



**Politecnico
di Torino**

POLITECNICO DI TORINO

Dipartimento di Elettronica e Telecomunicazioni

Corso di Laurea Magistrale in Ingegneria Elettronica

Classe LM-29

Tesi di Laurea Magistrale

**Advanced Monte Carlo simulations of
carrier transport in materials and devices
for optoelectronics**

Relatori

Prof. Michele GOANO

Prof. Enrico BELLOTTI

Prof. Francesco BERTAZZI

Dr. Alberto TIBALDI

Dr. Matteo G. C. ALASIO

Candidato

Domenico ABBONDANDOLO

Anno Accademico 2023/2024

Contents

1	Introduction to Monte Carlo method	2
1.1	Transport model	3
1.2	Algorithm	15
2	Advanced models for Monte Carlo simulations	23
2.1	Analytic band electronic structure	24
2.2	Full band electronic structure	40
3	Mobility of InAs, InSb and InAsSb	51
3.1	Simulation context	51
3.2	Simulation framework	51
3.3	InAs and InSb	53
3.4	InAsSb	62
4	Ge-on-Si photodetector frequency response	68
4.1	Simulation context	68
4.2	Simulation framework	70
4.3	Results	74
5	Conclusions	79
A	Numerical methods	80
A.1	Bisection method	80
A.2	Simpson integration method	81
A.3	Gauss elimination method	82

Chapter 1

Introduction to Monte Carlo method

The Monte Carlo method is a technique for the simulation of carrier transport in semiconductors. It is a semi-classical approach, meaning that one part of the system is treated classically and the other part is described quantum mechanically. Differently from the most common simulation technique, i.e. drift-diffusion, in which transport is described by a differential equation that is solved numerically (using Finite-difference or Finite-Element methods), in the Monte Carlo simulations the transport equations (for electrons and holes) are solved through a statistical analysis on the possible interactions carriers can be subjected to. In order to do that, one or more particles are simulated for a certain period of time; in this time, carriers undergo a certain number of interactions, coming from an external electric field or from the perturbing potential of the underlying crystal atoms (where perturbation means any displacement from the condition of ideal crystal with atoms frozen at their equilibrium position) from which the energy or the velocity distribution can be obtained.

This level of insight of the particles' dynamics allows to predict and simulate quite accurately some transport properties, such as velocity-field curves, low-field mobility in function of temperature, doping, etc. These quantities are used by other simulation techniques, such as drift-diffusion, as input parameters. Moreover, the Monte Carlo method allows the possibility of studying high energy transport, such as impact ionization, that is quite impossible to study with drift-diffusion based on low-energy approximations. Finally, Monte Carlo transport can be used for the simulation of electronic devices as well, making it an incomparable method in the study of detectors (such as Avalanche photodetectors) and transistors (e.g. MOSFETs or HFETs) in which significantly high electric field is reached.

One of the most important constituent of the Monte Carlo technique, on which the accurateness of the simulations relies, is the physical model of the crystal-carriers system, namely the description of the **electronic bandstruc-**

ture and the description of interactions between perturbed crystal potential and carriers (**scatterings**). For example, the electronic bandstructure can be approximated by the minimum of the first conduction band or the minima of its valleys (the so-called *analytic band* approximation), or a more accurate multi-band dispersion can be accounted for. On the other hand, high degree of approximation is used for the treatment of scatterings. The main research trends of Monte Carlo are focused upon finding solutions for a better description of this physical model in a computationally-efficient way, and this work is aimed at the investigation of some of those.

In the following section, the basics of Monte Carlo simulation framework is explained, in order to provide a basis to understand the choices and the implementations made throughout the thesis. Most of information is taken from [24].

1.1 Transport model

Under certain approximations, that will be discussed later, electronic transport in semiconductors can be modeled through the Boltzmann transport equation (BTE):

$$\frac{df(\mathbf{r}, \mathbf{k}, t)}{dt} = -\nabla_{\mathbf{r}}f(\mathbf{r}, \mathbf{k}, t) \cdot \mathbf{v} - \frac{q\mathbf{F}(\mathbf{r})}{\hbar} \cdot \nabla_{\mathbf{k}}f(\mathbf{r}, \mathbf{k}, t) + \left(\frac{\partial f(\mathbf{r}, \mathbf{k}, t)}{\partial t}\right)_{\text{coll}}, \quad (1.1)$$

where:

- the unknown f is the probability density of a carrier (electron or hole) in function of time, position and k-vector. It is the probability of finding a particle in the infinitesimal interval $d^3\mathbf{r}d^3\mathbf{k}$ centered at position \mathbf{r} and at point \mathbf{k} at time t ;
- the first term in the right-hand side is the contribution to the time evolution of f due to a diffusion mechanism: at a fixed point in real space and k-space, f decreases at a rate proportional to the modulus of the velocity \mathbf{v} in that point and to the modulus of the space-gradient of f (remembering that the direction of the gradient is the direction in which the function increases most quickly from a given point and the magnitude is the rate of increase in that direction). In conclusion, this contribution tends to move f in real space according to the velocity of the particle and to spread the distribution over the space;
- the second term in the right-hand side is a term related to drift under the effect of a field \mathbf{F} ; it has the same shape of the diffusion term but it involves the coordinates in the k-space only, so it can be seen as a diffusion in k-space;
- the third term of the right-hand side is the temporal evolution of the distribution due to collision events. This term describes the carrier interaction with the lattice and with other carriers.

The Boltzmann transport equation is a semiclassical approach to the problem of electronic transport, meaning that apart from the collision term, that is evaluated using quantum-mechanics, as will be shortly illustrated later, the structure of the equation comes from classical physics. In fact, the BTE is used to describe thermodynamic systems like gas particles or fluids; In order to well describe electronic transport as well, the following conditions have to occur:

1. since the particles have a well-defined position and momentum, the characteristic length of materials or devices simulated must be larger than the size of the electron wave-packet given by the mean-free-path (the average distance traveled by the electron between two collisions); moreover, the energy scale of interest should be much larger than the uncertainty implied by the spread of the electron momentum;
2. all scattering processes are independent (no memory is conserved of where and when the previous collisions happened), local in space (they involve no change of \mathbf{r}), instantaneous (their duration is negligible with respect to the free-flight time between successive collisions), so they depend only on the initial and final \mathbf{k} ;
3. k-states are electronic states of an infinite ideal crystal, with ion cores frozen in their equilibrium position without any external potential applied, as given by the Bloch theorem [18].

1.1.1 Scattering rates

The collision term in equation 1.1 can be modeled as

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \sum_{\mathbf{k}'} f(\mathbf{r}, \mathbf{k}', t) S(\mathbf{k}', \mathbf{k}) [1 - f(\mathbf{r}, \mathbf{k}, t)] - f(\mathbf{r}, \mathbf{k}, t) S(\mathbf{k}, \mathbf{k}') [1 - f(\mathbf{r}, \mathbf{k}', t)], \quad (1.2)$$

where $S(\mathbf{k}', \mathbf{k})$ is the scattering (collision) probability per unit time. The first term of the sum describes the *in-scattering*, the rate at which particles enter, from any state \mathbf{k}' , the infinitesimal volume $d^3\mathbf{r}d^3\mathbf{k}$ centered at point (\mathbf{r}, \mathbf{k}) . This rate is proportional to the probability that the state \mathbf{k}' is occupied, the probability that the state \mathbf{k} is not occupied and the probability this event can happen $S(\mathbf{k}', \mathbf{k})$. In non-degeneracy conditions, Fermi-Dirac statistics is approximated by Boltzmann statistics and all states are considered always available, i.e. $[1 - f(\mathbf{r}, \mathbf{k}, t)] = 1$, but in case of degeneracy, this term could be relevant and change the simulated material properties. The second term of the sum describes the *out-scattering*, i.e. the rate at which particles exit the volume $d^3\mathbf{r}d^3\mathbf{k}$ and the state \mathbf{k} to go into any state \mathbf{k}' .

Finally, the term $S(\mathbf{k}', \mathbf{k})$ is evaluated considering the different effects of the carrier interaction with lattice vibrations, impurities etc. as perturbations of the simple time-independent Hamiltonian \hat{H}_0 for which the electronic states are evaluated (the Bloch states remembering item 3. of 3)

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{H}'(t), \quad (1.3)$$

where $\lambda\hat{H}'(t)$ is the Hamiltonian introduced by the perturbation.

Any possible state $|\phi_n(t)\rangle$ of the new Hamiltonian $\hat{H}(t)$ can be expressed as a linear combination of the time-dependent eigenstates of the unperturbed problem \hat{H}_0 :

$$\hat{H}_0 |\psi_n\rangle = E_n |\psi_n\rangle, \quad (1.4)$$

$$|\psi_n\rangle(t) = |\psi_n\rangle e^{-E_n(t-t_0)/\hbar}, \quad (1.5)$$

$$|\phi_n(t)\rangle = \sum_n a_n(t) |\psi_n(t)\rangle. \quad (1.6)$$

Solving the Schroedinger equation for $\hat{H}(t)$, making a power expansion of the solution over λ and retaining only the first order term (i.e. proportional to λ^1), it is possible to write a differential equation for the time evolution of the coefficients a_n . As a consequence, if, for example, at time $t = t_0$ the total state $|\phi_n(t)\rangle$ is equal to an unperturbed eigenstate $|\psi_i(t)\rangle$ (so $a_i(t_0) = 1$), as time passes the probability coefficient $a_i(t)$ can decrease and the coefficient of another state a_f can increase. That is, a transition between state i and f has occurred. Considering that the time for the interaction is sufficiently long, i.e. that two different interaction (and so two different perturbations) are not too close in time, the first-order transition probability per unit time is

$$S(i, f) = \frac{2\pi}{\hbar} |\langle f | \hat{H}' | i \rangle|^2 \delta(E_f - E_i \pm \hbar\omega), \quad (1.7)$$

where $|i\rangle$ is the initial state, $|f\rangle$ is the final state, E_f is the final energy after the state transition, E_i is the initial energy and $\hbar\omega$ is the energy transferred in the process. The delta term in 1.7 indicates that the S is different from zero only when $E_f - E_i \pm \hbar\omega = 0$ i.e. when energy is conserved. Equation 1.7 is known as **Fermi Golden Rule**. The most important perturbations considered in a Monte Carlo simulation are the followings.

Phonon deformation potential scattering It is the perturbation of the electrostatic potential felt by the carrier due to the displacement of ionic cores of the lattice. The displacement is caused by lattice vibrations that becomes more relevant as the lattice temperature increases. These vibrations are modeled as quasi-particles called phonons. In a crystal, several vibration modes exists, that can be visualized in the phonon dispersion relation (figure 1.1b), i.e. the relation between phonon energy and phonon wavevector $q = |\mathbf{k}' - \mathbf{k}|$. The most important classification of these modes is in **acoustic**, that have a linear dispersion around $q = 0$ (Γ point of the Brillouin zone, figure 1.1a) and **optical** phonons, that have non-zero energy and approximately constant dispersion at $q = 0$. The scattering rates for the acoustic and optical carrier-phonon interaction are described by the following formulas [5]:

$$\frac{1}{\tau_{ac}^{\pm}(n, \mathbf{k})} = A \sum_{\mathbf{n}'} \int_V d\mathbf{k}' \frac{q^2}{\omega_{ac}} \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) \times |I_{ovp}|^2 \times \delta(E_f - E_i \mp \hbar\omega_{ac}), \quad (1.8)$$

$$\frac{1}{\tau_{\text{op}}^{\pm}(n, \mathbf{k})} = B \sum_{n'} \int_V d\mathbf{k}' \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) \times |I_{\text{ovp}}|^2 \times \delta(E_f - E_i \mp \hbar\omega_{\text{op}}), \quad (1.9)$$

Here, the form of equation 1.7 and of equation 1.2 can be recognized. In particular:

- the minus sign corresponds to phonon absorption, while the plus sign corresponds to a phonon emission;
- ω is the phonon frequency. In order to be suitable for a Monte Carlo simulation, the phonon dispersion relation is approximated, and ω is constant for optical phonons (green line in figure 1.1b). For acoustic phonons, it can be approximated around $q = 0$ with a linear expression $\omega_{\text{ac}}(q) = qv_l$, where v_l is called the mode sound velocity ([24]), or, alternatively, it can be approximated as a piecewise function (red line in figure 1.1b, [5]):

$$\omega_{\text{ac}}(q) = \begin{cases} (4v_l/a)\sqrt{1 - \cos(qa/4)} & \text{if } q \leq 1, \\ 4v_l/a & \text{otherwise,} \end{cases}$$

where a is the crystal lattice constant;

- N_q is the phonon number having wavevector modulus equal to q . Since phonons are bosons, they follow the Bose-Einstein statistics at thermal equilibrium

$$N_q = \frac{1}{\exp(\hbar\omega_q)/(k_B T) - 1}. \quad (1.10)$$

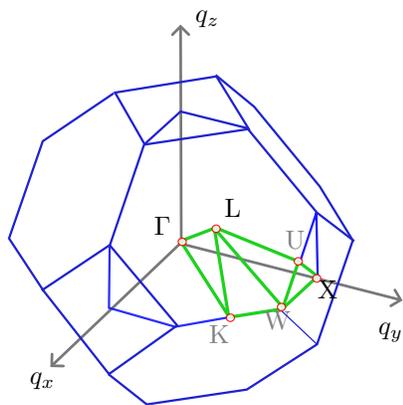
The same statistics is used out of equilibrium as well, since it is supposed that the movement of the electrons under the effect of the electric field does not affect too much the lattice temperature;

- $A = (2\pi^2 D_{\text{ac}}^2)/(\hbar V \rho)$, where ρ is the crystal density, V is the crystal volume and D_{ac} is called **acoustic deformation potential**;
- $B = (\pi(D_t K)^2)/(V \rho \omega_{\text{op}})$, where $(D_t K)$ is the **optical deformation potential**;
- I_{ovp} is the overlap factor between the initial and the final wavefunction

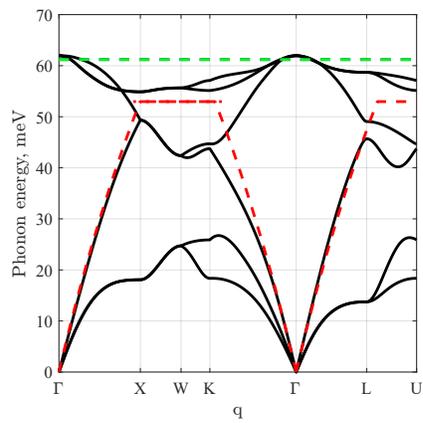
$$I_{\text{ovp}}(n, \mathbf{k}, n', \mathbf{k}') = \int d\mathbf{r} \psi_i^*(n, \mathbf{k}) \psi_f(n', \mathbf{k}'). \quad (1.11)$$

It is worthwhile giving particular attention the terms D_{ac} and $D_t K$. They can be obtained with numeric simulations (such as DFT, [1]) or can be obtained through fitting experimental data.

The term $|I_{\text{ovp}}(n, \mathbf{k}, n', \mathbf{k}')|^2$ is actually present in all types of scattering, because it depends on the initial and final state and not on the particular scattering mechanism. However, it is very difficult to take into account: it is usually approximated to 1 or with an analytical formula [24].



(a). Visualization of the full Brillouin zone for a zincblende crystal. The green lines represent the set of direction of the irreducible wedge



(b). Phonon dispersion relation in silicon obtained with *ab initio* calculations. The x -axis corresponds to a set of directions in the reciprocal space depicted in figure 1.1a. The red dashed line represent the approximation of the dispersion of acoustic phonons used in Monte Carlo simulations; the green dashed line represents the approximation of the dispersion of the optical phonon used in Monte Carlo simulations

Figure 1.1: Dispersion relation in reciprocal space

Polar phonon scattering In polar material, such as GaAs, InAs or InSb, the vibrations of opposite charged atoms in polar materials generate long-range dipole fields leading to a macroscopic electric field that can scatter the carriers. The polar interaction may be due to either acoustic or optical phonons, leading to piezoelectric and polar optical scattering, respectively.

Ionized impurity scattering It is due to the carrier interaction with the Coulomb potential of ionized impurities, e.g. doping.

Impact ionization It is the interaction among carriers that can bring to the generation of e/h pairs. It occurs when carriers have high energy, e.g. in presence of high electric field.

Alloy scattering It is due to the fluctuation of the local crystal potential in space, present in alloy materials.

In the following sections, other useful physical quantities are described related to scatterings, in particular to the selection of the state after scattering. Finally, a more detailed insight is dedicated to ionized impurity scattering, since it is part of the focus of this work.

1.1.2 Useful physical quantities

Momentum relaxation time

The majority of scattering mechanisms is not isotropic, i.e. the initial momentum is not randomized after the scattering and carrier retain a certain memory of the incident momentum. A quantity that gives information on the time required to relax (randomize) the initial \mathbf{k} is the momentum relaxation time [33]

$$\frac{1}{\tau_m(\mathbf{k})} = \sum_{\mathbf{k}'} S(\mathbf{k}, \mathbf{k}') \left(1 - \frac{k'}{k} \cos(\theta)\right), \quad (1.12)$$

where θ is the angle between the incident and scattered momenta. The physical sense of this formula is that each scattering rate is weighted by their ability to change the scattering angle: if for a scattering rate $\theta = 0$, i.e. momentum not relaxed, its contribution to the sum is minimum; if instead $\theta = \pi$, the contribution is maximum.

Cross section

The differential cross section σ , is defined as the ratio of the number of particles scattered into the direction (θ, ϕ) per unit time, per unit solid angle, divided by the incident flux of particles [7]. The normalized cross section gives information about the probability density of the scattering angle, in fact

$$\sigma_{\text{norm}} d\Omega, \quad (1.13)$$

where $\sigma_{\text{norm}} = (C \cdot \sigma)$ such that $\int_{\Omega} \sigma_{\text{norm}} d\Omega = 1$, is the probability of finding the scattered particle in the infinitesimal interval $d\Omega$ centered at solid-angle Ω . In order to evaluate the scattering rate W , one has to integrate the cross section over the solid angle. If the particle scatters with a certain number of centers with density N_i

$$W = \int_{\Omega} N_i v_g \sigma d\Omega, \quad (1.14)$$

where v_g is the group velocity of the particle. Since the effect of the scattering is to change the direction of the particle wavevector \mathbf{k} (as well as particle's energy), under an appropriate choice of the reference system for the collision, the cross section depends only on θ . Knowing also that

$$d\Omega = \sin(\theta) d\theta d\phi, \quad (1.15)$$

equation 1.14 becomes

$$W = 2\pi \int_0^{\pi} N_i v_g \sigma(\theta) d\theta, \quad (1.16)$$

and the probability density becomes

$$f(\theta) = C \cdot 2\pi \sin(\theta) \sigma(\theta), \quad (1.17)$$

where C is the coefficient for the normalization.

Impact parameter

Related to the cross section, another important quantity in the collision theory is the impact parameter, that can be defined as the distance between the straight line representing the trajectory of the particle before scattering, and the straight line parallel to the latter and passing through the scattering centre. A visualization of the impact parameter and the scattering angle is shown in figure 1.2.

In classical physics, cross section and impact parameter are related by the following equation

$$\sigma(\theta) d\Omega = 2\pi b |db|. \quad (1.18)$$

1.1.3 Ionized impurity scattering

The scattering is elastic (i.e. $|\mathbf{k}'| = |\mathbf{k}|$) and the perturbing Hamiltonian is

$$H'(\mathbf{r}) = \sum_{i=1}^{N_i} V(\mathbf{r} - \mathbf{r}_i), \quad (1.19)$$

where N_i is the total number of impurities and $V(\mathbf{r} - \mathbf{r}_i)$ is the electrostatic potential of the impurity located at position \mathbf{r}_i . In the perturbation theory

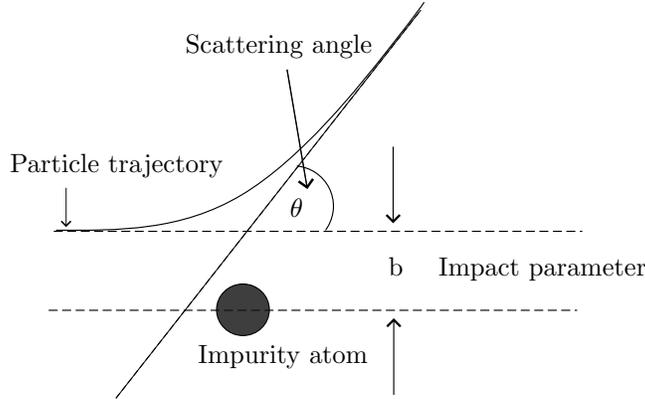


Figure 1.2: Visualization of the scattering angle and the impact parameter

framework, the contribution of the potential of all impurities should be considered; however, it is assumed that only one scattering center is active and all the others average out, so that the interaction is considered as two-body mechanism:

$$|\langle \mathbf{k}' | H' | \mathbf{k} \rangle|^2 = N_i |\langle \mathbf{k}' | H'_{\text{singleimp}} | \mathbf{k} \rangle|^2. \quad (1.20)$$

In principle, this perturbation modifies the carrier wavefunction locally, i.e. near the interaction, according to the time-independent Schrodinger problem:

$$\left[-\frac{\hbar^2}{2m^*} \nabla^2 + \hat{H}'(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (1.21)$$

where m^* is the carrier (electron or hole) effective mass, that takes into account the periodic potential of the unperturbed crystal ($H' = 0$). Applying the concepts of quantum collision theory it is possible to show that the presence of this perturbation modifies the transition probability between the unperturbed eigenstates (as it has been seen in the Fermi golden rule), and the transition rate is proportional to the transition matrix [7]

$$T(\mathbf{k}', \mathbf{k}) = \int e^{-i\mathbf{k}' \cdot \mathbf{r}} V(\mathbf{r}) \psi_k(\mathbf{r}) d\mathbf{r}, \quad (1.22)$$

where $\psi_k(\mathbf{r})$ is the solution of 1.21.

The use of the transition matrix as eq.1.22 allows to evaluate the transition rate exactly without approximations. To obtain the first-order perturbation approximation, as in the Fermi-Golden rule (equation 1.7), $\psi_k(\mathbf{r})$ is approximated by the unperturbed wave $e^{i\mathbf{k} \cdot \mathbf{r}}$. This is also called *first Born approximation*. It is used in Monte Carlo simulations because it provides simple formulas to implement at low computational cost.

The two simplest and most famous models for ionized impurity scattering are Brooks-Herring and Conwell-Weisskopf models.

Conwell-Weisskopf model This model [9] supposes that a particle scatters only with the nearest impurity centre, and all the perturbing effects lying further away than half the average distance between centres add up to zero. The potential $V(\mathbf{r})$ is Coulombian truncated at distance $b = (4/3\pi n_i)^{-1/3}$ where n_i is the impurity density:

$$V(\mathbf{r}) = \frac{eZ}{4\pi\epsilon r}, \quad (1.23)$$

where e is the elementary charge, Z is the number of charge units of the impurity, ϵ is the dielectric constant of the material. Using the first Born approximation, and approximating the overlap integral to 1, the differential cross section can be written as

$$\sigma_{\text{CW}}(\theta) = \begin{cases} \frac{R^2}{4\sin^4(\theta/2)} & \text{if } \theta \geq \theta_{\min}, \\ 0 & \text{otherwise,} \end{cases}$$

where the minimum scattering angle is evaluated according to

$$\tan\left(\frac{\theta_{\min}}{2}\right) = \frac{e^2 m^*}{4\pi\epsilon b \hbar^2 k^2} \quad (1.24)$$

and R is equal to

$$R = \frac{Ze^2}{4\pi\epsilon m^* v_g^2}. \quad (1.25)$$

The total scattering rate for spherical parabolic bands is [24]

$$W_{\text{CW}} = \pi N_i Z^2 b^2 \sqrt{\frac{2E}{m^*}}, \quad (1.26)$$

where E is the energy of the particle.

In figure 1.3 the scattering rates for doping level $N_i = 1 \times 10^{15} \text{ cm}^{-3}$ and $N_i = 1 \times 10^{17} \text{ cm}^{-3}$ are reported. The material properties used to obtain this result are the followings:

- effective mass $m^* = 0.05 \text{ eV}$;
- relative dielectric constant $\epsilon_r = 15.15$;
- temperature $T = 300 \text{ K}$;
- spherical parabolic bands.

i.e. the parameters related to the Γ valley of the first conduction band of InAs.

