoir-induced thermalization in nonequilibrium quantum nanodevices," *Phys. Rev. B*, vol. 88, no. 11, pp. 5675–5683, 2013.

[111] G. Lin, D. Meyaard, J. Cho, E. F. Schubert, H. Shim, and C. Sone, "Analytic model for the efficiency droop in semiconductors with asymmetric carrier-transport properties based on drift-induced reduction of injection efficiency," *Appl. Phys. Lett.*, vol. 100, pp. 161106–1–161106–4, 2012.

[112] M. Luisier, "Many-body effects in semiconductor lasers.," *Masters thesis, Departement of Information Technology and Electrical Engineering, ETH Zurich*, 2003.

[113] L. D. Marks and D. R. Luke, "Robust mixing for ab initio quantum mechanical calculations," *Phys. Rev. B*, vol. 78, no. 7, pp. 075114–1–075114–12, 2009.

[114] R. Zitko, "Convergence acceleration and stabilization of dynamical mean-field theory calculations," *Phys. Rev. B*, vol. 80, no. 12, pp. 125125–1–125125–4, 2009.

[115] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode transistor," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 64–77, Jan. 1969.