In figure 1.4 the normalized differential cross sections for the same doping levels are reported. The material properties used are the same; the particle energy for which the figure is obtained is $E = 0.01 \text{ eV}$;

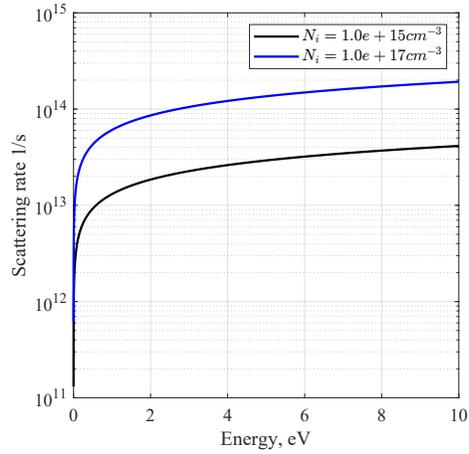


Figure 1.3: Scattering rates obtained using CW model. $T = 300$ K, $E = 0.01$ eV, $m^* = 0.05m_0$, $\epsilon_r = 15.15$

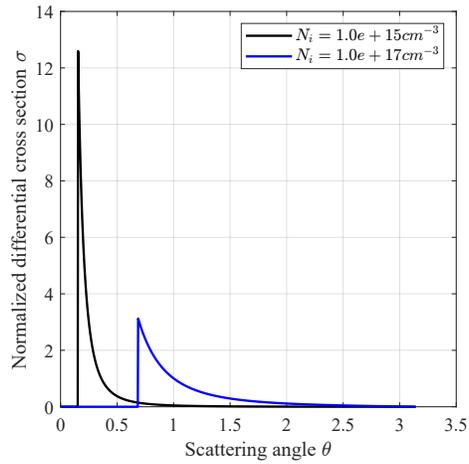


Figure 1.4: Differential cross section obtained using CW model. $T = 300$ K, $E = 0.01$ eV, $m^* = 0.05m_0$, $\epsilon_r = 15.15$

Brooks-Herring model This is the most used model [8], since it removes the truncation of the electric field of the CW model that is, in a certain sense, arbitrary. The BH model assumes that when a particle scatters with the Coulombian potential of a centre, nearby carriers screen such potential, leading to

$$V(\mathbf{r}) = \frac{eZ}{4\pi\epsilon r} e^{-q_0 r}, \quad (1.27)$$

with q_0 as the inverse Debye screening length

$$q_0 = \sqrt{\frac{e^2 n}{\epsilon k_B T}}. \quad (1.28)$$

In equation 1.28, n is the carrier density (electron or holes), T is the electron temperature (approximated with the lattice temperature, but in case of high carrier density it can be different [34]) and k_B is the Boltzmann constant. Using the first Born approximation, and approximating the overlap integral to 1, the differential cross-section can be written as

$$\sigma_{\text{BH}}(k, \theta) = \frac{R^2}{4(\sin^2(\theta/2) + \beta^{-1})^2}, \quad (1.29)$$

where $\beta^{-1} = ((\hbar q_0)/(2m^*v_g))^2$.

The total scattering rate for spherical parabolic bands is [24]

$$W_{\text{BH}} = \frac{2^{5/2} \pi N_i Z^2 e^4}{(4\pi\epsilon)^2 E_\beta^2 (m^*)^{1/2}} \frac{\sqrt{E}}{1 + 4E/E_\beta}, \quad (1.30)$$

where $E_\beta = \frac{\hbar^2 q_0^2}{2m^*}$ and E is the energy of the particles.

In figures 1.5 the scattering rates for doping levels $N_i = 1 \times 10^{15} \text{cm}^{-3}$, $N_i = 1 \times 10^{17} \text{cm}^{-3}$ and $N_i = 1 \times 10^{19} \text{cm}^{-3}$ are reported. In figure 1.6 the normalized differential cross sections are reported. The used material properties are the same employed for CW model figures.

Comparing the results for the two models it is possible to notice that:

- the BH model presents a very high scattering rate at low energy that increases as doping level decreases. However, the scattering angle is very likely to be small, so the velocity components of the particles at each scattering event is slightly modified;
- on the other hand, CW model presents low scattering rates at low energy that increases as doping increase. The scattering angle is higher on average compared to the BH model, so the velocity of the particle is more affected at each scattering.

Since the particles' energy in Monte Carlo simulations is usually less than 1 eV, the CW model is more computationally efficient than the BH one, since a scattering event is less likely to happen.

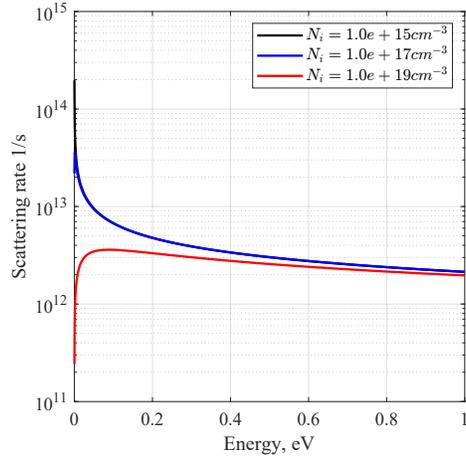


Figure 1.5: Scattering rates obtained using BH model. $T = 300$ K, $m^* = 0.05m_0$, $\epsilon_r = 15.15$

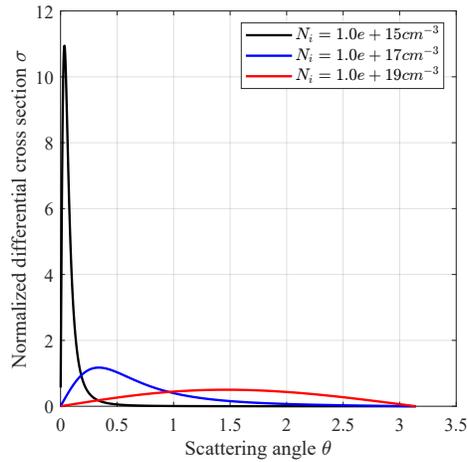


Figure 1.6: Differential cross section obtained using BH model. $T = 300$ K, $E = 0.01$ eV, $m^* = 0.05m_0$, $\epsilon_r = 15.15$

Advanced models for ionized impurity scattering

The main limitation of the models described above are the followings.

1. They are two different models to describe the same phenomenon. CW model works better than BH when high degree of compensation is present, i.e., when both donor and acceptor concentration is significant. In this case, few carriers are able to screen the ionized impurities, but the latter are able to screen each other. On the other hand, in case each of the ionized impurities contributes to free carriers and the screening is high, as in the case of majority electrons or holes, the BH model is better. A model that works for every condition of screening would be preferable instead of using the two models separately.
2. They both make use of the first Born approximation. Attempts have been made to include the exact transition rate in Monte Carlo simulation, solving 1.21 using spherical harmonic expansion ([27], [26], [35]); however, for the materials analyzed in that papers, this approach resulted more significant that the Born approximation only at low temperature ($T = 0 \sim 80K$) and relatively low doping concentration (low degeneracy).
3. Each scattering is the result of an interaction with a single impurity center which acts on the electron independently of all others. To visualize the validity of this approximation, it is possible to use the quantity $\xi = (4\pi N_i \lambda_d^3)/3$, where

$$\lambda_d = \frac{1}{q_0} \cdot \sqrt{\frac{\mathcal{F}_{1/2}(\eta)}{\mathcal{F}_{-1/2}(\eta)}}, \quad \eta = \frac{E_F - E_C}{k_B T}, \quad (1.31)$$

q_0 is 1.28 and \mathcal{F}_x are the Fermi integrals of order x . It represents the number of carriers present in the screening sphere: if $\xi \gg 1$, the approximation of the perturbation as a single screened Coulomb potential with the Debye length as screening length is valid. If $\xi < 1$, a *multi-ion* scattering regime is present and the electron reacts to many impurity atoms simultaneously. Figure 1.7 reports the ξ quantity for silicon, data are reproduced from [27].

1.2 Algorithm

The Monte Carlo method is a way to solve BTE 1.1 numerically. The algorithm is the following:

1. a certain number of particles is initialized with a certain distribution (e.g. Boltzmann distribution);
2. for a time τ between two collisions, that can be constant or an outcome of a random variable, particles are accelerated (**drift**) under the effect of

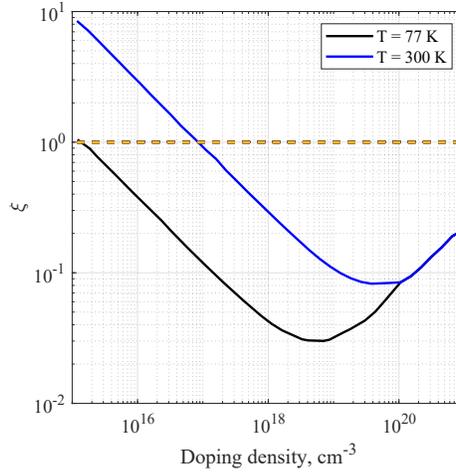


Figure 1.7: Number of screening charges ξ in the Debye sphere as a function of N_i in silicon at 77 and 300 K. The condition for which the single-ion scattering approximation is valid is $\xi \gg 1$ (yellow dashed line)

the field; this causes a change of momentum and position of the particles:

$$\Delta \mathbf{k} = -\frac{e\mathcal{E}}{\hbar} \tau, \quad (1.32)$$

$$\Delta r = \langle v \rangle_{\tau} \tau, \quad (1.33)$$

where the symbol

$$\langle . \rangle$$

indicates the time average over τ . Only the components of momentum and position along the direction of the field are affected.

3. after the drift part, particles **scatter**, selecting a scattering mechanism among all included according to their probability to happen (the scattering rates), determined by particles' energy.
4. after the scattering mechanism is selected for each particles, the **final state** is chosen, selecting the new **k**-vector according to the related probability distribution. For some scatterings the distribution is uniform in θ and ϕ and the scattering is called isotropic; for some others the distribution is anisotropic and is expressed by an analytical formula.
5. the algorithm is repeated until the simulation time has passed; it has to be sufficiently long in order to reach steady state.

The algorithm described above is illustrated as a flowchart in Figure 1.8.

It's important that the time step τ is chosen such that the probability distribution of the selection of the scattering mechanism is well-defined: indicating with W_T the sum of each scattering rate, the following relation should be satisfied:

$$\tau W_T \leq 1. \quad (1.34)$$

The outcome of the simulation is not the general solution of the BTE $f(\mathbf{x}, \mathbf{k}, t)$ but some moments of it, like the energy and velocity distribution, some of which are described below.

- The temporal average of the velocity is equal to

$$v = \frac{1}{FTN} \sum_{t=0}^T (\epsilon_f - \epsilon_i), \quad (1.35)$$

where N is the number of particles, ϵ_i and ϵ_f are the initial and final energy at time t and T is the simulation time. An example of this quantity, extracted from a simulation, is reported in figure 1.9. Looking at this figure, it is possible to see if the simulation has reached convergence: steady-state is reached if the velocity is constant after a certain period of time.

- Another quantity of interest is the ensemble average of the velocity in the direction of the field, reported in figure 1.11.

$$v(t) = \frac{1}{N} \sum_{i=0}^N v_i(t). \quad (1.36)$$

- Finally, the the ensemble-averaged energy is also useful because it gives information about convergence and on the reliability of the results: for example, the energy should increase with increasing temperature and/or increasing electric field. An example of this outcome is shown in figure 1.10

If there is no interaction between simulated particles, the time evolution is an ergodic process (i.e. the ensemble average is equal to the time average) and material properties can be extracted even by simulating a single particle. The advantages of using more particles is that the statistic analysis is improved for the same simulation time and it becomes possible to study effects that involves interaction between carrier themselves (e.g. screening).

Direct, combined and rejection technique

In order to choose the scattering mechanism after the drift part, or to select the final k-state after a scattering mechanism is chosen, it is necessary to extract a possible outcome of a random variable following a certain probability

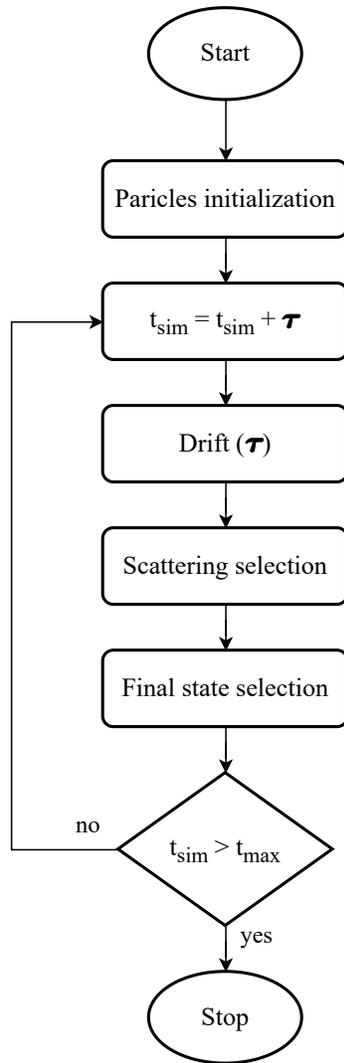


Figure 1.8: Flow chart representing the Monte Carlo algorithm for carrier transport

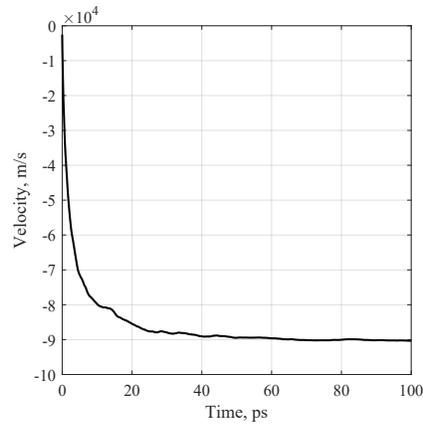


Figure 1.9: Time average of carrier velocity extracted from a Monte Carlo simulation

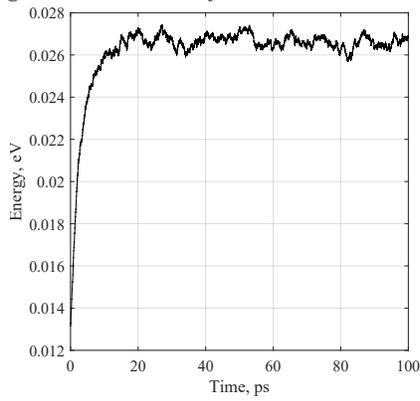


Figure 1.10: Ensemble average of carrier energy extracted from a Monte Carlo simulation

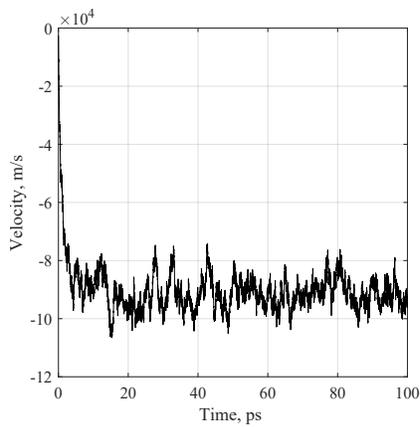


Figure 1.11: Ensemble average of carrier velocity along the direction of the electric field, extracted from a Monte Carlo simulation

distribution. The procedure involves using a random variable R with uniform distribution between 0 and 1.

Given a continuous or discrete random variable with probability density function $f(x)$, and r_1 is an outcome of the random variable R , it is possible to extract a value x_1 following the probability distribution $f(x)$ from the following equation

$$r_1 = \int_0^{x_1} f(x) dx . \quad (1.37)$$

If the solution of the integral 1.37 is known analytically and it is invertible, the evaluation of x_1 is straightforward.

$$x_1 = F^{-1}(r_1), \quad (1.38)$$

where F is a primitive of f . Such method to find x_1 is called **direct technique**.

However, if that is not the case, an iterative algorithm can be applied. This algorithm consists in finding a function $g(x)$, whose primitive over dx is known analytically and invertible, such that

$$g(x) \geq f(x) \quad \forall x \in [a, b], \quad (1.39)$$

where $[a, b]$ is the interval where the function $f(x)$ is defined. Once $g(x)$ is found, a number x_1 is chosen using the direct technique with g as probability distribution. Then, another uniform random number r_2 is chosen between 0 and 1. If the following condition applies

$$f(x_1) > r_2 g(x_1), \quad (1.40)$$

then x_1 is an outcome of the random variable following the probability density function $f(x)$; otherwise, another value x_1 must be chosen and the procedure is repeated. According to the type of $g(x)$ this method is called:

- **rejection technique** if $g(x) = const$, i.e., g is a uniform distribution over $[a, b]$;
- **combined technique** if $g(x)$ is a generic function.

It is worthwhile pointing out that the more $\int_a^b |g(x) - f(x)| dx$ is small, the less number of attempts are required to find a correct x_1 . Another important thing to mention is that a good algorithm for the generation of pseudo-random number must be used to ensure a good statistics of the simulation.

Focusing on the selection of the final k-state, the angle ϕ is usually given by a uniform distribution between 0 and 2π :

$$\phi = 2\pi r_\phi, \quad (1.41)$$

where r_ϕ is one of the possible outcomes of the random variable R . As for θ , an angle θ_r is usually selected with the direct or rejection technique.

Device Monte Carlo

Using the Monte Carlo method it is possible to simulate electronic devices as well. The BTE 1.1 is coupled with the Poisson equation

$$\nabla^2\phi = -\frac{\rho}{\epsilon} = -\frac{e}{\epsilon}(-n(\mathbf{x}) + p(\mathbf{x}) + N_D^+(\mathbf{x}) - N_A^-(\mathbf{x})), \quad (1.42)$$

where ϕ is the electrostatic potential, ρ is the charge density, $n(\mathbf{x})$, $p(\mathbf{x})$, $N_D^+(\mathbf{x})$ and $N_A^-(\mathbf{x})$ are the space distributions of electron density, hole density, ionized donors density and ionized acceptors density, respectively. The two coupled equations are solved self-consistently according to the following procedure:

1. the geometry of the device is designed, the boundary conditions (i.e. the borders of the device for the transport equation and the potential at the contacts for Poisson equation) are defined. Then, a mesh for the Poisson equation is established; it can be solved using a finite difference (FD) or finite element method (FEM), from which the choice of the mesh is derived (structured or unstructured).
2. A number of simulated particles is defined, each having a portion of the actual charge of the free-carriers of the device, obtained solving the neutrality equation.

$$-n(\mathbf{x}) + p(\mathbf{x}) + N_D^+(\mathbf{x}) - N_A^-(\mathbf{x}) = 0. \quad (1.43)$$

The particles distribution inside the device will determine the profile of the electric field, so high accuracy is obtained by simulating a high number of particles.

3. The Poisson equation is solved to determine the electric field inside the device.
4. The BTE is solved with a Monte Carlo approach using the electric field evaluated previously as input, to determine the displacement of carrier in real and reciprocal space.

The step 3. and 4. are repeated until a steady state is reached.

A critical aspect of device Monte Carlo simulations is that charge is stored in particles, that can occupy any position in the device within the machine precision, whereas the Poisson equation is solved on the nodes of the mesh. Therefore, it is necessary to map the particles' charge on the mesh nodes: this procedure is called *particle-mesh coupling* [24]. When using an unstructured mesh, particle-mesh coupling can lead to nonphysical self-forces that needs to be properly corrected [2].

The algorithm described above is illustrated as a flowchart in Figure 1.12.

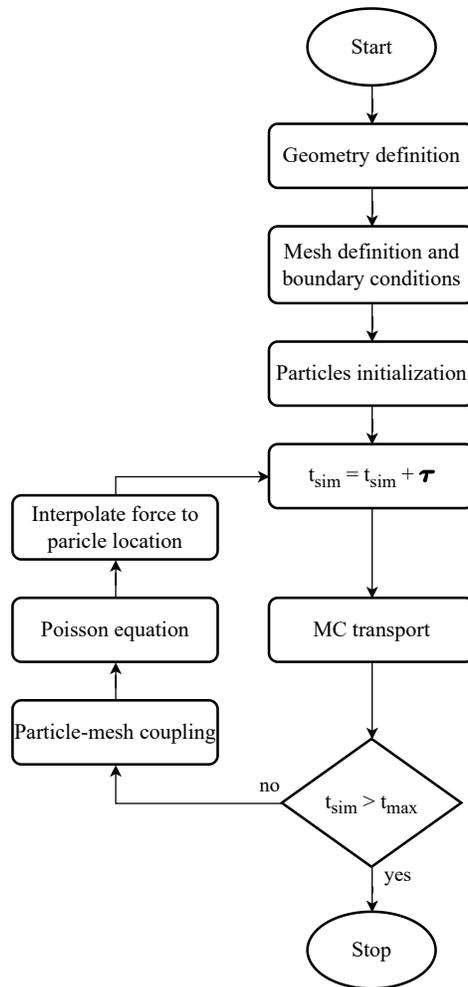


Figure 1.12: Flow chart of the Monte Carlo method for device simulation. MC transport is an abbreviation for the main loop in figure 1.8

Chapter 2

Advanced models for Monte Carlo simulations

In this work, advanced physical models for the Monte Carlo method have been implemented and/or investigated in order to obtain a better description of carrier transport in the following conditions:

- high doping density and degeneracy;
- high energy.

Two ionized impurity scattering models, that attempt to improve Brooks-Herring one are presented. The first was proposed by Ridley [38]; it has been implemented because it is able to unify the physics behind BH and CW models, trying to deal with ionized impurity scattering at any level of screening in the single-ion scattering approximation. The second was proposed by Kosina [29] and has been analyzed as an attempt to overcome the limitation of the BH model such as the single-ion scattering, the use of the first Born cross section and the approximations on the screening function. It was chosen because it proved to work quite well in predicting silicon mobility at 300 K for doping densities up to $1 \times 10^{20} \text{ cm}^{-3}$, as can be seen in figure 2.1.

Since high doping density is related to high carrier density and degeneracy, Pauli exclusion principle was included as well, following the work of [32].

These models make use of an analytic description of the electronic band-structure and are used and compared in Chapter 3, where Monte Carlo method is used for the evaluation of the electron low-field mobility of InAs, InSb and InAsSb.

Finally, a novel method to include a multi-band electronic dispersion relation has been analyzed. It is better compared to a simpler analytic band description, because it provides a more exact evaluation of the scattering rates and a better description of transport especially in high energy conditions, since carrier can scatter to higher bands and can undergo impact ionization ([15], [16]).

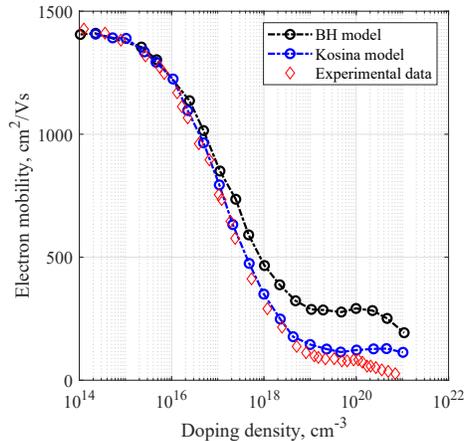


Figure 2.1: Electron mobility in silicon at $T = 300$ K, data reproduced from [29]

This model has been used in Chapter 4 to simulate a device based on a novel photodetector design [13].

2.1 Analytic band electronic structure

In this framework, the electronic bandstructure is approximated by a set of valleys, each valley having the following dispersion relation

$$E(1 + \alpha_v E) = \frac{\hbar^2 k^2}{2m_v^*}, \quad (2.1)$$

where E is the electron energy in the valley reference system (i.e. $E = 0$ at the bottom of the valley), m_v^* and α_v are the valley effective mass and non-parabolicity factor, respectively. The effective mass can be isotropic (corresponding to spherical valleys) or anisotropic (corresponding to elliptic valleys). However, the anisotropic case can be reduced to the isotropic case through the Herring and Vogt transformations [24]. The analytic band description relies on the fact that, due to scattering processes that release energy (e.g. phonon relaxation) electrons tend to stay in the local minima of the lowest energy bands. Usually, the valleys employed in a Monte Carlo simulation are the first two or three of the conduction band (such as the ones corresponding to Γ , X, L points of the irreducible wedge) and the ones of the heavy holes and light holes bands. This description allows for the calculation of realistic transport properties even at moderately high electric field [14]. In the following analytic models, the overlap integral is always approximated to 1 and valleys are considered spherical. From equation 2.1, the wavevector k can be obtained from

$$k = \sqrt{\frac{2m_v^*}{\hbar^2} E(1 + \alpha E)} \quad (2.2)$$

and the group velocity v as

$$v(k) = \frac{1}{\hbar} \frac{\partial E}{\partial k} = \frac{\hbar k}{m^*(1 + 2\alpha E)}. \quad (2.3)$$

2.1.1 Ridley model theory

This model investigates the meaning of the two-body approximation, finding a way to re-conciliate the two seemingly distinct models for ionized impurity scattering in Monte Carlo simulations: the Brooks-Herring and Conwell-Weisskopf models.

From classical collision theory, a relation exists between the scattering differential cross section $\sigma(\theta)$ and the impact parameter b (equation 1.18) that can be rewritten in integral form as:

$$\int_{\Omega} \sigma d\Omega = \int_b 2\pi b db. \quad (2.4)$$

Recalling that $d\Omega = \sin(\theta) d\theta d\phi$, equation 2.4 becomes:

$$b^2 = \int_{\theta}^{\pi} 2\sin(\theta')\sigma(\theta') d\theta'. \quad (2.5)$$

This means that, in a two-body problem, the scattering angle is linked to a length quantity, namely the impact parameter b .

Ridley compared this quantity with the average distance between impurities

$$a = (2\pi n_i)^{-1/3}. \quad (2.6)$$

He supposed that the probability of scattering with an angle θ does not depend only on the ionized center with which the particle scatters, but also on the centers lying between them, namely the centers distant from the particle less than the impact parameter b : the effect of this centers is to reduce the effect of former one, decreasing the probability of scattering at angle θ . The probability of there being no scattering center with impact parameter less than b is:

$$P(b) = e^{-\pi b^2 N a}, \quad (2.7)$$

where N is the density of ionized impurities. This contribution is multiplied to the original cross section giving:

$$\sigma'(\theta) = \sigma(\theta) e^{-N a \int_{\theta}^{\pi} 2\pi \sin(\theta') \sigma(\theta') d\theta'}. \quad (2.8)$$

The next step is to substitute $\sigma(\theta)$ with the Brooks-Herrings cross-section 1.29, obtaining the Ridley model

$$\sigma_R(\theta) = \frac{R^2}{4(x^2 + \beta^{-1})^2} e^{-\pi R^2 N a (1-x^2)/[(x^2 + \beta^{-1})(1 + \beta^{-1})]}, \quad (2.9)$$

where $x = \sin(\theta/2)$ (see BH and CW models for the values of R and β^{-1}). The total scattering rate can then be evaluated using equation 1.16.

It is possible to notice that:

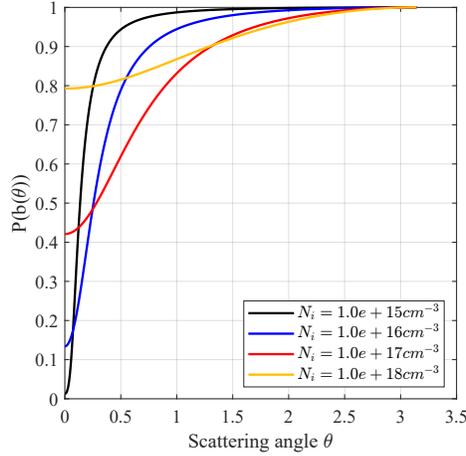


Figure 2.2: $P(b(\theta))$ in function of the scattering angle, evaluated using $m^* = 0.05m_0$, $E = 10meV$, $T = 300$ K and $\epsilon_r = 15.15$

- when q_0 (the screening parameter, inverse of Debye length) is high, the values of σ_R are similar to BH total cross section: this is motivated by the fact that if screening is high the possible range of impact parameters are small compared to the distance between impurities ($b \ll a$) and so the term $P(b)$ has little influence;
- when q_0 is small, the possible range of impact parameters are comparable with the distance between impurities ($b \sim a$) and then the term $P(b)$ is significant. The differential cross section could be imagined as the result of the interaction with a Coulomb potential with low screening (then quite similar to a bare Coulomb potential) that is suppressed for impact parameters comparable to the average distance between the impurities. In this way the CW model is retrieved.

The probability $P(b)$ is shown in figure 2.2. It is evaluated considering carrier density equal to doping density, i.e. in case of no compensation. It is possible to see deviation from the BH cross section (where $P(b)$ is neglected and supposed equal to 1) especially at small scattering angles.

Comparison between Ridley model and BH and CW models In figure 2.3 the scattering angle probability amplitude of BH, CW and Ridley models are compared. For doping density equal to carrier density (i.e. high screening, figure 2.3b) the Ridley and BH cross sections are quite similar, both different from the CW one. When the carrier density is lower than the doping density, as in the case of compensated semiconductors (figure 2.3a), the Ridley cross section is more similar to the CW one.

Figure 2.4 shows the comparison of the scattering rates. It can be seen that Ridley line lies almost below the minimum of BH and CW lines: by tuning

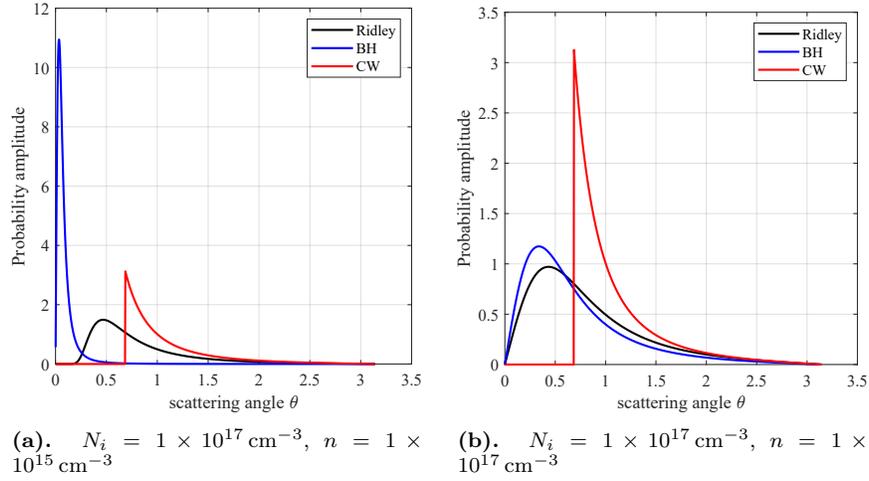


Figure 2.3: Probability distribution of scattering angle θ at $T = 300 \text{ K}$, $m^* = 0.05m_0$, $E = 10 \text{ meV}$ and $\epsilon_r = 15.15$

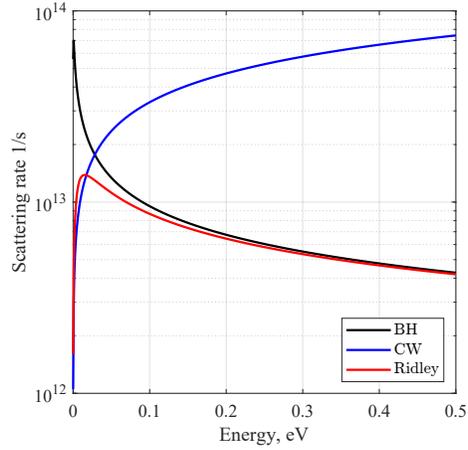


Figure 2.4: Comparison between BH, CW and Ridley scattering rate for $N_i = 1 \times 10^{17} \text{ cm}^{-3}$, $m^* = 0.1m_0$, $T = 300 \text{ K}$ and $\epsilon_r = 15.15$

the average distance between impurities in Ridley model, this feature could be retrieved exactly.

Another important advantage of the Ridley model is that it removes the high scattering rate for low energy, making this model more computationally efficient compared to BH model.

2.1.2 Ridley model implementation algorithm

Total scattering rate It is possible to integrate Ridley cross section to obtain an analytical formula for the scattering rate:

$$W_R = \frac{v_g}{a} (1 - e^{-\frac{aW_{BH}}{v_g}}), \quad (2.10)$$

where v_g is the velocity of the particle, a is the average distance between impurities 2.6 and W_{BH} is the Brooks-Herring scattering rate 1.30.

Final state The procedure for the evaluation of the final state follows [11]. The solution for the equation 2.5 using BH cross-section as $\sigma(\theta)$ is [38]

$$b^2 = \frac{R^2(1-x^2)}{(x^2 + \beta^{-1})(1 + \beta^{-1})}. \quad (2.11)$$

The impact parameter for a scattering particle varies from 0 to the maximum impact parameter obtainable with the Brooks-Herring model, putting $\theta = 0$ in equation 2.11.

$$b_m^2 = \frac{R^2}{\beta^{-1}(1 + \beta^{-1})}. \quad (2.12)$$

Recalling equations 1.18 and 2.9, the probability distribution of b is proportional to

$$f(b) = 2\pi b e^{-\pi a b^2 N_i}. \quad (2.13)$$

This distribution is integrable and invertible so the direct technique can be used to sample a value for the impact parameter following the distribution f

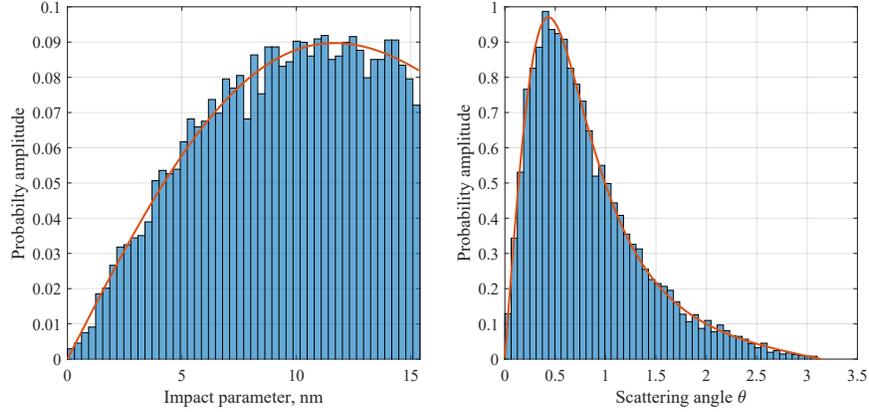
$$r_1 = \frac{e^{-\pi a b^2 N_i} - e^{-\pi a b_m^2 N_i}}{1 - e^{-\pi a b_m^2 N_i}}, \quad (2.14)$$

$$b^2 = -\frac{1}{\pi a N_i} \log(r_1(1 - e^{-\pi a b_m^2 N_i}) + e^{-\pi a b_m^2 N_i}). \quad (2.15)$$

The passage from b to θ is obtained again from equation 2.11

$$\cos(\theta) = 1 - 2 \frac{R^2 - (1 + \beta^{-1})\beta^{-1}b^2}{(1 + \beta^{-1})b^2 + R^2}. \quad (2.16)$$

In figure 2.5a, the histogram representing the extraction of b values from a uniform random variable between 0 and 1 (as in the Monte Carlo method) is shown. As is can be seen, it follows the theoretical distribution 2.13. In figure 2.5b, the histogram representing the extraction of θ values with the above mentioned procedure is shown. It follows the theoretical distribution given by equation 2.9.



(a). Impact parameter distribution: the continuous line represents the theoretical distribution given by equation 2.13; the histogram represents the distribution obtained through the direct technique

(b). Scattering angle θ distribution: the continuous line represents the theoretical distribution given by equation 2.16; the histogram represents the distribution obtained through the direct technique

Figure 2.5: Distribution of impact parameters and scattering angle for the Ridley model for $N_i = 1 \times 10^{17} \text{ cm}^{-3}$ and $T = 300 \text{ K}$

2.1.3 Kosina model theory

This model was developed as an attempt to better predict majority electrons mobility in silicon compared to the BH model, since the latter proved to over-estimate it for doping levels higher than $1 \times 10^{18} \text{ cm}^{-3}$. The Kosina model is evaluated using the Born approximation, as the BH model, but introduces the following improvements.

Screening function The screening function q_0 is not constant and equal to 1.28, as in the case of BH model, but depends on the exchanged momentum between the final and the initial state $q = |\mathbf{k}' - \mathbf{k}|$ (and so it depends on the scattering angle θ):

$$\beta_s^2 = q_0^2 \frac{\mathcal{F}_{-1/2}(\eta)}{\mathcal{F}_{1/2}(\eta)} G(q, \eta), \quad (2.17)$$

where $\mathcal{F}_x(\eta)$ is the Fermi-Dirac integral of order x and η is equal to

$$\eta = \frac{E_F - E_c}{k_b T} \quad (2.18)$$

for electrons.

$G(q, \eta)$ is a complicated function obtained using Lindhard screening theory [17]:

$$G(\xi, \eta) = \frac{1}{\mathcal{F}_{-1/2}(\eta)} \frac{1}{\xi \sqrt{\pi}} \int_0^\infty \frac{x}{1 + \exp(x^2 - \eta)} \log \left(\left| \frac{x + \xi}{x - \xi} \right| \right) dx, \quad (2.19)$$

f_1	$-\frac{2\mathcal{F}_{-3/2}(\eta)}{3\mathcal{F}_{-1/2}(\eta)}$	g_1	$\frac{\mathcal{F}_{1/2}(\eta)}{2\mathcal{F}_{-1/2}(\eta)}$
f_2	$\frac{4\mathcal{F}_{-5/2}(\eta)}{15\mathcal{F}_{-1/2}(\eta)}$	g_2	$\frac{\mathcal{F}_{3/2}(\eta)}{4\mathcal{F}_{-1/2}(\eta)}$
f_3	$-\frac{8\mathcal{F}_{-1/2}(\eta)}{105\mathcal{F}_{-1/2}(\eta)}$	g_3	$\frac{3\mathcal{F}_{5/2}(\eta)}{8\mathcal{F}_{-1/2}(\eta)}$

Table 2.1: Coefficients of the screening function linear system

where

$$\xi^2 = \frac{\hbar^2 q^2}{m^* k_B T}. \quad (2.20)$$

In order to be implemented in a Monte Carlo simulation, this formula is approximated by an rational function

$$G(q, \eta) \approx \frac{1 + a\xi^2 + b\xi^4}{1 + c\xi^2 + d\xi^4 + e\xi^6}. \quad (2.21)$$

The coefficients of 2.21 are found computing the Taylor series expansions of 2.19 for $\xi \rightarrow 0$ and $\xi \rightarrow \infty$ and equating them to the Taylor expansions of 2.21 [29]. This results in a 3x3 linear system that can be solved with the Gauss elimination method (A.3)

$$\begin{bmatrix} f_1 & 1 & -g_1 \\ -1 & g_1 & g_2 \\ uf_2 + g_1 & uf_1 + g_2 & u + g_3 \end{bmatrix} \begin{bmatrix} c \\ d \\ e \end{bmatrix} = \begin{bmatrix} -f_2 \\ f_1 \\ 1 - uf_3 \end{bmatrix},$$

$$a = c + f_1,$$

$$b = g_1 e.$$

The list of coefficients for the system is reported in table 2.1 and $u = 7.2(1 + 10^{\eta/5})$.

This rational approximation deteriorates for high degeneracy ($\eta > 5$, [29]). The shape of the momentum-dependent screening function is reported in figure 2.6.

Multiple scattering The perturbing potential is the one generated by two ionized atoms

$$V_0(r) = \frac{Ze}{4\pi\epsilon_0\epsilon_r} \left(\frac{1}{r} + \frac{1}{r-R} \right), \quad (2.22)$$

where $R = (2\pi n_i)^{-1/3}$.

Using the Born approximation, the same procedure used to obtain Brooks-Herring model, one finds the following differential cross section:

$$\sigma(q) \propto |V(q)|^2 = \frac{1}{(q^2 + \beta_s^2 G(\xi, \eta))^2} \left(1 + \frac{\sin(qR)}{qR} \right). \quad (2.23)$$

In this formula, both two-ions correction and the q -dependent screening function are considered. Is it possible to notice that, using the approximations $G = 1$ and $R \rightarrow \infty$ the BH model is retrieved.

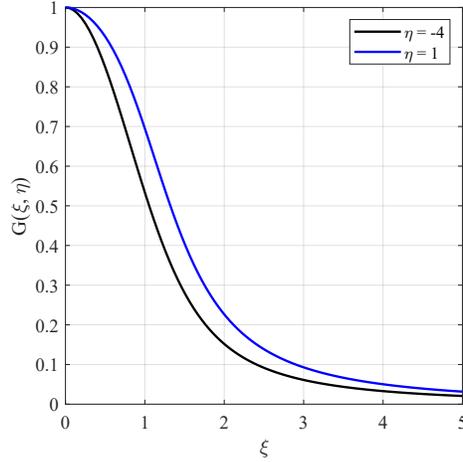


Figure 2.6: q -dependent screening function $G(q, \eta)$

Second Born correction Since this model is evaluated using the first Born approximation, a further correction is included to take in considerations the local modification of the wave-function near the impurity center. The correction to the cross section is the following:

$$\Delta\sigma(q) = s_0(k)\sigma_{B1}(q), \quad (2.24)$$

$$s_0 = \frac{a}{1 + \frac{4k^2}{\beta_s^2} - a} \quad a = \frac{U_0}{\beta_s^2} \left(1 - \frac{U_0}{4\beta_s}\right), \quad (2.25)$$

where U_0 is defined as

$$U_0 = \frac{2m^*}{\hbar^2} \frac{Ze^2}{4\pi\epsilon_0\epsilon_r}. \quad (2.26)$$

$\sigma_{B1}(q)$ is the BH cross section, meaning that this correction is interpreted as a further scattering mechanism of BH-type, whose scattering rate is wighter by the term s_0 . Since it is found that the correction term for the first Born cross section is always grater than 1 [35], and that the correction term s_0 becomes negative for $F = \frac{|U_0|}{\beta_s} > 4$, this correction is applied only if $F < 4$, namely if doping is sufficiently high.

Comparison between Kosina and BH models In this paragraph, the scattering rate and the final state probability distribution using the Kosina model are calculated and compared to the ones obtained from the BH model.

Kosina scattering rate, considering two-ion scattering and momentum-dependent screening function is equal to

$$\lambda(k) = C(k) \int_0^{2k} \frac{1}{(q^2 + \beta_s^2 G(\xi, \eta))^2} \left(1 + \frac{\sin(qR)}{qR}\right) q dq, \quad (2.27)$$

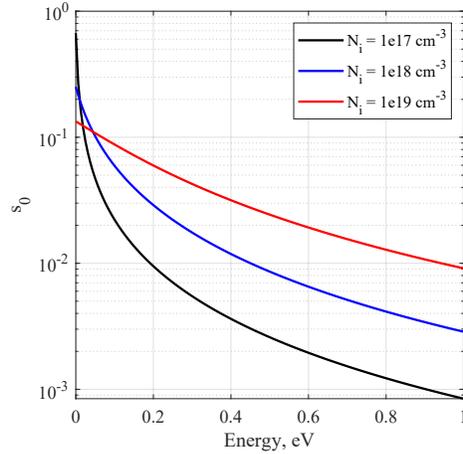


Figure 2.7: Values of the second Born correction parameters s_0 for different doping levels, $T = 77$ K, $m^* = 0.05m_0$, $\eta = 1$, $\epsilon_r = 15.15$, parabolic bands

with

$$C(k) = \frac{N_i Z^2 e^4}{2\pi \hbar^2 (\epsilon_0 \epsilon_r)^2 v_g(k)}. \quad (2.28)$$

The second Born correction is evaluated as:

$$\Delta\lambda(k) = s_0(k)C(k) \int_0^{2k} \frac{q}{(q^2 + \beta_s^2)^2} dq. \quad (2.29)$$

The integrals are performed using the Simpson method A.2 with number of integration steps $N = 1001$ to ensure the correct evaluation of the integral in the energy range of interest $[0 \div 10]$ eV.

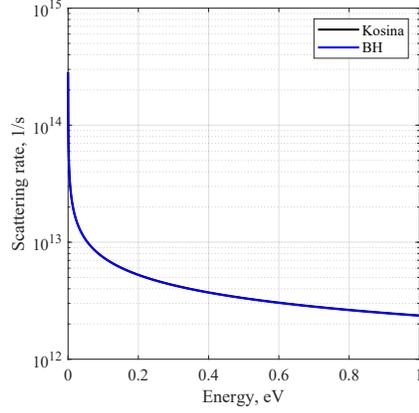
The values of s_0 are reported in figure 2.7. They are all smaller than 1 and decrease rapidly as the electron energy increases. Compared to BH model, this contribution results in a slight increase of the scattering rate, especially in the low-energy range and no impact on the final state distribution.

The other two contribution are more interesting and are investigated more in details.

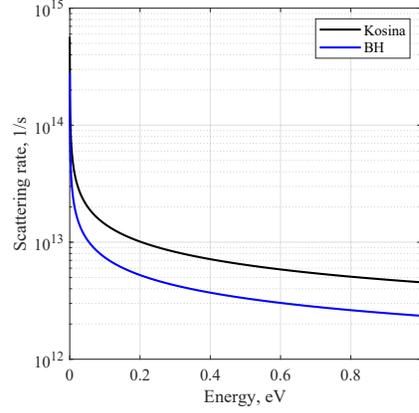
Figure 2.8 reports the scattering rate at $T = 300$ K for low doping level ($N_i = 1 \times 10^{15}$). In figure 2.8a only the q -dependent screening function and the second Born correction are considered, while in figure 2.8b the complete model is shown. It is possible to see that for this doping level the screening function does not impact on the scattering rate, while the two-ion contribution is appreciable and causes an increase of the scattering rate of almost two times.

Figure 2.10 reports the probability distribution of the final state at $N_i = 1 \times 10^{15} \text{ cm}^{-3}$. It is very similar to the BH probability distribution, with a peak for small q (and as a consequence, a small scattering angle θ).

Scattering rates for high doping density $N_i = 1 \times 10^{18} \text{ cm}^{-3}$ at $T = 300$ K are reported in figure 2.9. In this case, the most relevant contribution is

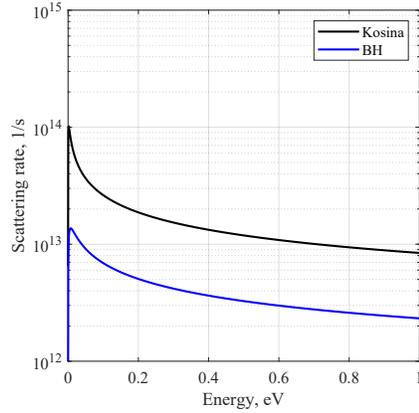


(a). q -dependent screening included only

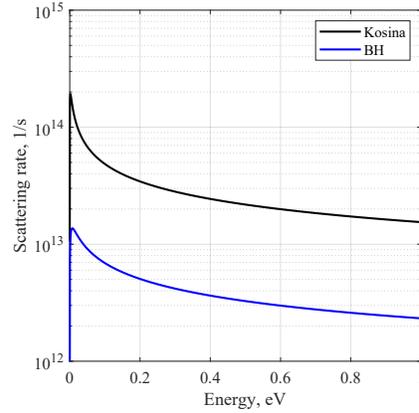


(b). final model

Figure 2.8: Comparison between BH and Kosina scattering rates, $T = 300$ K, $N_i = 1 \times 10^{15} \text{ cm}^{-3}$, $m^* = 0.05m_0$, $\eta = -4$, $\epsilon_r = 15.15$, parabolic bands



(a). q -dependent screening included only



(b). final model

Figure 2.9: Comparison between BH and Kosina scattering rates, $T = 300$ K, $1 \times 10^{18} \text{ cm}^{-3}$, $m^* = 0.05m_0$, $\eta = 5$, $\epsilon_r = 15.15$, parabolic bands

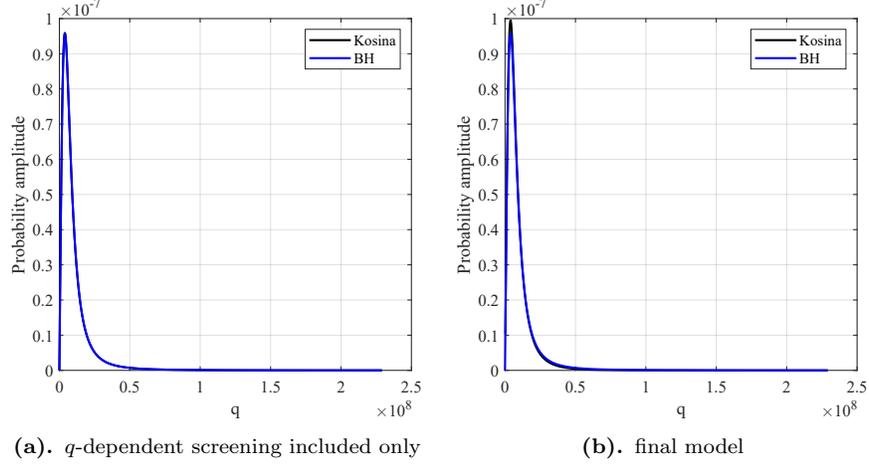


Figure 2.10: Comparison between BH and Kosina probability distribution for the state after scattering, $T = 300$ K, $N_i = 1 \times 10^{15} \text{ cm}^{-3}$, $m^* = 0.05m_0$, $\eta = -4$, $\epsilon_r = 15.15$, $E = 10$ meV, parabolic bands

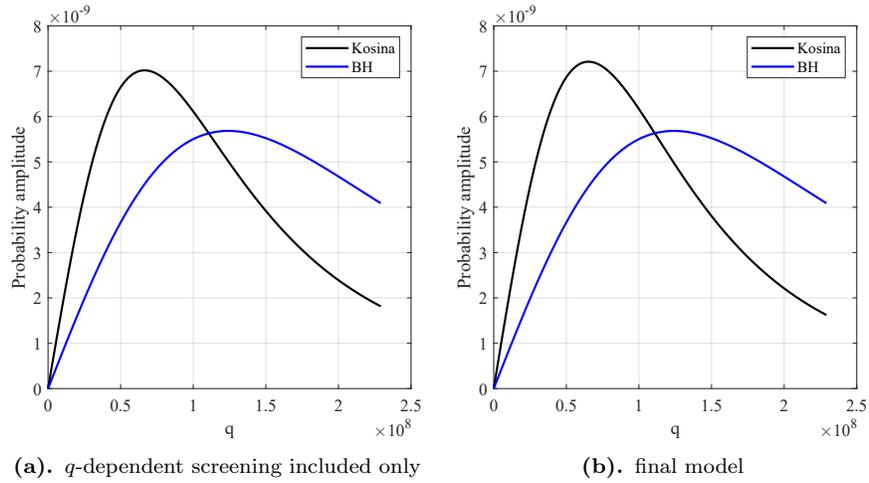


Figure 2.11: Comparison between BH and Kosina probability distribution for the state after scattering, $T = 300$ K, $N_i = 1 \times 10^{18} \text{ cm}^{-3}$, $m^* = 0.05m_0$, $\eta = 5$, $\epsilon_r = 15.15$, $E = 10$ meV, parabolic bands

effect of the degeneracy (included in the screening function part), that increases the scattering rates of almost one order of magnitude. The two-ions correction provides a further increase.

As for the final state distribution for $N_i = 1 \times 10^{18} \text{ cm}^{-3}$ and $T = 300 \text{ K}$, it is reported in figure 2.11. Also in this case the most relevant contribution is given by the screening function, that shifts the average exchanged momentum to lower q .

2.1.4 Kosina model implementation algorithm

As preliminary calculation for the scattering rate and the final state, the Fermi-Dirac integral employed in the evaluation of β_s 2.17 and the q -dependent screening function 2.21 is obtained using [20].

The position of the Fermi level is obtained solving self-consistently the neutrality equation and the equations for carrier densities:

$$\begin{cases} -n + p + N^+ = 0, \\ n = \frac{8\pi\sqrt{2}m^{*3}}{\hbar^3} \int_{E_g}^{\infty} \sqrt{\gamma(E)} d\gamma f_d(E, E_f) dE, \\ p = \frac{8\pi\sqrt{2}m^{*3}}{\hbar^3} \int_{-\infty}^0 \sqrt{\gamma(E)} d\gamma f_d(E, E_f) dE, \end{cases} \quad (2.30)$$

where $\gamma(E) = E(1 + \alpha E)$, $d\gamma = 1 + 2\alpha E$, α is the non-parabolicity factor and $f_d(E, E_f)$ is the Fermi-Dirac distribution

$$f(E, E_f) = \frac{1}{e^{(E-E_f)/k_bT} + 1}. \quad (2.31)$$

For the neutrality equation the bisection method was used (A.1).

The second Born correction is treated as an additional Brooks-Herring mechanism. The scattering rate is evaluated as

$$W_{2\text{Born}} = s_0 \frac{2^{5/2}\pi N_i Z^2 e^4}{(4\pi\epsilon)^2 E_\beta^2 (m^*)^{1/2}} \frac{\sqrt{E}}{1 + 4E/E_\beta}, \quad (2.32)$$

where $E_\beta = \frac{\hbar^2 \beta_s^2}{2m^*}$ and E is the energy of the particles (as in equation 1.30). As for the final state, the following procedure is applied in order to follow the distribution given by 1.29 [24]:

- A random number between 0 and 1 is extracted;
- the scattering angle θ is obtained through

$$\cos(\theta) = 1 - \frac{2(1-r)}{1 + 4rE/E_\beta}. \quad (2.33)$$

The implementation of the other two corrections is described below.

Method No.1

Scattering rate The scattering rate can be evaluated using numerical integration of equation 2.27. This has the advantage of not having the necessity of recurring to self-scattering, but it can be computationally expensive. However, using parallelization over the values of initial energy, the processing time is comparable to the other scattering rate calculations.

Final state As for final state, the combined technique is used, using a Brooks-Herring-type cross-section as a supremum:

$$p_{\text{sup}}(q) = \frac{2}{(q^2 + \beta_{\text{min}}^2)^2} \geq \frac{1}{(q^2 + \beta_s^2 G(\xi, \eta))^2} \left(1 + \frac{\sin(qR)}{qR}\right), \quad (2.34)$$

where $\beta_{\text{min}}^2 = \beta_s^2 G(2k)$.

1. A scattering angle θ_r is generated according to equation 2.33;
2. θ_r is converted to q_r using

$$q = \sqrt{2k^2(1 - \cos(\theta_r))}; \quad (2.35)$$

3. a second random number between 0 and 1 r_2 is generated;
4. if the following condition is satisfied:

$$\frac{1}{(q_r^2 + \beta_s^2 G(\xi, \eta))^2} \left(1 + \frac{\sin(q_r R)}{q_r R}\right) > r_2 p_{\text{sup}}(q_r), \quad (2.36)$$

then θ_r is accepted, otherwise the procedure is repeated from step 1.

Method No.2

Scattering rate In order to avoid the numerical integration of 2.27, the self-scattering technique can be employed: the used scattering rate is higher than the actual one; then, at the selection of the final state, scatterings are accepted or rejected by means of the combined technique. The scattering rates are evaluated from p_{sup}

$$W_{\text{sup}} = 2 \frac{2^{5/2} \pi N_i Z^2 e^4}{(4\pi\epsilon)^2 E_\beta^2 (m^*)^{1/2}} \frac{\sqrt{E}}{1 + 4E/E_\beta}, \quad (2.37)$$

where $E_\beta = \frac{\hbar^2 \beta_{\text{min}}^2}{2m^*}$. However, since the scattering rate related to p_{sup} would be too high to be used in a Monte Carlo simulation ($> 1 \times 10^{15} \text{ s}^{-1}$), the momentum relaxation times are used instead as equivalent isotropic scattering rates. In fact, for an anisotropic scattering process with a high preference for forward scattering (as ionized impurity scattering), the momentum relaxation rate τ_m is

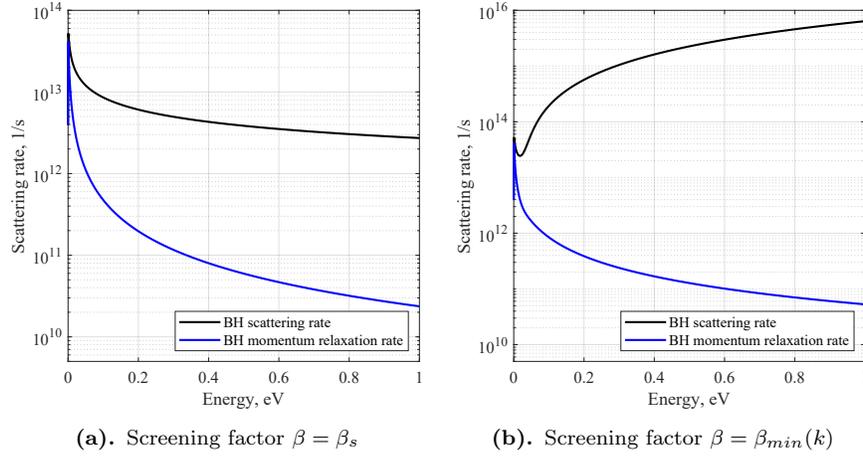


Figure 2.12: Comparison between BH scattering rates and momentum relaxation rates for $N_i = 1 \times 10^{17} \text{ cm}^{-3}$, $T = 300 \text{ K}$, $m^* = 0.05m_0$, $\eta = 0.05$, $\epsilon_r = 15.15$, parabolic bands

always smaller than the total scattering rate [33]. Recalling equation 1.14 and the fact that, for an elastic process

$$q^2 = |\mathbf{k}' - \mathbf{k}|^2 = 2k^2(1 - \cos(\theta)) \quad (2.38)$$

equation 1.12 can be rewritten as

$$\tau_m^{-1} = N_i v_g \left(\frac{\pi}{k^4} \int_0^{2k} \sigma(q) q^3 dq \right). \quad (2.39)$$

And for the BH model, it is equal to

$$\tau_m^{-1} = C(k) \frac{1}{4k^2} \left(\log(1+b) - \frac{b}{1+b} \right), \quad (2.40)$$

where $b = 4k^2/\beta_s^2$. Figure 2.12 shows a comparison between BH scattering rate and momentum relaxation rate: figure 2.12a rates are evaluated considering the screening factor $\beta = \beta_s$, while figure 2.12b rates are evaluated with $\beta = \beta_{min}$ as in p_{sup} . From the latter figure it is evident that the anisotropic scattering rates cannot be used in a Monte Carlo code.

Final state As for the final state, in order to apply the combined technique, the following procedure was applied:

1. a random number $q_r \in [0, 2k]$ is chosen according to the density function

$$f(q) = \frac{q^3}{(q^2 + \beta_{min}^2)^2}. \quad (2.41)$$

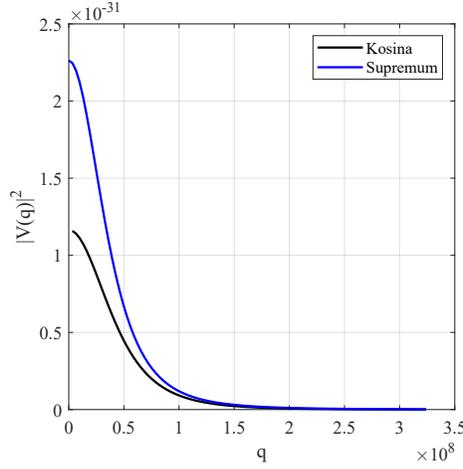


Figure 2.13: profile of $p_{sup}(q)$ (Eq. 2.34, blue line) and Kosina $|V(q)|^2$ (Eq. 2.23, black line) for $N_i = 1 \times 10^{17} \text{ cm}^{-3}$ and $T = 300 \text{ K}$, $E = 20 \text{ meV}$

The shape of the function f allows for the use of rejection technique without having too much iterations. The used supremum is the maximum of the function obtained for $q = \sqrt{3\beta_{\min}}$ (red dashed line in figure 2.14a). Figure 2.14a shows the profile of $f(q)$ and figure 2.14b shows the distribution of the samples obtained using the rejection technique;

2. a random number p_r is chosen evenly distributed between 0 and $p_{sup}(q_r)$;
3. if $p_r < |V(q_r)|^2$ (equation 2.23) then the scattering is accepted, otherwise it is rejected and self-scattering is performed. Figure 2.13 shows the profile of $p_{sup}(q)$ and Kosina $|V(q)|^2$;
4. the accepted scattering is elastic so a random ϕ and $\cos(\theta)$ are chosen using a uniform distribution.

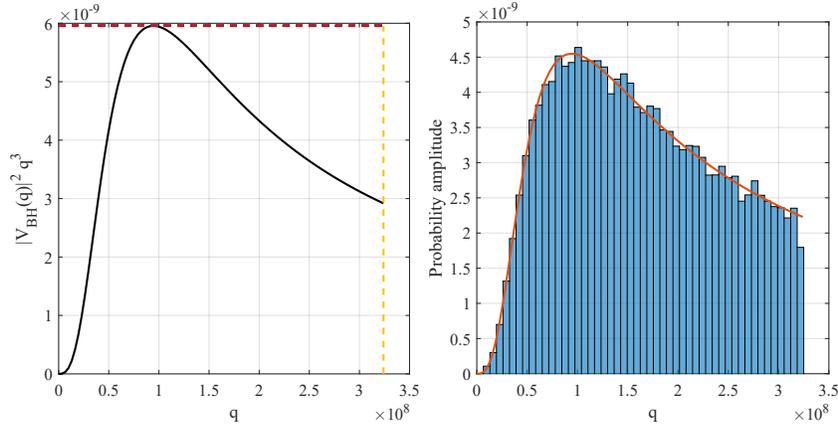
2.1.5 Pauli principle

In a semiconductor of finite size, only a finite number of \mathbf{k} -states are allowed. Considering a chain of N atoms spaced by the quantity a , periodic boundary conditions are imposed on the shape of the electron wavefunction, i.e. the Bloch function:

$$\psi(z) = e^{ikz} u(z) = e^{ik(z+Na)} u(z+Na) = \psi(z+Na). \quad (2.42)$$

This leads to have only discrete values of k

$$k = \frac{2\pi l}{Na} \quad l = 1, \dots, N. \quad (2.43)$$



(a). Theoretical distribution $f(q)$ (equation 2.41) not normalized. The yellow line represents the theoretical distribution dashed line corresponds to $q = 2k$; the red dashed line corresponds to $f(\sqrt{3}\beta_{min})$

(b). Samples distribution: the continuous given by equation 2.41; the histogram represents the distribution obtained with the rejection technique

Figure 2.14: Distribution of wavevector q according to the BH isotropic distribution $f(q)$ (equation 2.41) for $N_i = 1 \times 10^{17} \text{cm}^{-3}$ and $T = 300\text{K}$, $E = 20\text{meV}$

Considering equation 2.43, it can be shown that in three dimensions, the density of states in k-space is [33]

$$\frac{V^3}{4\pi^3}, \quad (2.44)$$

where $V = (Na)^3$ is the sample volume.

The fact that only a finite number of states for each \mathbf{k} exists is generally neglected, but becomes more important when with high carrier density is involved and degenerate behaviour arises in semiconductors. In the BTE 1.1 framework, the main effect of degeneracy is that the term $[1 - f(\mathbf{r}, \mathbf{k}, t)]$ in equation 1.2 is not negligible anymore. A way to consider this term was implemented following the work of [32] for ensemble Monte Carlo simulations. The proposed algorithm is the following:

1. The volume of the simulated device is evaluated as $V = N/n$, where N is the number of simulated particles and n is the carrier density. The carrier density is evaluated using equation 2.30. In this way, the density of states in k-space can be evaluated using 2.44;
2. a grid is set in k-space, choosing the maximum wavevector magnitude k_{\max} and the number of points n_{points} . In such a way, a three-dimensional grid, with the range $[-k_{\max}, k_{\max}], [-k_{\max}, k_{\max}], [-k_{\max}, k_{\max}]$, each dimension having points equal to n_{points} is allocated. The volume of each

k-cell is therefore equal to

$$\Omega_c = \left(\frac{2k_{\max}}{n_{\text{points}} - 1} \right)^3 \quad (2.45)$$

and the maximum number of state that each cell can accommodate is then

$$N_c = \frac{\Omega_c V}{4\pi^3}; \quad (2.46)$$

3. all carrier are tracked in this k-grid, both when they drift and when they scatter, such that a non-normalized occupation probability $f(\mathbf{k})$ can be defined counting the number of particles in each k-cell;
4. when a scattering mechanism occurs, a random number r_1 between 0 and 1 is generated and it is compared to the normalized occupation probability $f(\mathbf{k})/N_c$
 - if $r_1 \leq f(\mathbf{k})/N_c$, the scattering is rejected and the self-scattering is performed;
 - if $r_1 > f(\mathbf{k})/N_c$, the scattering is accepted.

In this way the term $[1 - f(\mathbf{r}, \mathbf{k}, t)]$ is statistically considered.

The algorithm works for any initial distribution of the particle, given that the simulation lasts long enough to reach convergence. A current problem of this implementation is that to avoid conflicts in writing the k-grid in memory, the execution of the simulations using this algorithm is single-thread.

Figure 2.15 shows the number of particles tracked in the k-grid during a Monte Carlo simulation, with and without the rejection algorithm described above: when degeneracy is included, the effect is that particles rearrange occupying higher k-states as well.

Figure 2.16 shows that this rearrangement cause an increase of the average energy of the particles. This was expected since in the parabolic approximation higher values of k correspond to higher values of energy. The effect of the inclusion of degeneracy on the time average of the velocity 1.35 is shown in figure 2.17.

2.2 Full band electronic structure

In this framework, the electronic band structure is numerically evaluated. The energy levels are more realistic than those obtained from the analytic band approximation, and each \mathbf{k} -state is associated with a Bloch wavefunction. This approach allows for a more accurate simulation of the scattering mechanisms.

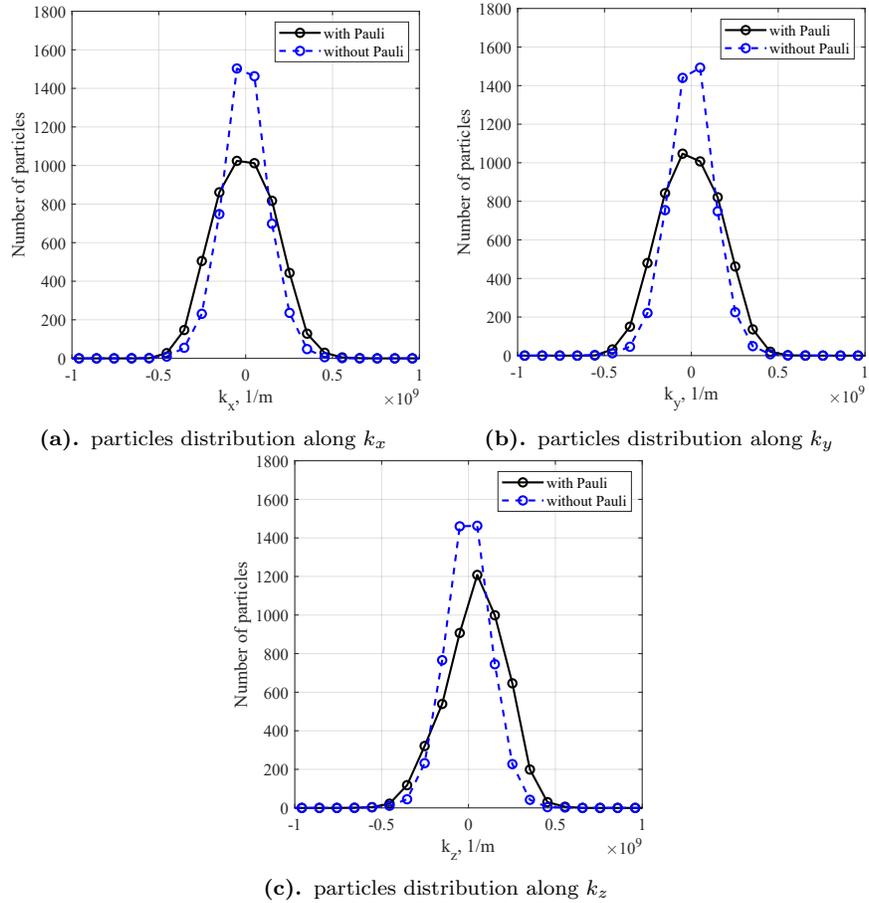


Figure 2.15: Particle distribution in k-grid. 5000 particles simulated, carrier density equal to $1 \times 10^{18} \text{ cm}^{-3}$, $T = 300 \text{ K}$, electric field $F = 1 \times 10^4 \text{ V/m}$ along x direction

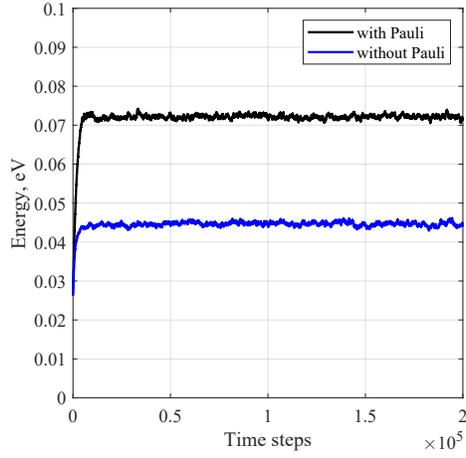


Figure 2.16: Ensemble average energy, 5000 particles simulated, carrier density equal to $1 \times 10^{18} \text{ cm}^{-3}$, $T = 300 \text{ K}$, electric field $F = 1 \times 10^4 \text{ V/m}$ along x direction

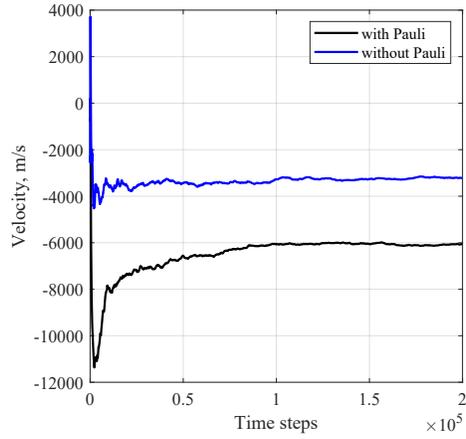


Figure 2.17: Time average velocity. 5000 particles simulated, carrier density equal to $1 \times 10^{18} \text{ cm}^{-3}$, $T = 300 \text{ K}$, electric field $F = 1 \times 10^4 \text{ V/m}$ along x direction

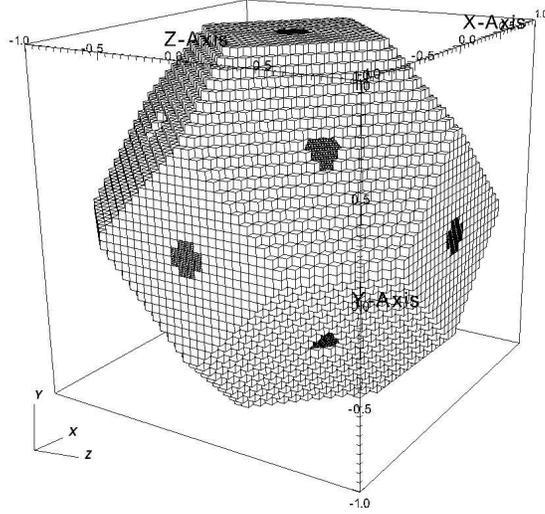


Figure 2.18: Cubic mesh of a zincblende unit cell in reciprocal space

2.2.1 Reciprocal space mesh

In a full-band Monte Carlo simulation, for each material, the mesh of the Full Brillouin zone of the reciprocal lattice is constructed. The nodes of the mesh are the k-vectors to which the data of bandstructure are associated, i.e. the energy levels $E_{n,\mathbf{k}}$ and the wavefunctions $|\psi_{n,\mathbf{k}}\rangle$. The mesh is cubic and the basis functions of this mesh are serendipity elements ([3]). Figure 2.18 shows the shape of the mesh for a zinc-blende material. It is possible to make refinements of the mesh in correspondence of the low-energy points, such as Γ , L or X points.

2.2.2 Empirical pseudopotential method

The next step in the setting of a full-band simulation framework is calculating the energy bands and the wavefunctions, i.e. solving the Schrodinger equation for electrons in the crystal

$$\hat{H}_e |\psi_{n,\mathbf{k}}\rangle = E |\psi_{n,\mathbf{k}}\rangle, \quad (2.47)$$

where \hat{H}_e takes into account the kinetic energy of electrons, the interaction between electrons themselves and the interaction between electrons and ion cores.

$$H_e = \sum_i \frac{p_i^2}{2m_i} + \frac{1}{2} \sum_{ii'} \frac{e^2}{|r_i - r_{i'}|} - \sum_{ij} \frac{Z_j e^2}{|r_i - R_{j0}|}. \quad (2.48)$$

This is in principle a very complex problem since it takes into account the contribution of a huge number particles, and can be only approached with several approximations. One way solving this problem is using *ab initio* calculations such as density functional theory DFT. However, it is highly computationally-expensive. A more computationally-efficient way is to use empirical methods, i.e. based on experimental data or pre-existing *ab initio* information. The one used here is the Empirical pseudopotential method EPM.

Under the mean-field approximation, the many-body Schroedinger equation 2.48 is transformed to a single body equation

$$\left(\frac{\hat{p}^2}{2m} + \hat{V} \right) |\psi_{n,\mathbf{k}}\rangle = E |\psi_{n,\mathbf{k}}\rangle. \quad (2.49)$$

The key point of the EPM is that eigenstates of a valence electron $|\psi_s\rangle$ can be accurately represented by a combination of a finite number (≈ 100) of plane waves, each plane wave identified by a reciprocal lattice vectors \mathbf{G} (decided before the actual calculations)

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} U_{\mathbf{G}}^n(\mathbf{k}) e^{i\mathbf{G}\cdot\mathbf{r}}. \quad (2.50)$$

In order to allow this representation of $|\psi_v\rangle$, the contribution of the core states $|b_t\rangle$ should be removed by the Schrodinger equation . That is because:

- they do not impact significantly the shape of ψ_v outside the relatively small core region;
- their shape is similar to bound states, i.e., concentrated in small regions of space. This means that their reciprocal space representation (i.e. their Fourier expansion) would require a much higher number of plane waves than the one actually needed to reach a satisfying degree of approximation.

Therefore, V in equation 2.49 is replaced by [43]

$$V_p = V + \left(\sum_t (E - E_t) |b_t\rangle \langle b_t| \right). \quad (2.51)$$

Despite the complex appearance of V_p , it can be approximated by using few points (≈ 5) in reciprocal space:

$$V_p(\mathbf{r}) = \sum_{\mathbf{G}} V_f(\mathbf{G}) S(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.52)$$

where $S(\mathbf{G})$ is called *structure factor*

$$S(\mathbf{G}) = (1/N) \sum_{j \text{ in unit cell}} e^{-\mathbf{G}\cdot\mathbf{r}_j} \quad (2.53)$$

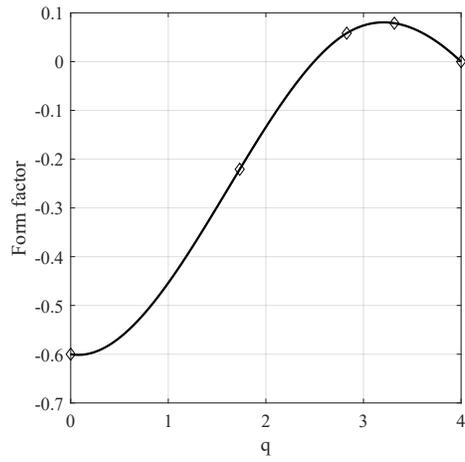


Figure 2.19: Silicon form factors

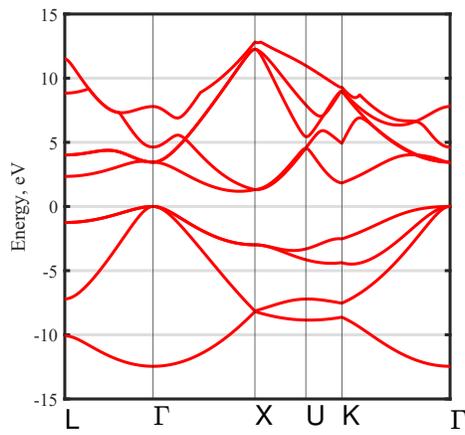


Figure 2.20: Silicon dispersion relation calculated with EPM

and $V_f(\mathbf{G})$ is called *form factor*. \mathbf{r}_j in 2.53 is the position of the j -th atom in the considered unit cell. The profile of the form factors in silicon is shown in figure 2.19.

The Schrodinger problem is then re-written as a matrix eigenvalue equation

$$\mathbf{H}\mathbf{U} = E\mathbf{U}. \quad (2.54)$$

The information of the bandstructure is in the energies E and in the \mathbf{U} coefficients, related to the presiously selected \mathbf{G} vectors (equation 2.50).

2.2.3 Scattering

Having the wavefunctions, the overlap integral (equation 1.11) can be considered exactly in the evaluation of the scattering rate and the selection of the final state. In fact, it is usually neglected or poorly approximated when an analytic description of the bandstructure is considered. This allows an higher degree of accuracy in the carrier interaction dynamics.

$$|\langle \mathbf{k}', n' | \hat{H}' | \mathbf{k}, n \rangle|^2 \propto |I_{\mathbf{k}, \mathbf{k}', n, n'; \mathbf{G}}|^2, \quad (2.55)$$

$$I_{\mathbf{k}, \mathbf{k}', n, n'; \mathbf{G}} = \frac{1}{\Omega_c} \int_{\Omega_c} d\mathbf{r} \sum_{\mathbf{G}' \mathbf{G}''} U_{\mathbf{G}'' \mathbf{k}'}^{(n')*} U_{\mathbf{G}' \mathbf{k}}^{(n')} e^{i(\mathbf{G} + \mathbf{G}' - \mathbf{G}'') \cdot \mathbf{r}}, \quad (2.56)$$

where \mathbf{G} is the momentum exchanged during the interaction.

Considering the overlap factor, the electron-phonon scattering rates in silicon are evaluated. The considered bandstructure is made of 4 conduction bands and 4 valence bands (figure 2.20). The phonon dispersion relation is approximated as figure 1.1b, and the deformation potentials are tuned in order to mach *ab initio* calculated scattering rates [1]; they are shown in figure 2.21.

The electron-phonon scattering rate in silicon from the first conduction band to the upper ones is reported in figure 2.22 for acoustic and optical phonons. The hole-phonon scattering rate from the first valence band to the lower ones is reported in figure 2.23 for acoustic and optical phonon. This rate is evaluated as function of \mathbf{k} -vector as in equations 1.8 and 1.9 and then mapped in function of energy (multiple points are present for the same energy because multiple \mathbf{k} -vectors correspond to the same energy in the dispersion relation).

The total acoustic phonon scattering rate, i.e. the transition rate at each energy considering the maximum scattering rate from each bands, is reported in figure 2.24 for electrons and holes. Likewise, the total optical scattering rate is reported in figure 2.25. As it can be observed, they are in good agreement with *ab initio* calculations (HSE rate in the figures).

Finally, as an example of the effectiveness of the full-band model, intrinsic silicon velocity-field curve are obtained and compared to experimental data. The included scattering mechanisms are

- acoustic phonon deformation potential scattering;
- non polar optical phonon deformation potential scattering;

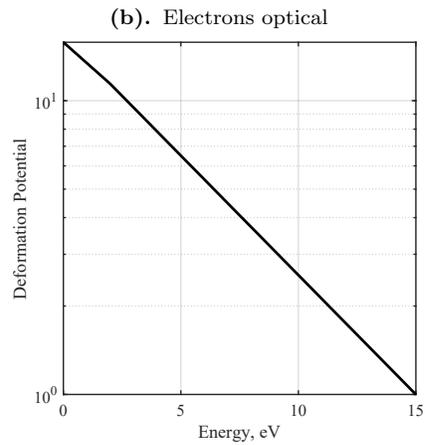
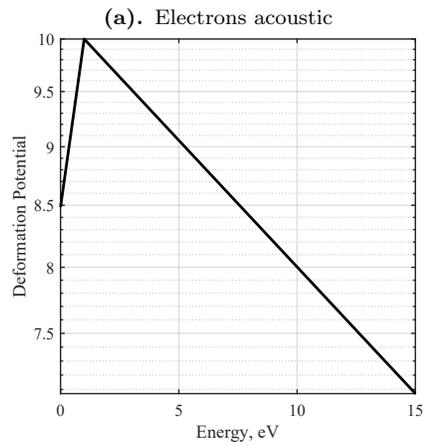
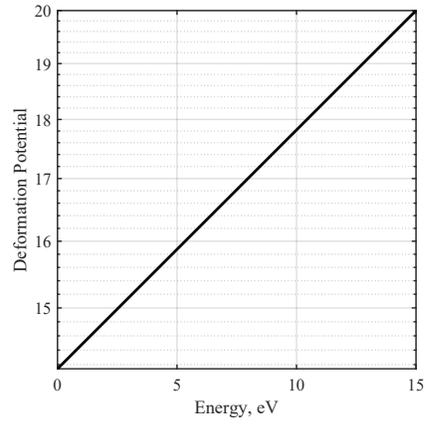
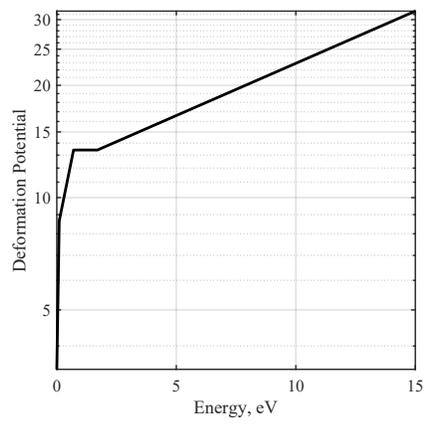


Figure 2.21: Silicon full-band carrier-phonon deformation potential

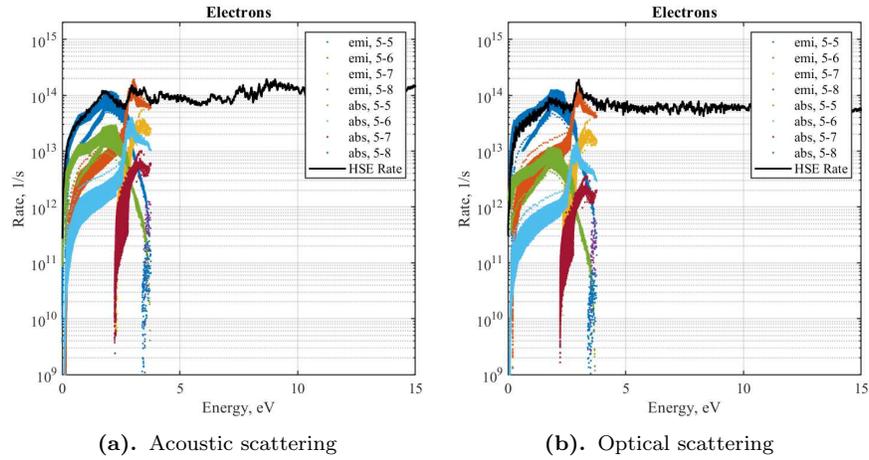


Figure 2.22: Electron-phonon full band scattering rates in silicon, from the lowest conduction band to the upper ones

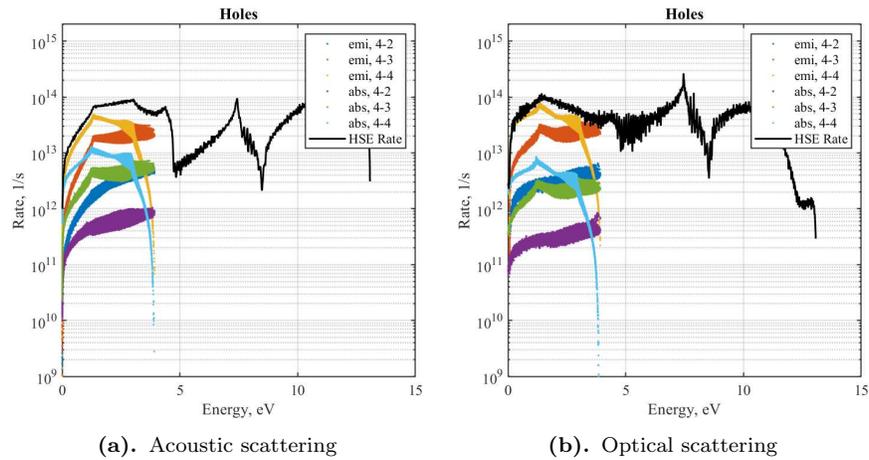


Figure 2.23: Hole-phonon full band scattering rates in silicon, from the highest valence band to the lower ones

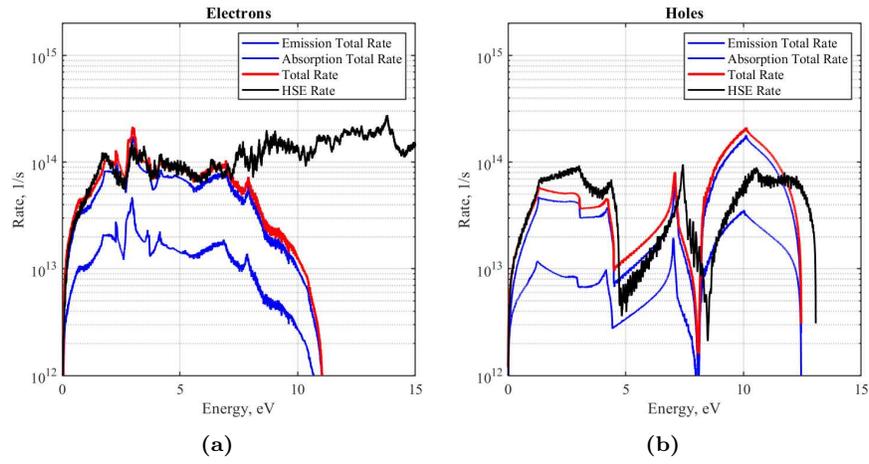
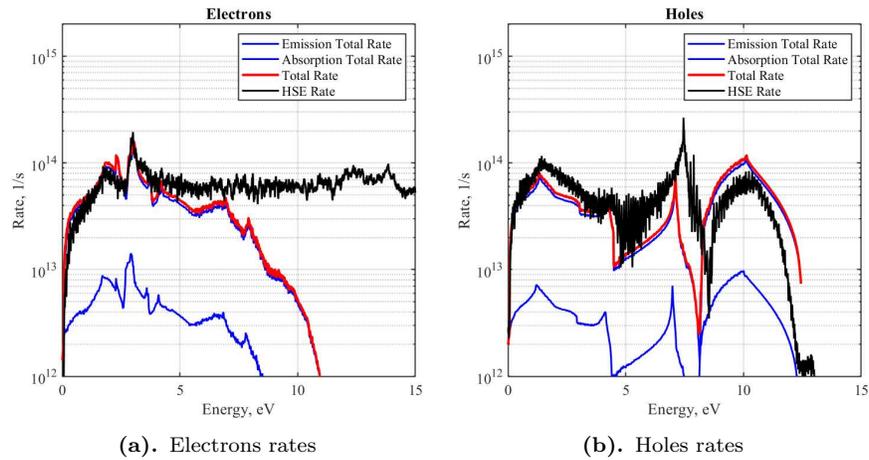


Figure 2.24: Silicon full band total acoustic scattering rates



(a). Electrons rates

(b). Holes rates

Figure 2.25: Silicon full band total optical scattering rates

Carrier	P_0 ($\text{eV}^{-\alpha_0} \text{s}^{-1}$)	α_0	E_0 (eV)
e	1.0×10^{11}	4.6	1.1
h	1.14×10^{12}	3.4	1.49

Table 2.2: Values for impact ionization rate

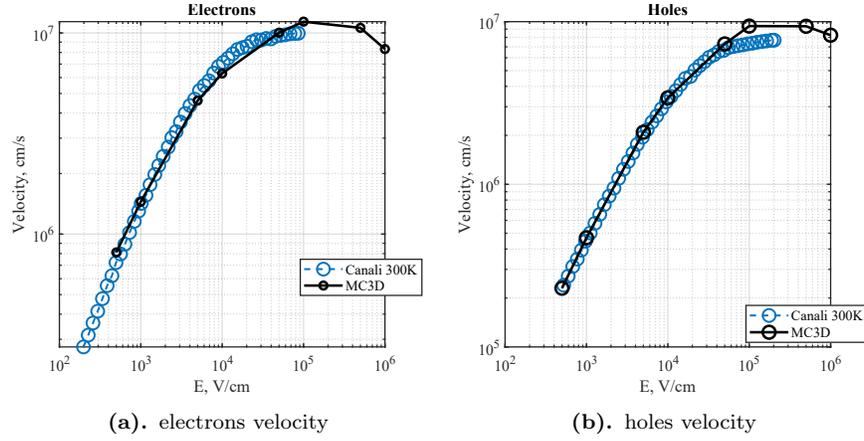


Figure 2.26: Silicon velocity-field curves, $T = 300$ K, experimental data taken from [23]

- impact ionization, approximated by the Keldysh rate [36]

$$P_{11}(E) = P_0 (E - E_0)^{\alpha_0} \quad (2.57)$$

where the values of P_0 , E_0 and α_0 are reported in table 2.2

As it can be seen, simulation are quite in agreement even at very high electric field.

Chapter 3

Mobility of InAs, InSb and InAsSb

3.1 Simulation context

Indium arsenide (InAs) and indium antimonide (InSb) are direct bandgap semiconductors with a small bandgap and high electron mobility. These properties make them attractive materials for use in Mid-Wave InfraRed (MWIR) photodetectors, both in bulk form and in superlattices [40], as well as for channel layers in high-electron-mobility transistors (HEMTs) [4]. Their alloy, InAsSb, has found significant applications in superlattice detectors and barrier detectors [28].

Despite their practical importance, there is limited understanding of the transport properties of these materials, both in experimental and theoretical investigations. This chapter provides an estimation of their low-field mobility as a function of temperature and doping.

3.2 Simulation framework

For the evaluation of the mobility of InAsSb, the mobility of the compounds InAs and InSb is evaluated first. A matching with experimental data is performed for these two materials and then the used simulation parameters were interpolated to create a model for the ternary alloy.

Since mobility μ is a low-field quantity, it is assumed that electrons stay in the bottom of the conduction band and don't scatter in upper valleys of bands. So the band-structure used to describe k-states is composed of a single spherical non-parabolic conduction band, described by the effective mass m^* and a non-parabolicity factor α evaluated according to the model

$$\alpha = \frac{1}{E_g} (1 - m^*)^2, \quad (3.1)$$

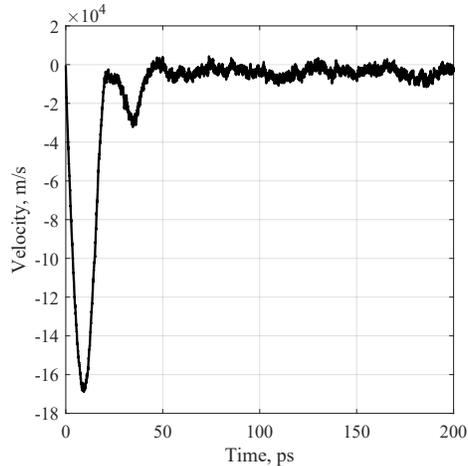


Figure 3.1: Ensemble velocity in the direction of the field in a simulation including Pauli exclusion principle

where E_g is the material energy gap.

The mobility is extracted from the following formula

$$\mu = \frac{|v|}{F}, \quad (3.2)$$

where v is the velocity and F is the electric field.

The simulation are performed using 5000 particles and an electric field of $F = 1 \times 10^4 \text{ V/m}$. Different electric fields have been tried, but the one chosen is small enough to belong to the linear region of the velocity-field curve but sufficiently high to cancel out the effect of thermal fluctuation. In the simulations that not include Pauli exclusion principle, the chosen simulation time was of 100 ps and the time step was $5 \times 10^{-16} \text{ s}$. The velocity is obtained from the temporal average (equation 1.35).

For the simulation that included Pauli blocking, since the initial Boltzmann distribution has to rearrange, the time needed to reach convergence is higher. Therefore, the simulation time was set to 200 ps and the time step was set to $1 \times 10^{-15} \text{ s}$. The number of simulated particles was increased to 10000 and v is extracted by performing the mean on the ensemble average of the velocity (equation 1.36) considering only the second half of the simulated points ($t \in [100 \div 200] \text{ ps}$). This because there is significant time where the particle distribution rearranges and the temporal-averaged velocity poorly converges within the simulation time used.

An example of ensemble velocity obtained from a simulation including Pauli principle is shown in figure 3.1.

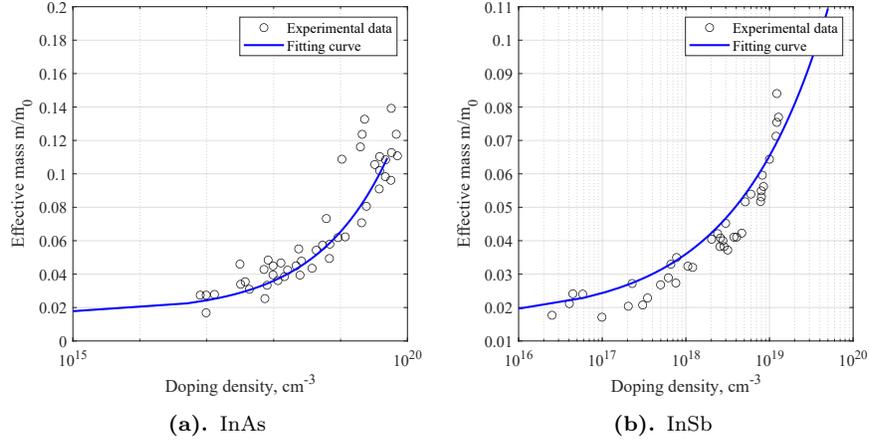


Figure 3.2: Effective mass vs doping level

	InAs	InSb
E_g^0	0.415 eV	0.24 eV
α	$2.76 \times 10^{-4} \text{ eV K}^{-2}$	$6.0 \times 10^{-4} \text{ eV K}^{-2}$
T_0	83 K	500 K

Table 3.1: Values for the evaluation of E_g with the Varshini model, for InAs and InSb

3.3 InAs and InSb

For the binary compounds InAs and InSb, the dependence of the effective mass from the electron concentration is obtained fitting experimental data. The model is the following:

$$m_e^* = aN_d^b + c, \quad (3.3)$$

where $a = 1.225 \times 10^{-9} \text{ cm}^3$, $b = 0.4$ and $c = 0.0166$. The comparison of the model with experimental data, taken from [44], is shown in figure 3.2.

The temperature dependence of the energy gap is obtained using the Varshini equation

$$E_g = E_g^0 - \frac{\alpha T^2}{T + T_0}. \quad (3.4)$$

The values E_g^0 , α and T_0 for InAs and InSb are reported in table 3.1 and the value of E_g is shown in figure 3.3.

The following scattering mechanisms are included in the simulations:

- acoustic phonon deformation potential scattering
- non-polar optical deformation potential phonon scattering
- polar optical phonon scattering

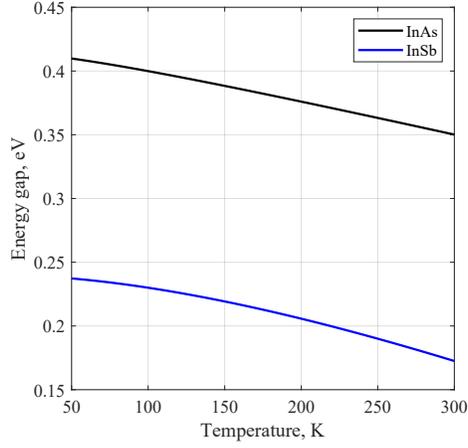


Figure 3.3: Energy gap vs temperature, Varshini model

	InAs	InSb
Lattice constant a , Å	6.058	6.479
Density ρ , g/cm ³	5.68	5.66
Sound velocity v_L , cm/s	3.83×10^5	3.4×10^5
Dielectric constant ($f \rightarrow 0$) ϵ_0	15.15	16.9
Dielectric constant ($f \rightarrow \infty$) ϵ_∞	12.3	15.7
Acoustic ph. def. pot. D_{AC}^e , eV	10.0	7.0
Optical ph. def. pot. D_{opt}^e , eV/cm	1×10^9	1×10^9
Optical phonon energy $\hbar\omega_{LO}$, eV	0.03	0.025

Table 3.2: Parameters necessary for the inclusion of the scattering mechanisms in the simulation

- ionized impurity scattering, using Brooks-Herrings, Ridley and Kosina models

For the implementation of phonon scattering, see [24]. All the parameters required for the implementation of the scatterings are reported in table 3.2

3.3.1 Results

Temperature dependence Figure 3.4 reports the simulation results of InAs and InSb electron mobility in function of temperature for doping density $N_i = 1 \times 10^{15} \text{ cm}^{-3}$. A good matching with experimental data can be observed in the trend and in the order of magnitude by all analyzed models. The BH and Ridley model gives the same result: the mobility is evaluated assuming the case of majority electrons in uncompensated material so the screening is high and the effect of the "third body exclusion" is negligible. The Kosina model gives a slightly lower value: this can be addressed to the two-ions scattering correction,

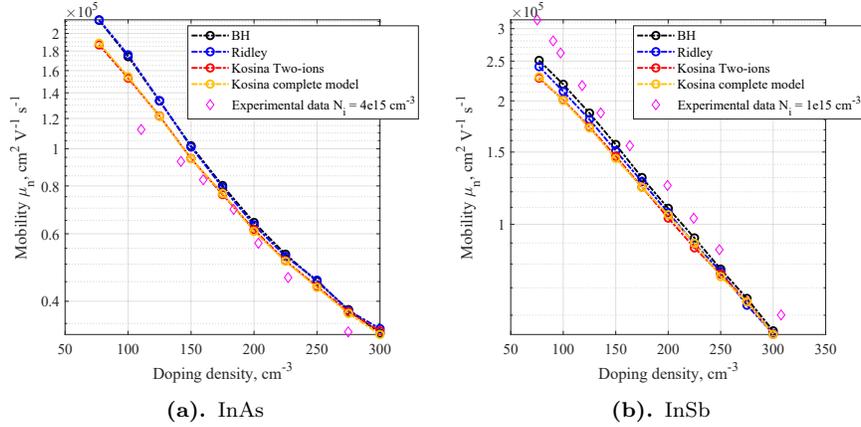


Figure 3.4: Mobility vs temperature, $N_i = 1 \times 10^{15} \text{ cm}^{-3}$, data taken from [10] for 3.4a and from [21] for 3.4b

that is appreciable even at this doping density.

Doping dependence In the following paragraph the results of simulations in function of doping are shown; experimental data are taken from [25] for InAs and from [31] for InSb.

Figure 3.5 reports the simulation results of InAs and InSb in function of doping density at $T = 300 \text{ K}$. Also here, the BH and Ridley model gives quite similar results, while Kosina model gives a slightly lower value for the mobility at $N_i = 1 \times 10^{18} \text{ cm}^{-3}$. The results are quite in agreement with experimental data for InAs; instead, the mobility seems underestimated for InSb. This may be due to the phonon scattering, whose deformation potential are not perfectly adjusted. At $T = 77 \text{ K}$ (figure 3.6), the mobility is the same for BH and Ridley model and in agreement with experimental data, but the Kosina model gives values that are almost two orders of magnitude lower. To explain this, a detailed analysis of all the corrections has been performed, and the results are reported in figures 3.7 for $T = 300 \text{ K}$ and in 3.8 for $T = 77 \text{ K}$.

First, it is possible to observe that the equivalent BH isotropic scattering, used for the implementation of Kosina model, gives the same results compared to the BH non-isotropic model. The second Born correction and the two-ion correction gives slightly lower values of the mobility conserving the same trend of the one obtained with BH. The q -dependent and degeneracy-dependent screening function provides a further decrease of the mobility at $T = 300 \text{ K}$. At $T = 77 \text{ K}$ the screening function contribution is the most relevant. In particular, at lower temperatures and high doping level the degeneracy is too high. In this conditions, the rational approximation $G(q, \eta)$ deteriorates with respect to the actual one 2.19, but the distribution of the final state is not so different from the BH (figure 2.11a shows a lower average scattering angle for the Kosina model compared to the BH model). Instead the term $\mathcal{F}(-1/2, \eta)/\mathcal{F}(1/2, \eta)$ that mul-

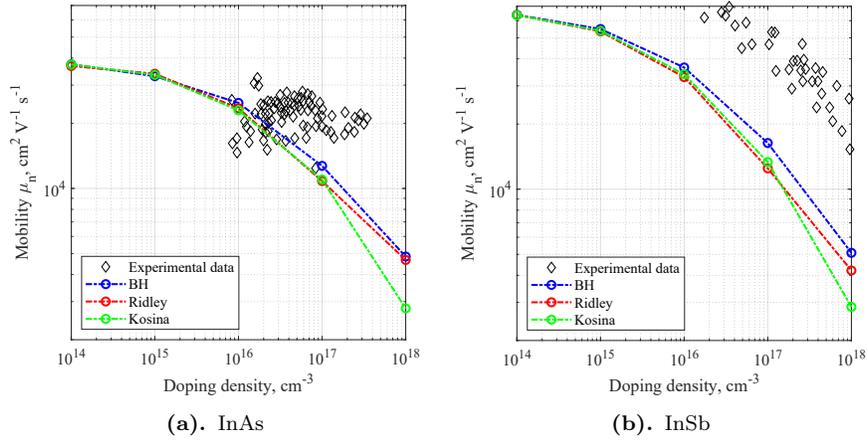


Figure 3.5: Mobility vs doping density, $T = 300$ K. Comparison between BH, Ridley and Kosina models

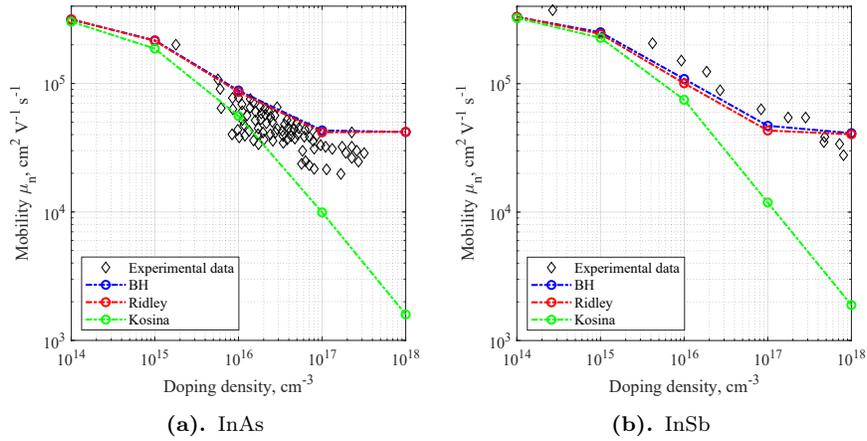


Figure 3.6: Mobility vs doping density, $T = 77$ K. Comparison between BH, Ridley and Kosina models

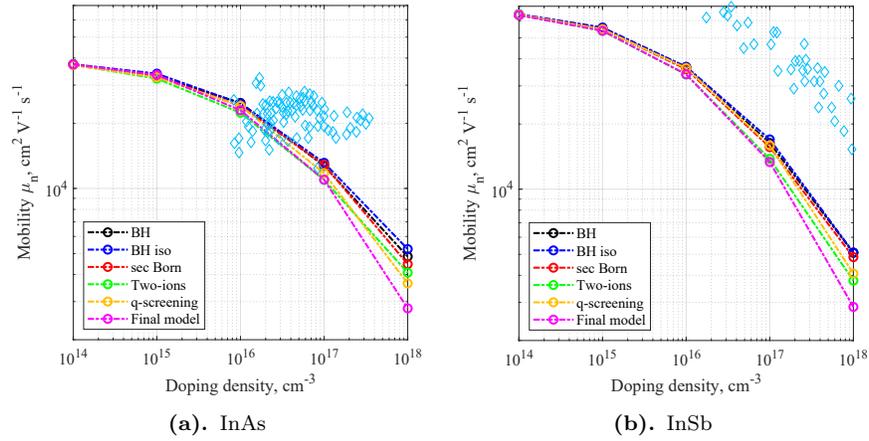


Figure 3.7: Mobility vs doping density, $T = 300 \text{ K}$. Analysis of the various correction of the BH model introduced by Kosina model

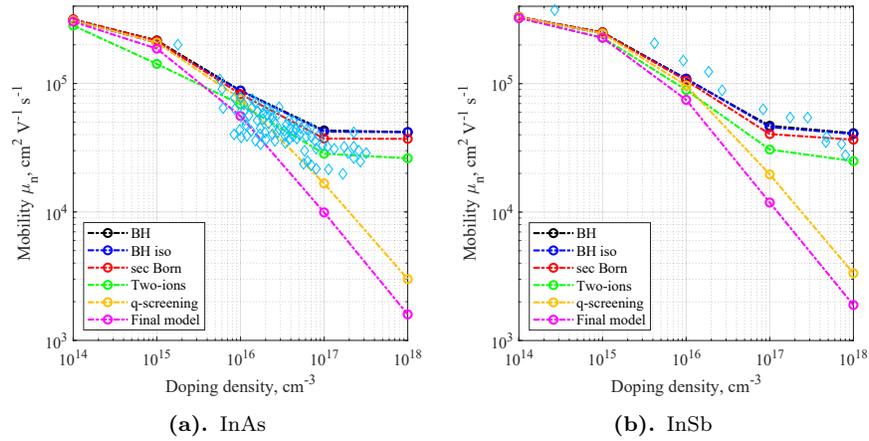


Figure 3.8: Mobility vs doping density, $T = 77 \text{ K}$. Analysis of the various correction of the BH model introduced by Kosina model

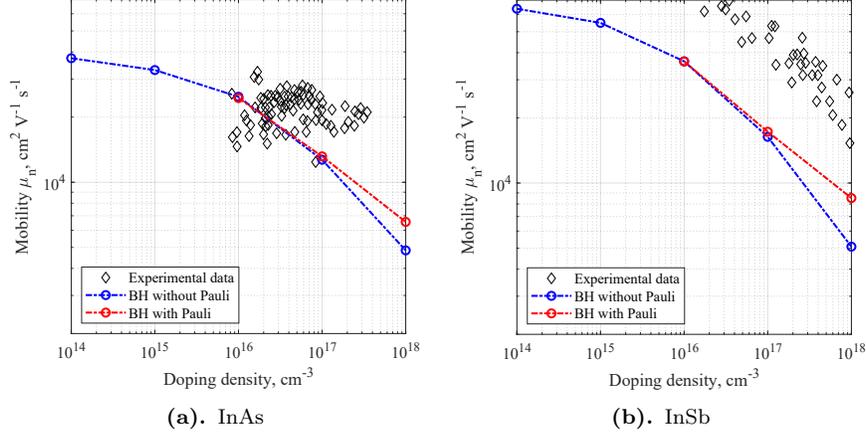


Figure 3.9: Mobility vs doping density, $T = 300$ K. Analysis of the effect of Pauli exclusion algorithm on BH model

multiplies the q_0^2 2.17 can reach significantly small number. With a lower screening, the scattering rate increases and for higher doping density, the probability of having a final state with large θ increases. This results in a huge decrease in electron velocity and so, mobility. In tables 3.4 the calculated values of η for InAs and InSb at $T = 300$ K and $T = 77$ K are reported. Since this correction to β 2.17 is physical and not related to any assumption the Kosina model could have made (e.g., the two-ions scattering and the rational approximation of the screening function) it is supposed that using a more accurate model (i.e. including electron-plasmon scattering, relevant at high degeneracy level and low temperatures), a more realistic behaviour of the mobility could be retrieved.

Doping	77 K	300 K	Doping	77 K	300 K
1e14	-4.33	-3.96	1e14	-4.37	-0.68
1e15	-2.05	-3.94	1e15	-2.98	-0.72
1e16	0.41	-2.14	1e16	0.34	-0.82
1e17	4.30	0.05	1e17	4.06	-0.11
1e18	12.6	2.72	1e18	11.5	2.16
1e19	27.8	6.65	1e19	24.2	5.34

Table 3.3: InAs

Table 3.4: InSb

Pauli principle The Pauli blocking algorithm is then used in the simulations. The k -grid is chosen such that for the maximum number of states per cell was around 40 for every simulations. The mobility in function of doping density using the BH model is reported in figure 3.9 for $T = 300$ K and in figure 3.10 for $T = 77$ K.

The Pauli principle is responsible of rearranging particle's states to higher k -vectors. This results in higher energy and increase of velocity and mobility. This behaviour can be observed both at $T = 300$ K and $T = 77$ K. A change in

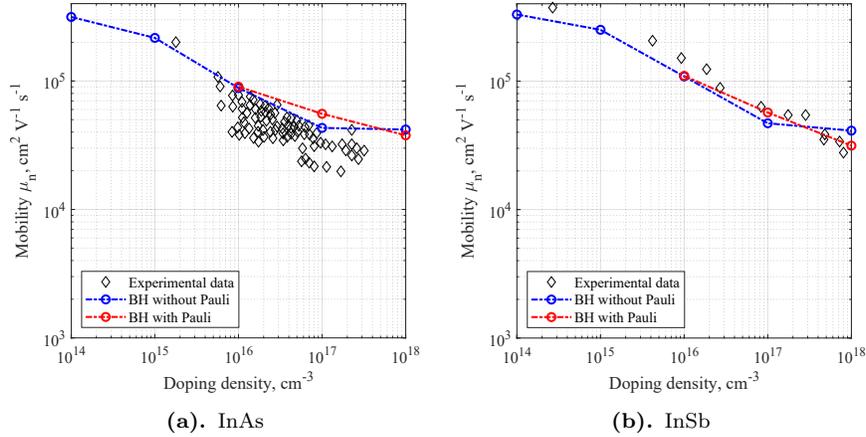


Figure 3.10: Mobility vs doping density, $T = 77$ K. Analysis of the effect of Pauli exclusion algorithm on BH model

mobility is appreciable for sufficiently high doping density: $N_i > 1 \times 10^{16} \text{ cm}^{-3}$ at $T = 77$ K and $N_i > 1 \times 10^{17} \text{ cm}^{-3}$ at $T = 300$ K. At $N_i = 1 \times 10^{18} \text{ cm}^{-3}$ and $T = 77$ K the mobility is lower using Pauli blocking, in accordance with the trend of experimental data.

Figures 3.11 show the mobility obtained using Kosina model and Pauli principle. It is evident that the equivalent BH isotropic model does not give the same results of the non-isotropic BH model. This may be due to the fact that the former does not entirely come from physics laws, but the cross section of the interaction is constructed "artificially" in such a way the scattering could be isotropic. When such model is modified by another "physical" model, such as the Pauli blocking, the equivalence does not hold anymore and unrealistic results are obtained, e.g. mobility that increases with doping as in figure 3.11a. In order to have a correct analysis of the Kosina model with the addition of Pauli blocking, the former model has to be evaluated using "physics laws", i.e. using the non-isotropic cross section (section 2.1.4). The mobility results using this model are reported in figure 3.12 and 3.13. At $T = 77$ K, it is possible to see that the mobility has increased by more than one order of magnitude, giving realistic results for InAs and InSb up to $N_i = 1 \times 10^{17} \text{ cm}^{-3}$.

As for the Ridley model, no simulations have been performed with the addition of the Pauli blocking, because it has a statistical correction not derived from "physics laws" as Kosina equivalent isotropic BH rate, so the application of Pauli blocking is, in principle, nonphysical.

To end this analysis, a comparison between BH and Kosina models is shown in figure 3.14 and 3.15.

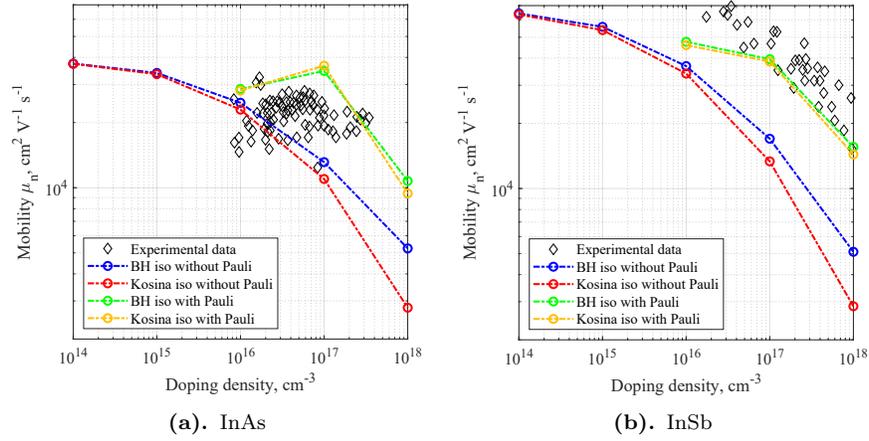


Figure 3.11: Mobility vs doping density, $T = 300$ K. Analysis of the effect of Pauli exclusion algorithm on Kosina model

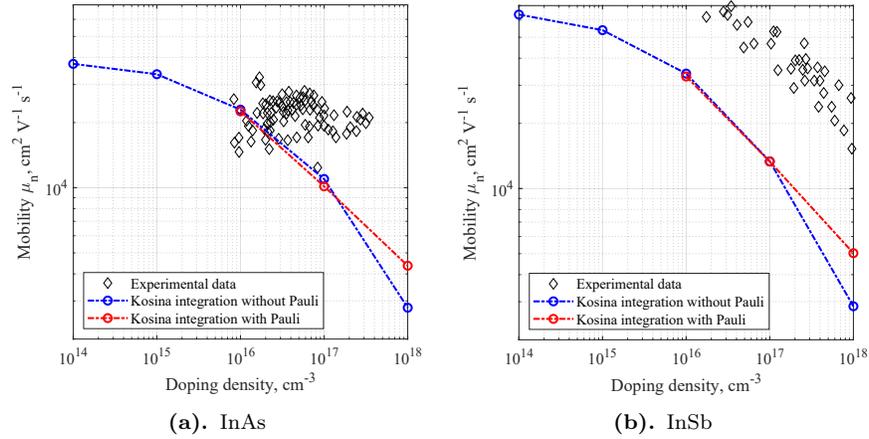


Figure 3.12: Mobility vs doping density, $T = 300$ K. Analysis of the effect of Pauli exclusion algorithm on Kosina model

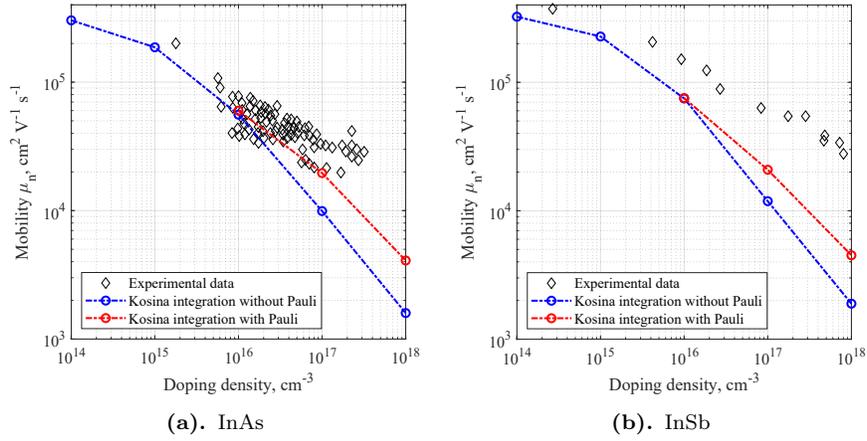


Figure 3.13: Mobility vs doping density, $T = 77$ K. Analysis of the effect of Pauli exclusion algorithm on Kosina model

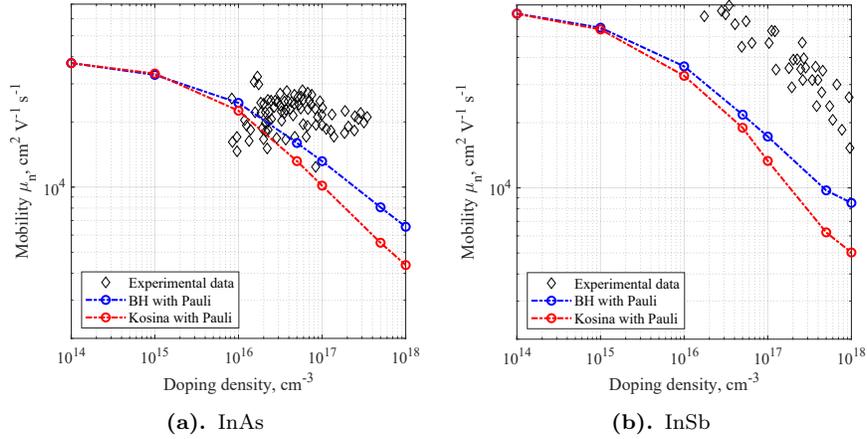


Figure 3.14: Mobility vs doping density, $T = 300$ K. Comparison between BH and Kosina models considering Pauli principle

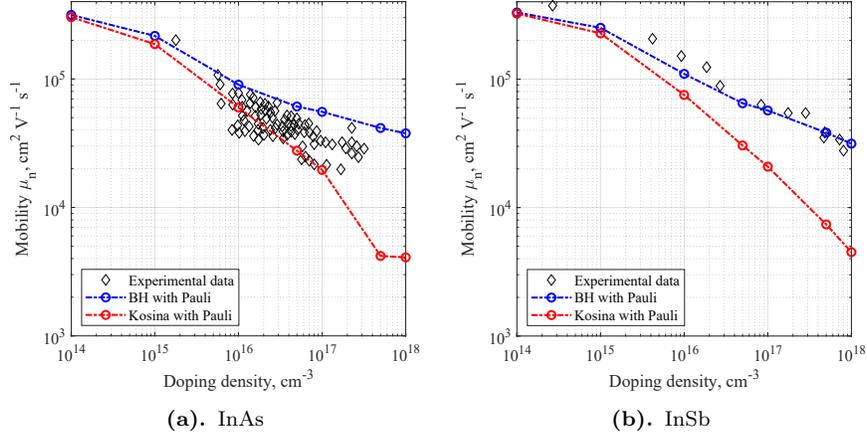


Figure 3.15: Mobility vs doping density, $T = 77$ K. Comparison between BH and Kosina models considering Pauli principle

3.4 InAsSb

As for InAsSb material properties, a search in literature was performed and the dependence of the effective mass from the molar fraction and the dependence of the gap from molar fraction and temperature was found. All the other parameters are interpolated linearly between the ones of InAs and of InSb according to the following formula:

$$Q_{\text{InAsSb}} = xQ_{\text{InSb}} + (1-x)Q_{\text{InAs}}, \quad (3.5)$$

where Q is a generic quantity (e.g. relative dielectric permittivity, density, optical phonon energy etc.) and x is the Sb molar fraction. The data for the effective mass were taken from [39] and the fitting equation is

$$m_e = 0.023 - 0.039x - 0.03x^2. \quad (3.6)$$

The effective mass in function of the Sb molar fraction is shown in figure 3.16.

The data for the energy gap were taken from [42] and the model is

$$E_g(x, T) = E_g^{\text{InSb}}(T)x + E_g^{\text{InAs}}(T)(1-x) - x(1-x)b_T, \quad (3.7)$$

where b_T is the bowing parameter

$$b_T = b_0 - \frac{k_B S_1 T_1}{e^{T_1/T} - 1} + \frac{k_B S_2 T_2}{e^{T_2/T} - 1}. \quad (3.8)$$

The values for the bowing parameters are reported in table 3.5 and the energy gap in function of the molar fraction for $T = 77$ K and $T = 300$ K is shown in figure 3.17.

As for the scattering rates, along with the scattering used for the binary compounds, the **alloy scattering** is added, that represents the fluctuation of

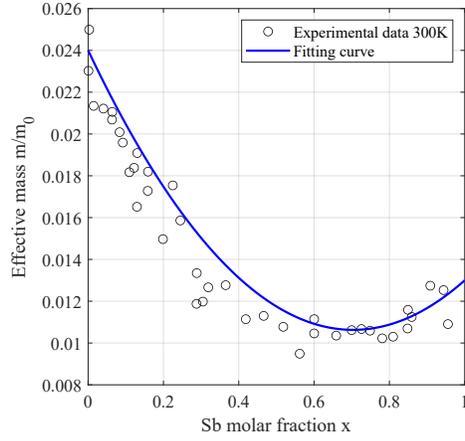


Figure 3.16: Effective mass of InAsSb in function of the Sb molar fraction, $T = 300$ K

b_0	S_1	T_1	S_2	T_2
0.938 eV	22.98	115.8 K	18.8	275.6 K

Table 3.5: Parameters for the evaluation of the bowing parameter for InAsSb, [42]

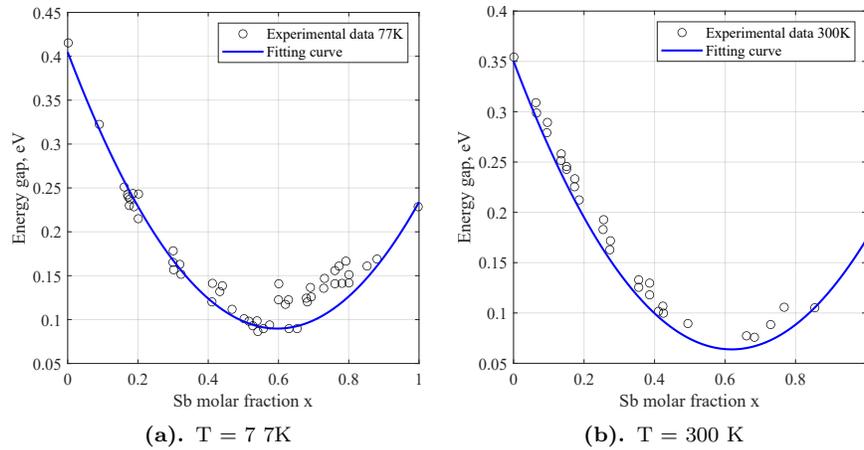


Figure 3.17: InAsSb bandgap E_g

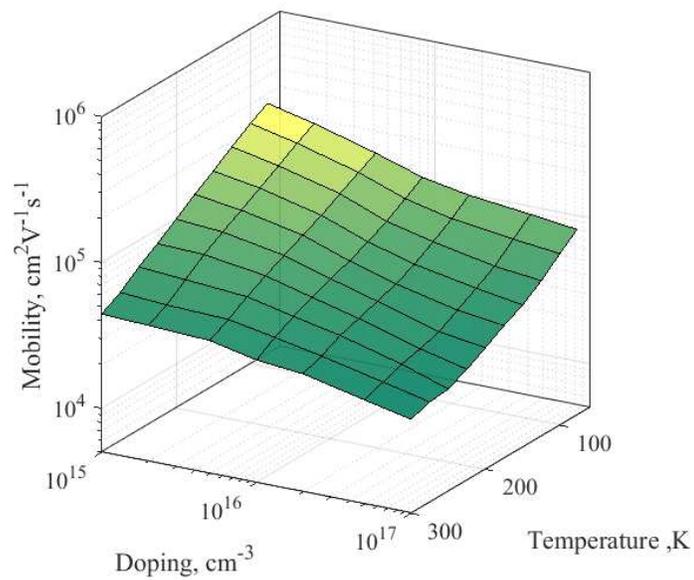
the local crystal potential in space [30]. It is modeled as an isotropic scattering characterized by a parameter ΔU . In the following simulations, it is equal to the difference in the electron affinity of the two materials $\Delta U = \chi_{InAs} - \chi_{InSb} = 0.31$ eV.

3.4.1 Results

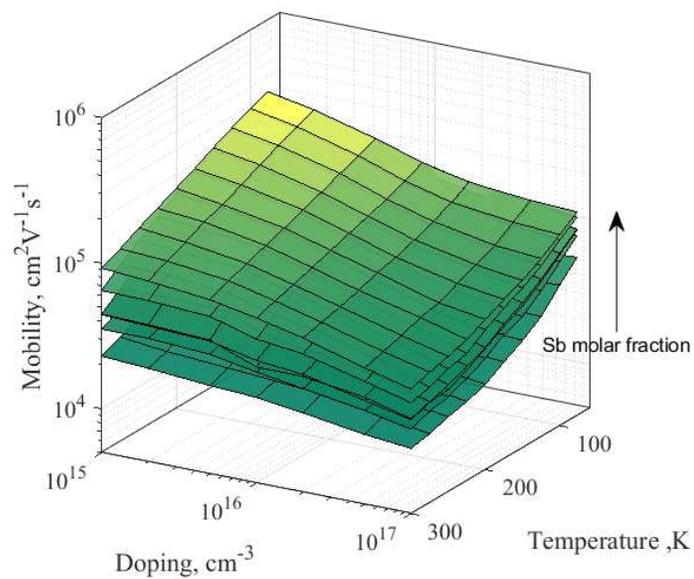
From the analysis of the ionized impurity scattering models performed for the materials InAs and InSb, it was observed that Ridley model gives the same mobility values of BH model and Kosina model gives a significantly lower mobility especially at $T = 77$ K. In order to comply with experimental data in the temperature range $T \in (77, 300)$ K, the mobility of InAsSb was calculated using BH model with the Pauli exclusion principle. However in this condition, an increase of mobility with doping for $N_i > 1 \times 10^{17} \text{ cm}^{-3}$ was observed, that is clearly unrealistic. This is due to the model of the effective mass, that does not take into account of any variation with the electron concentration. For this reason the simulation field is reduced to the interval $N_i = (1 \times 10^{15}, 1 \times 10^{17}) \text{ cm}^{-3}$, where the effective mass is not supposed to change much, as was observed in InAs and InSb. The overview of the the results of the simulations is shown in figure 3.18.

In the following figures, some cut plots of 3.18 are presented, to have a better understanding of mobility in the various combination of temperature, doping and molar fraction.

It is worthwhile commenting figure 3.21: at $T = 300$ K, the mobility increases with increasing Sb molar fraction except between $x = 0.4$ and $x = 0.6$ where it does not seem to change. This behaviour could be explained by the fact that the InAsSb effective mass is not linear with the Sb composition, but has a minimum around $x = 0.6$. In this region the linear interpolation of some quantities, such as the phonon deformation potentials or the phonon longitudinal sound velocity could not be so accurate. For example, other phonon modes, called **alloy modes** [12], should be taken into account.

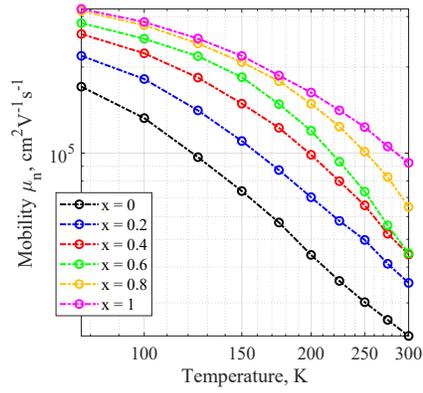


(a). Sb molar fraction $x = 0.3$

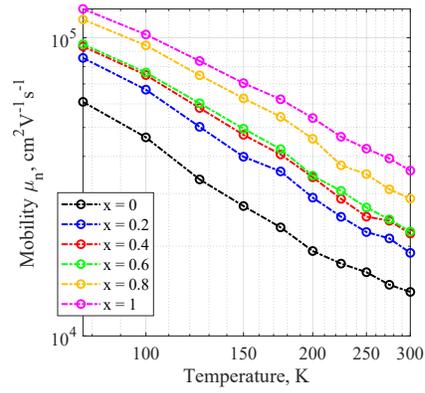


(b). Surfaces correspond to different values of Sb molar fraction

Figure 3.18: InAsSb electron mobility in function of temperature, doping density and molar fraction

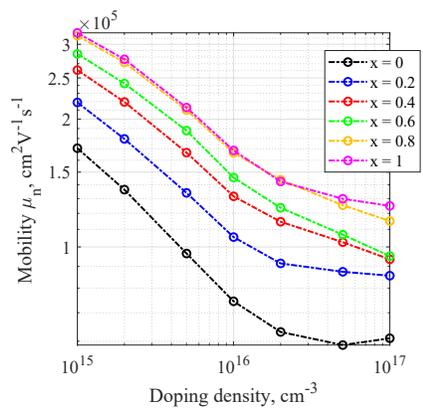


(a). $N_i = 1 \times 10^{15} \text{ cm}^{-3}$

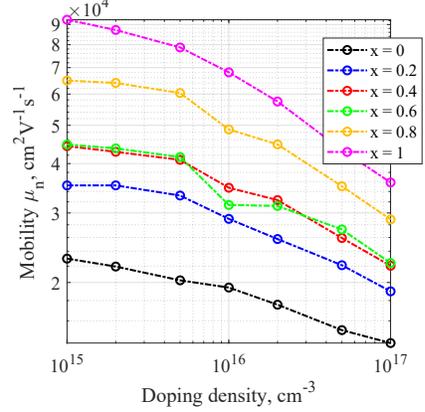


(b). $N_i = 1 \times 10^{17} \text{ cm}^{-3}$

Figure 3.19: InAsSb electron mobility in function of temperature

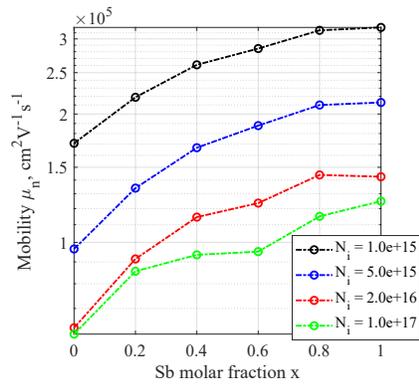


(a). $T = 77 \text{ K}$

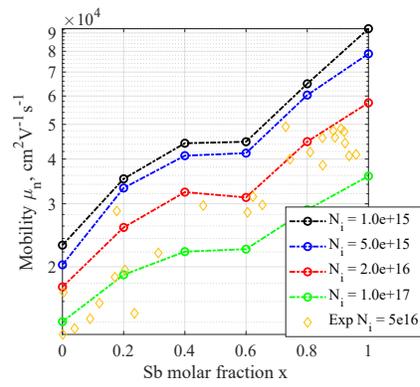


(b). $T = 300 \text{ K}$

Figure 3.20: InAsSb electron mobility in function of doping



(a). $T = 77\text{ K}$, doping density in cm^{-3}



(b). $T = 300\text{ K}$, doping density in cm^{-3}

Figure 3.21: InAsSb electron mobility in function of molar fraction

Chapter 4

Ge-on-Si photodetector frequency response

4.1 Simulation context

Photodetectors are devices capable of converting optical power in an electrical signal. The most important figure of merit of a photodetector is the **responsivity** R , defined as the ratio between the photogenerated current I_{PD} and the input optical power P_{op} . It depends mainly on the input optical wavelength, since it is related to the capacity of the absorbing layer to convert photons in electron/hole pairs.

When a non-DC optical power is provided as input, R depends also on its modulation frequency. Supposing that the photodetector is a linear time-invariant LTI system, input and output are related by a complex transfer function

$$\hat{I}_{PD}(f) = R(f)\hat{P}_{op}(f). \quad (4.1)$$

The complex responsivity $R(f)$ has typically a low-pass behaviour and its bandwidth is called **electrical bandwidth**. The reasons behind this frequency limitations are due to two main reasons:

- extrinsic mechanism, i.e. external parasitic resistance and capacitance, that forms low-pass RC filters with the input impedance of the photodetector;
- intrinsic mechanisms, related to the device capability of driving generated carriers towards the contacts and they are different according to the specific type of photodetector.

The device analyzed in this work is shown in figure 4.1. It is a waveguide p-i-n detector, made by an intrinsic germanium absorbing region (**1.** in figure), sandwiched between two silicon regions n - and p -doped (**2.** and **3.** respectively). The two Si regions are connected by a Si bridge grown on top of germanium. The

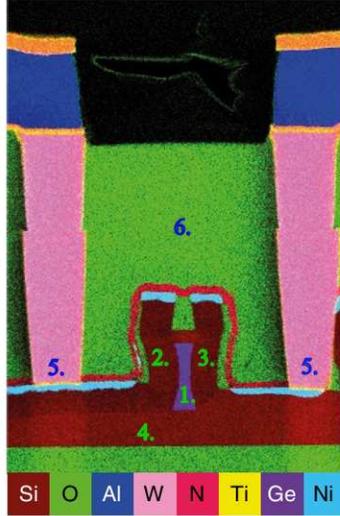


Figure 4.1: Ge-on-Si photodetector realized by Lischke et al. [13]

input optical signal is provided to the device through an underlying Si waveguide (4.) and the output signal is collected by the contacts made of tungsten (5.). The whole device is covered by a SiO₂ layer that is used as passivation and isolation (6.).

The device works in **reverse bias**, i.e., a negative voltage is applied to the contact of the p-side if the contact at the n-side is connected to ground. A high electric field is created in correspondence of the absorber layer, where e/h pairs are generated, and low current passes through the device in case of no illumination (called **dark current**). Under an applied bias $V = -2V$, for wavelength $\lambda = 1550$ nm, the study [13] shows that this device can reach an electrical bandwidth of 265 GHz.

In this type of devices, the most important intrinsic mechanisms to the limitation of the electrical bandwidth is the transit time, i.e. the time spent by generated carriers to reach the contacts.

$$\tau_{tr} = \frac{W}{v}, \quad (4.2)$$

where v is an average effective speed of electrons and holes and W is the absorption region width. Equation 4.2 is a poor approximation of the transit time, since

- the geometry of the device is complex and its contribution cannot be related only on the parameter W ;
- velocity v it is a difficult quantity to estimate since it depends on the type

of particles (electrons or holes), on the electric field they are subjected to and on the geometry of the device.

The first problem is solved considering the field in each point of the device, i.e. by a 3D simulation; an accurate estimation of the velocity of carriers is obtained using a full-band description of the materials.

4.2 Simulation framework

Following the theory of LTI systems, the frequency response is evaluated starting from the impulse response. An impulse of optical power $\delta(t)$ is provided in input and the output current is the impulse response of the system. Performing the Fourier transform \mathcal{F} on $I_{\text{out}}(t)$ one obtains the frequency response:

$$I_{\text{out}}(t) = R(t) * \delta(t), \quad (4.3)$$

$$|\mathcal{F}(I_{\text{out}})| = |R(f)|. \quad (4.4)$$

The simulation of the impulse response is performed as follows:

- first, the electric field under reverse bias $V = -2$ V is evaluated;
- then, to simulate the impulse of optical power, carriers are injected in one single instant and move according to the electric field previously calculated. This is called **frozen field** simulation and is based on the assumption that injected carriers produce a very small perturbation in the charge density of the device, causing a negligible variation of the electric field.

Geometry

The simulated geometry is reported in figure 4.2; the dimensions are reported in table 4.1. Since only the carrier transport inside the device is investigated, everything else of the initial structure is removed, i.e. the SiO_2 layer and the tungsten contacts. The mesh is obtained through the program GMSH [19], that given a geometry, creates an unstructured mesh using Delaunay triangulation [22]. In the XY plane (figure 4.2a), the mesh is finer inside and around the germanium layer: in fact, the main variation of the electric field are present in this area. Inside the silicon doped regions the mesh is coarser since the electric field is supposed to be low and not to change too much. Along the Z direction (4.2b), the geometry of the device does not change and the electrical properties are supposed to be constant: therefore, the mesh is coarse. The contacts, visible in 4.2b are ideal and the external surface of the device represents a reflecting boundary for the simulated particles.

Full-band setup

Following the procedure described in the Methods section, the full band method is applied to the description of the germanium layer. In the silicon layers,

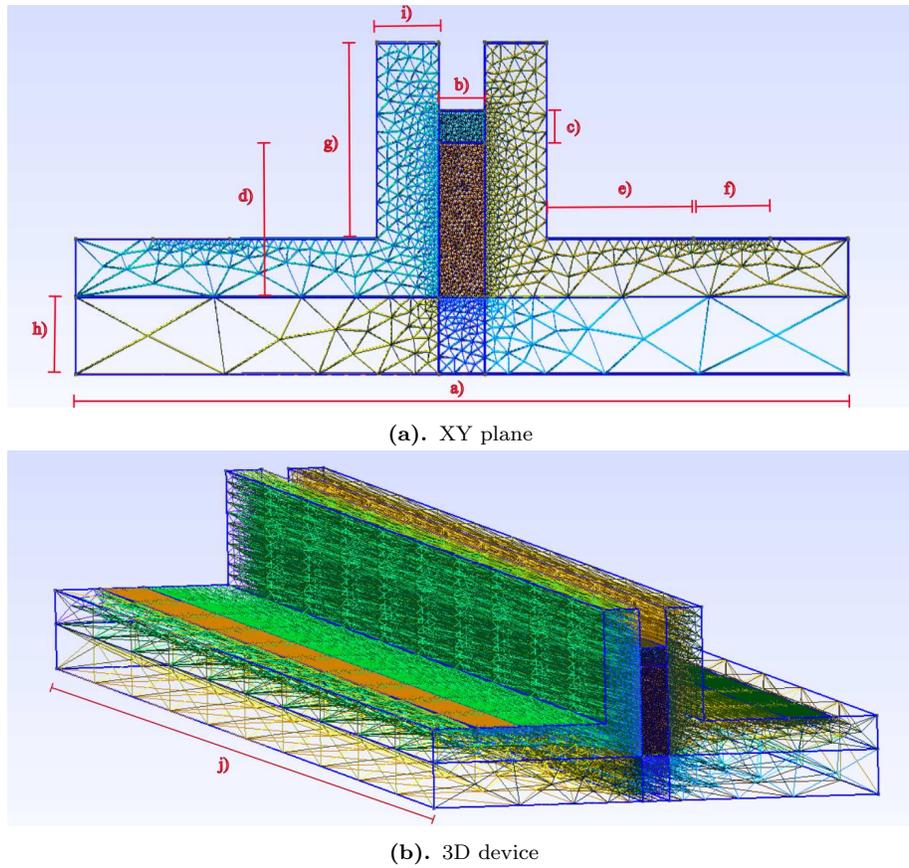


Figure 4.2: Ge-on-Si detector simulated geometry

Si waveguide width a)	$2 \mu\text{m}$
Ge layer width b)	120 nm
Si bridge height c)	82 nm
Ge layer height d)	400 nm
Distance between contact and Si fin e)	380 nm
Contact width f)	200 nm
Si doped height g)	658 nm
Si waveguide height h)	200 nm
Si doped width i)	160 nm
Device length j)	$10 \mu\text{m}$

Table 4.1: Dimensions of the simulated Ge-on-Si photodetector, bold letters correspond to the indications in figure 4.2

Carrier	P_0 (eV $^{-\alpha_0}$ s $^{-1}$)	α_0	E_0 (eV)
e	4.0×10^{11}	4.8	0.8
h	1.2×10^{12}	3.8	1.0

Table 4.2: Values for impact ionization rate 4.5

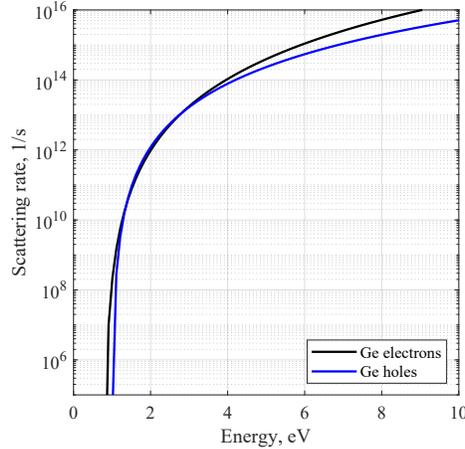


Figure 4.3: Impact ionization rates of Germanium using Keldysh formula

since the electric field is supposed to be low and carriers do not scatter to upper valleys, analytic non-parabolic approximation is used. Germanium is a zincblende crystal, so the shape of the full Brillouin zone in the same as silicon shown in figure 2.18. As for the evaluation electronic band-structure and wavefunctions, EPM was used. The used pseudopotential are reported in figure 4.4 and the energy levels in figure 4.5. With this setup, scattering rates are calculated. The included scattering rates are the following:

- acoustic phonon scattering;
- non-polar optical phonon scattering
- impact ionization

Keldysh formula was used for the impact ionization scattering rate [36]

$$P_{II}(E) = P_0 (E - E_0)^{\alpha_0}, \quad (4.5)$$

where the values of P_0 , E_0 and α_0 are reported in table 4.2. They are shown in figure 4.3.

Material properties are then simulated and confronted with experimental data in order to determine fitting parameters, i.e. acoustic and non-polar optical phonon deformation potential. The results of the simulation of bulk germanium velocity are shown in figure 4.6. As it is possible to observe, a good agreement with experimental data ([41]) was reached.

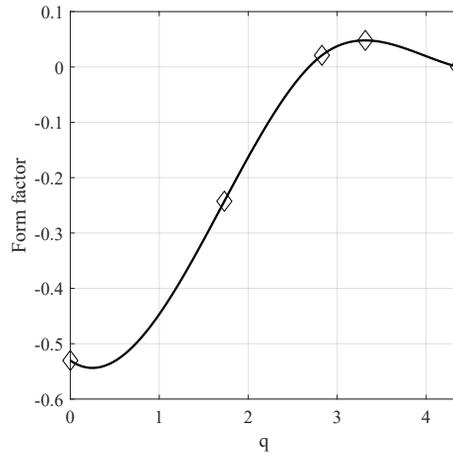


Figure 4.4: Germanium pseudopotential

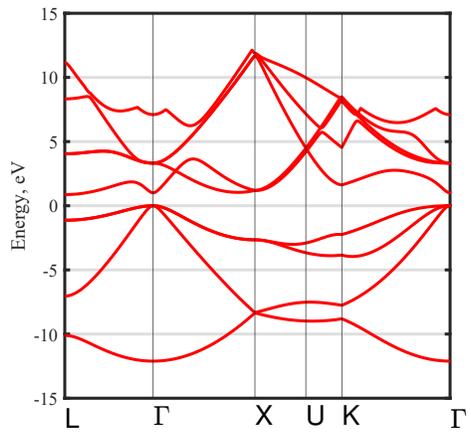


Figure 4.5: Germanium electronic band-structure calculated with EPM

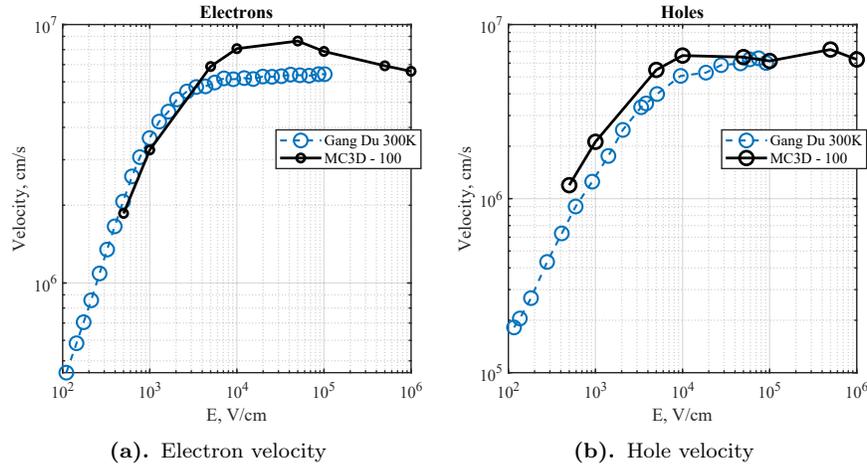


Figure 4.6: Germanium velocity-field curve, $T = 300$ K, experimental data taken from [41]

4.3 Results

The results of the device Monte Carlo simulation for the evaluation of the electric field are reported in figure 4.7. In order to satisfy neutrality equation, the simulated particles' charge must compensate the charge of the ionized donors of p- and n-sides of the device. For this reason, the higher the doping, the higher the number of simulated particles, and the higher the simulation time. The chosen doping for p- and n-sides was equal to $|N^+| = 1 \times 10^{16} \text{ cm}^{-3}$ corresponding to around 140000 simulated particles.

It is possible to visualize the distribution of carriers due to the coupled solution of Boltzmann and Poisson equations in figure 4.8. The points are the electrons, while the tetrahedrons are holes. As expected, carriers rearrange away from the center, where the electric field is higher. A small layer of holes is present at the interface between germanium and silicon because, presumably, an heterostructure is formed between the two materials and holes that have not sufficient energy remain trapped.

However, since the doping level is supposed to be much higher in the real device (about $1 \times 10^{20} \text{ cm}^{-3}$), a simulation of the same device was performed using drift-diffusion tools (i.e. Synopsys Sentaurus), where the high doping density is not as problematic; the results are reported in figure 4.9. As it is possible to observe, the electric fields are different by one order of magnitude in the germanium layer: for this reason, the electric field evaluated with drift-diffusion approach is used as input for the frozen field analysis. It is possible to observe a little spike of electric field near the left Si-Ge interface, better visible in figure 4.10: that is probably due to a molar grading insertion in that area in the Sentaurus device description. In Monte Carlo simulations this molar grading profile is absent and in fact the resulting electric field is symmetric along the

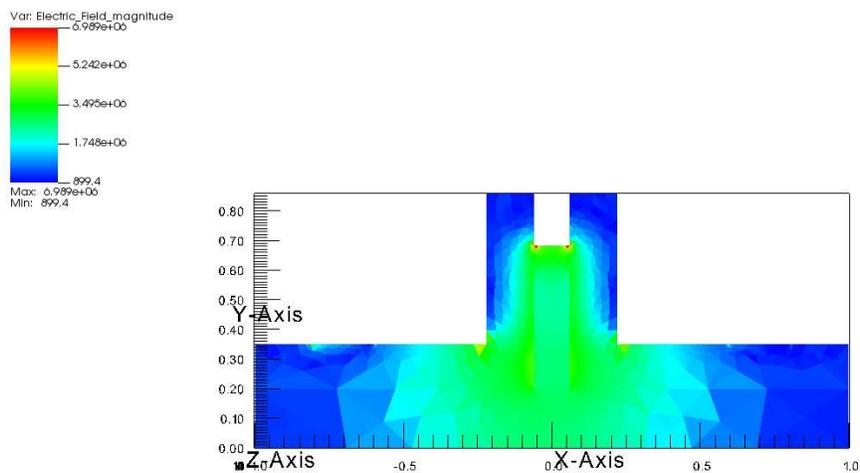


Figure 4.7: Electric field inside the Ge-on-Si photodetector: $V_{\text{bias}} = -2$ V; doping density in n- and p-sides: $1 \times 10^{16} \text{ cm}^{-3}$

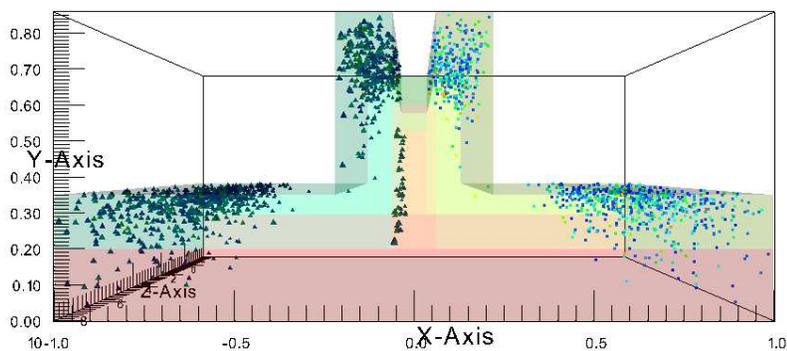


Figure 4.8: Distribution of particles inside the Ge-on-Si photodetector: $V_{\text{bias}} = -2$ V; doping density in n- and p-sides: $1 \times 10^{16} \text{ cm}^{-3}$

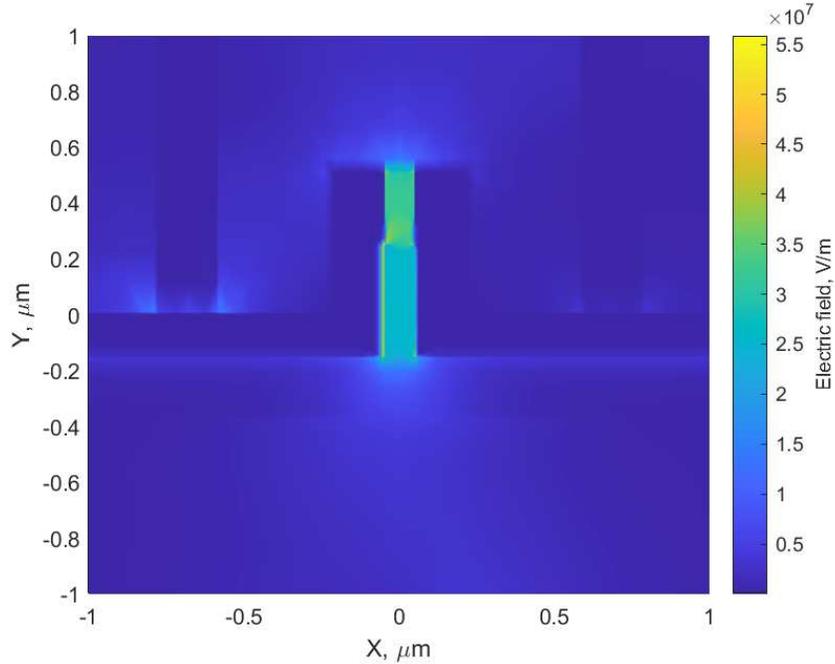


Figure 4.9: Electric field evaluated with drift-diffusion approach (Synopsys Sentaurus)

vertical axis.

The current resulting from the injection is reported in figure 4.11, evaluated with Ramo formula [37]. In order to have an estimation of the frequency response two methods can be applied:

- The FFT can be performed on the current, as explained before. Some problems can arise with this technique according to the choice of the time step and the time span of the simulation;
- Since the frequency response is low-pass, the system can be approximated as single pole, for which the bandwidth is approximately

$$BW \approx \frac{0.35}{t_f}, \quad (4.6)$$

where t_f is the fall time, measured in the falling edge of the signal as the time difference between the instant in which the signal is at 10% of its maximum value t_1 and the instant in which the signal is at 90% of its maximum value t_2 .

Looking at figure 4.11, $t_1 \approx 0.487$ ps and $t_2 \approx 1.702$ ps, resulting in a bandwidth $BW \approx 288$ GHz. The FFT is shown in figure 4.12. In this case, the

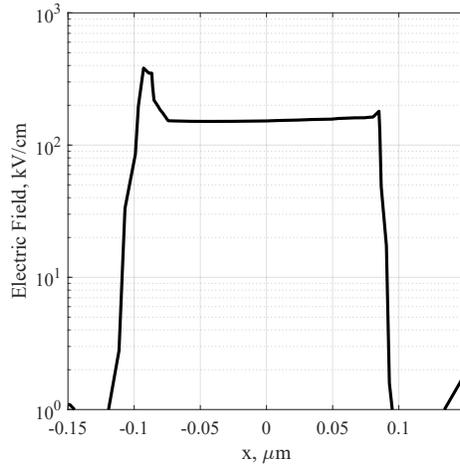


Figure 4.10: Photodetector electric field along x ; $y = 0$, $z = 0$

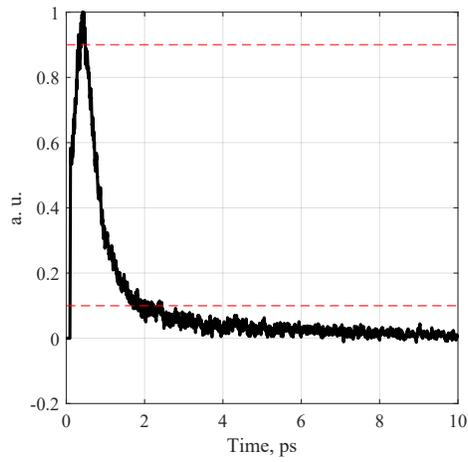


Figure 4.11: Current resulting from carrier injection in the detector under applied bias of $V = -2$ V. The red dashed lines are related to the levels at which the signal is at 10% and at 90% of its maximum value

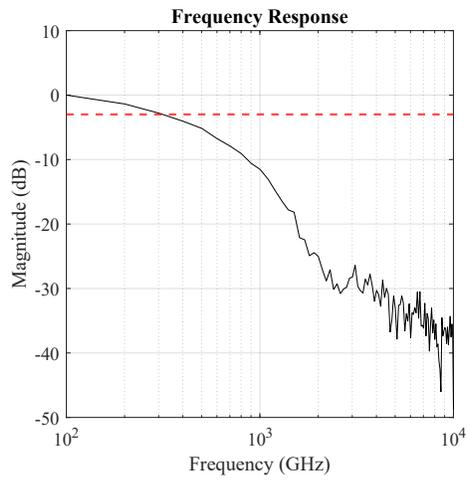


Figure 4.12: FFT of signal in figure 4.11, describing the frequency response of the device

-3 dB bandwidth reaches 300 GHz, greater than the one obtained with the single pole approximation.

Chapter 5

Conclusions

In this work, advanced models have been analyzed to improve the physics framework in Monte Carlo simulations. The full-band model results capable of simulate and predict material properties and electronic devices behaviour quite correctly. However, currently, only the interaction with phonons is properly included in the simulations. A hint for future works can be the integration in the full-band model of the interaction with the ionized impurities, since an increasingly high doping level is used for device materials. In this regard, Brooks-Herring, Ridley and Kosina model are put in comparison under the analytic band approximation in the evaluation of low-field mobility. It was observed that Ridley model gives the same results of the BH model in the case of majority electrons. It would be interesting to perform a comparison in case of compensated semiconductor where these models could lead to different carrier dynamics. The Kosina model returns a lower mobility compared to BH and Ridley model, and in particular this difference is really high (i.e. one order of magnitude) for high degeneracy and low temperature. A possible hint for future work could be improving the model for the evaluation of the mobility introducing other interaction, as electron-plasmon scattering, important at high degeneracy and low temperature. Finally, the introduction of Pauli exclusion principle causes an increase of the mobility for high doping and degeneracy. The largest increase is observed using Kosina model in the condition of high degeneracy. The used algorithm for the Pauli blocking is simple: it can be useful for the evaluation of material properties, but becomes cumbersome when it is applied to device simulations, where multi-threading is fundamental. Improvements to this algorithm, such as the one proposed in [6] could be investigated and implemented in a full-band framework to efficiently take into account degeneracy as well.

Appendix A

Numerical methods

A.1 Bisection method

This algorithm is used to find the root of a continuous function $f(x) = 0$ defined over an interval $[a, b]$. It is effective if $f(a)$ and $f(b)$ have opposite sign, since it can be demonstrated that if this condition applies at least one root exists in the interval (a, b) . A number of iterations is performed until the root is found with the proper tolerance, or the maximum number of iterations has been reached. During the iterations, the midpoint x_2 of the interval in which the root is present $[x_0, x_1]$ ($[a, b]$ in the first iteration) is calculated and two new intervals are created $[x_0, x_2]$ and $[x_2, x_1]$. Then, the interval that contains the root, i.e. whose f evaluated at its extrema have opposite signs, is selected to be the starting interval for the next iteration.

The function used in the code that implements the bisection method is the following.

```
double bisection_method(  
    double (*func)(double, BisectionParams*),  
    double guessmin,  
    double guessmax,  
    BisectionParams *params,  
    const int verbosity  
) {  
    const int max_ref = 100;  
    int ref = 0;  
    const double tol = 1.0e-6;  
    double fmin, fmax, fmid, guessmid;  
    double error;  
  
    // Bisection initialization  
    guessmid = (guessmin + guessmax) / 2.0;
```

```

fmin = func(guessmin, params);
fmax = func(guessmax, params);
fmid = func(guessmid, params);
error = fabs(fmid) / tol;

if (verbosity) {
    printf("Bisection refinement for equilibrium level\n");
    printf("Interval: [%.3f, %.3f];\nMax iterations: %d;\n"
           "Tolerance: %.1e;\n", guessmin, guessmax,
           max_ref, tol);
    printf("Iteration %4d - Error %.3e\n", ref, error);
}

if (signbit(fmin) != signbit(fmax)) {
    while (ref < max_ref && error > 1.0) {
        if (signbit(fmid)) {
            guessmin = guessmid;
            fmin = fmid;
        } else {
            guessmax = guessmid;
            fmax = fmid;
        }
        guessmid = (guessmin + guessmax) / 2.0;
        fmid = func(guessmid, params);
        ref++;
        error = fabs(fmid) / tol;
        if (verbosity) {
            printf("Iteration %4d - Error %.3e\n", ref, error);
        }
    }
} else {
    ERROR_THROW(SIM_ERROR,
               "Bisection cannot operate. Solution is not enclosed "
               "between the two extrema");
}

return guessmid;
}

```

A.2 Simpson integration method

It is used to evaluate definite integrals. Let $f(x)$ be a generic function defined over the interval $[a, b]$. Given an uniform subdivision of $[a, b]$ in n points, with

n an even number, the integral of f can be approximated as

$$\int_a^b f(x) dx \approx \frac{1}{3}h \left[f(x_0) + 4 \sum_{i=1}^{n/2} f(x_{x_{2i-1}}) + 2 \sum_{i=1}^{n/2-1} f(x_{2i}) + f(x_n) \right] \quad (\text{A.1})$$

with $h = (b - a)/n$ and $x_i = a + ih$ for $0 \leq i \leq n$.

The error committed by this algorithm is

$$\delta \leq \frac{1}{180}h^4(b - a) \max_{\xi \in [a,b]} |f^{(4)}(\xi)| \quad (\text{A.2})$$

therefore, in order to have a good approximation, a suitable n must be chosen.

The algorithm used in this work is the following:

```
double simpson(double (*func)(double), double a, double b, int n) {

    double h = (b - a) / n;
    double sum = func(a) + func(b);
    double x;

    for (int i = 1; i < n; i++) {
        x = a + i * h;
        if (i % 2 == 0) {
            sum += 2 * func(x);
        } else {
            sum += 4 * func(x);
        }
    }

    return (h / 3) * sum;
}
```

A.3 Gauss elimination method

The Gaussian elimination method is an algorithm for solving systems of linear equations. Given a system represented by the following augmented matrix:

$$\left(\begin{array}{ccc|c} a_1 & a_2 & a_3 & d_1 \\ b_1 & b_2 & b_3 & d_2 \\ c_1 & c_2 & c_3 & d_3 \end{array} \right) \quad (\text{A.3})$$

the method applies a series of row operations to transform the matrix into an upper triangular form, where all elements below the main diagonal are zero. The allowed operations are:

- Multiplying a row by a nonzero scalar.

- Adding a multiple of one row to another row.

This transformation results in an equivalent system described by a matrix of the form:

$$\left(\begin{array}{ccc|c} a'_1 & a'_2 & a'_3 & d'_1 \\ 0 & b'_2 & b'_3 & d'_2 \\ 0 & 0 & c'_3 & d'_3 \end{array} \right) \quad (\text{A.4})$$

Where a', b', c', d' are the new coefficients after elimination. The unknowns are then calculated using back substitution, starting from the last row (which corresponds to the last variable) and proceeding upwards through the system until all variables are determined.

The algorithm used in this work for the solution of a 3x3 linear system is the following

```
void gaussian_elimination(double A[3][3], double b[], double x[]) {
    int i, j, k;
    float factor;

    // Forward elimination
    for (i = 0; i < 2; i++) {
        for (k = i + 1; k < 3; k++) {
            factor = A[k][i] / A[i][i];
            for (j = i; j < 3; j++) {
                A[k][j] -= factor * A[i][j];
            }
            b[k] -= factor * b[i];
        }
    }
    // Back substitution
    x[2] = b[2] / A[2][2];
    x[1] = (b[1] - A[1][2] * x[2]) / A[1][1];
    x[0] = (b[0] - A[0][1] * x[1] - A[0][2] * x[2]) / A[0][0];
}
```

Bibliography

- [1] M. G. C. Alasio, M. Zhu, M. Matsubara, M. Goano, and E. Bellotti. Ab initio model of carrier transport in diamond. *Phys. Rev. Appl.*, 2024.
- [2] M. Aldegunde and K. Kalna. Energy conserving, self-force free Monte Carlo simulations of semiconductor devices on unstructured meshes. *Computer Physics Communications*, 2015.
- [3] D. Arnold and G. Awanou. The serendipity family of finite elements. *Foundations of Computational Mathematics*, 2011.
- [4] T. Ashley, M. T. Emeny, D. G. Hayes, K. P. Hilton, R. Jefferies, J. O. Maclean, S. J. Smith, A. W-H. Tang, D. J. Wallis, and P. J. Webber. High-performance InSb based quantum well field effect transistors for low-power dissipation applications. In *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009.
- [5] E. Bellotti, L. Tirino, M. Weber, and K. Brennan. A general Monte Carlo model including the effect of the acoustic deformation potential on the transport properties. *Journal of Computational Electronics*, 2004.
- [6] P. Borowik and L. Adamowicz. Improved algorithm for Monte Carlo studies of electron transport in degenerate semiconductors. *Physica B: Condensed Matter*, 2005.
- [7] B.H. Bransden and C.J. Joachain. *Introduction to Quantum Mechanics*. Longman Scientific & Technical, 1989.
- [8] H. Brooks and C. Herring. Scattering by ionized impurities in semiconductors. *Phys. Rev.*, 1951.
- [9] E. Conwell and V. F. Weisskopf. Theory of impurity scattering in semiconductors. *Phys. Rev.*, 1950.
- [10] G. Cronin and S. Borello. Epitaxial InAs on InAs substrates. *J. Electrochem. Soc.*, 1967.
- [11] T. G. Van de Roer and F. P. Widdershoven. Ionized impurity scattering in Monte Carlo calculations. *Journal of Applied Physics*, 1986.

- [12] J. D. Dow, W. E. Packard, H. A. Blackstead, and D. W. Jenkins. Chapter 5: Phonons in semiconductor alloys. In *Phonon Physics: The Cutting Edge*. North-Holland, 1995.
- [13] S. Lischke et al. Ultra-fast germanium photodiode with 3-db bandwidth of 265 Ghz. *Nature Photonics*, 2021.
- [14] M. Farahmand, C. Garetto, E. Bellotti, K.F. Brennan, M. Goano, E. Ghillino, G. Ghione, J.D. Albrecht, and P.P. Ruden. Monte Carlo simulation of electron transport in the III-nitride wurtzite phase materials system: binaries and ternaries. *IEEE Transactions on Electron Devices*, 2001.
- [15] M. Fischetti and S. E. Laux. Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects. *Phys. Rev. B*, 1988.
- [16] M. Fischetti, N. Sano, S. E. Laux, and K. Natori. Full-band Monte Carlo simulation of high-energy transport and impact ionization of electrons and holes in Ge, Si, and GaAs. *1996 International Conference on Simulation of Semiconductor Processes and Devices*, 1996.
- [17] M. Fischetti and W. Vandenberghe. *Advanced Physics of Electron Transport in Semiconductors and Nanostructures*. Springer, 2016.
- [18] S. Fujita and K. Ito. *Bloch Theorem*. Springer New York, 2007.
- [19] C. Geuzaine and J. F. Remacle. Gmsh.
- [20] M. Goano. Algorithm 745: computation of the complete and incomplete Fermi-Dirac integral. *ACM Trans. Math. Softw.*, 1995.
- [21] H. J. Hrostowski, F. J. Morin, T. H. Geballe, and G. H. Wheatley. Hall effect and conductivity of InSb. *Phys. Rev.*, 1955.
- [22] Y. Ito. *Delaunay Triangulation*. Springer Berlin Heidelberg, 2015.
- [23] C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quaranta. A review of some charge transport properties of silicon. *Solid-State Electronics*, 1977.
- [24] C. Jacoboni and P. Lugli. *The Monte Carlo Method for Semiconductor Device Simulation*. Springer-Verlag Wien New York, 1989.
- [25] V. V. Karataev, M. G. Mil'vidsky, N. S. Rytova, , and V. I. Fistui. Compensation in n-type InAs. *Sov. Phys. Semicond*, 1977.
- [26] L.E. Kay. *Monte Carlo studies of ionized impurity scattering in silicon and germanium alloys*. PhD thesis, University of Massachusetts, 1991.
- [27] L.E. Kay and T.W. Tang. An improved ionized-impurity scattering model for Monte Carlo simulations. *J. Appl. Phys.*, 1991.

- [28] P. C. Klipstein, E. Avnon, D. Azulai, Y. Benny, R. Fraenkel, A. Glozman, E. Hojman, O. Klin, L. Krasovitsky, L. Langof, I. Lukomsky, M. Nitzani, I. Shtrichman, N. Rappaport, N. Snapi, E. Weiss, and A. Tuito. Type II superlattice technology for LWIR detectors. In *Infrared Technology and Applications XLII*. SPIE, 2016.
- [29] H. Kosina and G. Kaiblinger-Grujin. Ionized-impurity scattering of majority electrons in silicon. *Solid-State Electronics*, 1998.
- [30] M.A. Littlejohn, J.R. Hauser, T.H. Glisson, D.K. Ferry, and J.W. Harrison. Alloy scattering and high field transport in ternary and quaternary III–V semiconductors. *Solid-State Electronics*, 1978.
- [31] E. Litwin-Staszewska, R. Piotrkowski, and W. Szymanska. Electron mobility and thermoelectric power in InSb at atmospheric and hydrostatic pressures. *Phys. Status Solidi B*, 1981.
- [32] P. Lugli and D. Ferry. Degeneracy in the ensemble monte carlo method for high-field transport in semiconductors. *IEEE Transactions on Electron Devices*, 1985.
- [33] M. Lundstrom. *Fundamentals of Carrier Transport*. Cambridge University Press, 2000.
- [34] N. S. Mansour, K. Diff, and K. F. Brennan. Effect of inclusion of self-consistently determined electron temperature on the electron–plasmon interaction. *Journal of Applied Physics*, 1996.
- [35] J. Meyer and F. Bartoli. Phase-shift calculation of ionized impurity scattering in semiconductors. *Physical review. B, Condensed matter*, 1981.
- [36] M. Michailat, D. Rideau, F. Aniel, C. Tavernier, and H Jaouen. Monte Carlo simulation of high-energy transport of electrons and holes in bulk SiGeC alloys. *J. Phys.: Conf.*, 2009.
- [37] S. Ramo. Currents induced by electron motion. *Proceedings of the I.R.E.*, 1939.
- [38] B.K. Ridley. Reconciliation of the Conwell-Weisskopf and Brooks-herring formulae for charged-impurity scattering in semiconductors: Third-body interference. *J. Phys. C: Solid State Phys.*, 1977.
- [39] A. Rogalski and K. Jóźwikowski. Intrinsic carrier concentration and effective masses in $\text{InAs}_{(1-x)}\text{Sb}_x$. *Infrared Physics*, 1989.
- [40] A. Rogalski, P. Martyniuk, M. Kopytko, P. Madejczyk, and S. Krishna. InAsSb-based infrared photodetectors: Thirty years later on. *Sensors*, 2020.
- [41] S. M. Sze, Y. Li, and K.K. Ng. *Physics of semiconductor devices*. John Wiley & sons, 2021.

- [42] P. T. Webster, N. A. Riordan, S. Liu, E. H. Steenbergen, R. A. Synowicki, Y.-H. Zhang, and S. R. Johnson. Measurement of InAsSb bandgap energy and InAs/InAsSb band edge positions using spectroscopic ellipsometry and photoluminescence spectroscopy. *J. Appl. Phys.*, 2015.
- [43] G. Weisz. Band structure and Fermi surface of white tin. *Physical review*, 1966.
- [44] W. Zawadzki. Electron transport phenomena in small-gap semiconductors. *Advances in Physics*, 2006.