Physics-based simulation of narrow and wide band gap photonic devices

POLITECNICO DI TORINO

SCUOLA DI DOTTORATO

Dottorato in Dispositivi elettronici – XXVIII ciclo

Tesi di Dottorato

Physics-based simulation of narrow and wide band gap photonic devices

Marco Vallone

Tutore
Prof. Michele Goano
Prof. Francesco Bertazzi

Coordinatore del corso di dottorato

Prof. Giovanni Ghione

Aprile 2016

Abstract

Historically, infrared (IR) detector technologies are connected mainly with controlling and night-vision problems: in a first stage, applications concerned simply with detection of IR radiation, but very soon capabilities to form IR images were developed, opening the way to systems for recognition and surveillance, especially for military purposes. Since the last decade of the twentieth century, the use of IR imaging systems for civil and peaceful purposes have increased continuously: these include medical and industrial applications, detection of earth resources, earth and universe sciences, etc. As an example, IR imaging is widely used in astronomy, to study interstellar medium and first-stages of stellar evolution; in medicine, IR thermography – IR imaging of the human body – is employed to detect cancers or other trauma; IR detectors are also widely used in automotive industry, chemical process monitoring, global monitoring of environmental pollution and climate changes, etc.

The discovery in 1959 by Lawson and co-workers of the wide tunability of the HgCdTe alloy allowed this compound to become one of the most important and versatile materials for detector applications over the entire IR range, spanning the short wavelength IR (SWIR: $1-3\,\mu$m), middle wavelength IR (MWIR: $3-5\,\mu$m), long wavelength IR (LWIR: $8-14\,\mu$m) and very long wavelength IR (VLWIR: $14-30\,\mu$m) bands.

Nevertheless, IR detector technology was and continues to be primarily driven by military applications. One of the negative aspects related to the support of defense agencies is the associated secrecy requirements that inhibit meaningful collaborations among research teams, especially on an international level. In addition, the main focus has been on the development of focal plane arrays for infrared cameras, rather than on establishing the basic knowledge.

A critical contribution to research is given by Technology Computer-Aided Design (TCAD), modeling and simulation. In the first part of this thesis, I present the main part of my research activity, focused on the development of abilities and methodologies for the simulation of realistic three-dimensional HgCdTe-based infrared photodetectors, in particular making use of the commercial simulator TCAD Sentaurus by Synopsys [1]. The purpose is the investigation of generation-recombination (GR) mechanisms and modeling of spectral photoresponse in narrow-gap HgCdTe-based photodetectors, with one-, two- and three-dimensional (1D, 2D, 3D) realistic TCAD models.

This activity – Chapters 1-5 – has been carried on within a collaboration with AIM Infrarot-Module GmbH, Theresienstraße 2, D-74072 Heilbronn, Germany, an industry leader in the production of IR photodetectors, mainly based on HgCdTe.

Many useful comments and discussions come from the Boston University (prof. E. Bellotti and collaborators), One Silber Way, Boston, MA 02215, U.S.A., especially about HgCdTe transport and absorption properties.

Coming into details, the structure of the first part of my thesis is the following:

Chapters 1-2 An overview of the most important types of IR detectors fabricated in the last decades is given. The HgCdTe material software library, developed as a pre-requisite for the realization of this work (a collection of Tcl, C++ and Sentaurus scripts), along with the electrical and optical simulation models here employed, are also presented.

Chapter 3 Considering an IR photodetector designed as a $N \times N$ focal plane array of pixels, the building blocks developed and described in Chapter 2 are employed to systematically compare 1D, 2D and 3D simulations of a single-color and dual band HgCdTe isolated pixel, discussing the non-trivial differences among 1D, 2D and 3D results. Concepts presented in this chapter have been published in Ref. [2].

Chapter 4 Still considering a photodetector fabricated as a $N \times N$ focal plane array of pixels, a combined experimental and numerical simulation study is presented on two sets of nominally identical HgCdTe single-color back-illuminated midwave infrared n-on-p isolated pixel grown by liquid-phase epitaxy, p-doped with Hg-vacancies and with Au, respectively. The present numerical model includes a novel formulation for band-to-band tunneling, which overcomes the intrinsic limitations of the classical Kane's description without introducing numerical issues typical of other approaches. A significant contribution to the dark current in both sets of devices is attributed to impact ionization, crucial to obtain a satisfactory explanation of the measured characteristics also at low-to-intermediate bias. Concepts presented in this chapter have been published in Ref. [3].

Chapter 5 Simulations are shown concerning HgCdTe-based LWIR detectors, focusing on methodological comparisons between the finite difference time domain (FDTD) and ray tracing optical models. Applicability and limitations of ray tracing method are discussed; in addition, the FDTD method allowed to enlighten and describe interesting inter-pixel optical interference and cavity effects. Concepts presented in this chapter have been published in Ref. [4].

Another important topic of industrial research in semiconductor physics deals with nitride-based light-emitting diodes (LEDs). From automotive to streetlights, from lights in our houses to the displays of TVs and smartphones, LED-based technology is making its way in the market. This proliferation would have been impossible without GaN-based LEDs, whose invention by Isamu Akasaki, Hiroshi Amano and Shuji Nakamura has been

rewarded with the 2014 Nobel Prize in Physics. Nevertheless, GaN-based LEDs performance is limited by a reduction (droop) of their internal quantum efficiency (IQE) as the driving current density is increased beyond ≈ 10 A/cm^2, whose physical origin is still under intense debate. In the second part of this thesis, I present a quantum model, based on condensed matter many-body theory, that allowed to obtain the electron capture time and hot-electron intraband relaxation times in a quantum well (QW)-barrier heterostructure, for longitudinal optic (LO) phonon emission.

Many useful comments and discussions about the IQE droop and general properties of GaN-based LEDs come from the Università di Modena e Reggio Emilia (prof. G. Verzellesi and collaborators), Via Università, 4, 41121 Modena, Italy, and from the Università di Padova (prof. M. Meneghini and collaborators), Via 8 Febbraio, 2 - 35122 Padova, Italy.

The second part of this thesis is structured as follows:

Chapter 6 After an introductory part about the general concept of self-energy, a quantum model is developed in order to obtain electron capture and hot electron intraband relaxation times in a quantum well, for electron LO phonon scattering. In particular, the effect of carrier density and electron energy have been investigated, obtaining semi-analytic expressions as function of carrier density, a topic often neglected in literature, despite its fundamental interest in semiconductor physics. It is shown that the scattering through the emission of pure LO-phonons is not a good approximation when the population increases, whereas the interplay between LO-phonon and collective plasma modes must be considered. Novel semi-analytic expressions in the single plasmon pole dynamical form of the random phase approximation were obtained, without making use of the more usual static limit of it. In this first chapter, the general theory is developed and numerical examples refer to III-V materials. Concepts presented in this chapter have been published in Ref. [5].

Chapter 7 The concepts exposed in Chapter 6 have been extended to nitride-based LEDs and tested against experimental data available in literature. The proposed approach produced a closed-form expression for capture time as a function of the carrier densities in LEDs QW and barrier states. Its application to simple two-population rate equations allowed to reproduce available experimental data with excellent agreement, offering an accurate yet practical alternative to the usual approximation of a constant capture time in modeling light emitting diodes and lasers. Concepts presented in this chapter have been published in Ref. [6].

Articles published on International Journals

- M. Vallone, “Quantum well electron scattering rates through longitudinal optic-phonon dynamical screened interaction: An analytic approach”, J. Appl. Phys., vol. 114, n. 5, p. 053704-9, 2013, doi: 10.1063/1.4817242

- M. Vallone, M. Goano, F. Bertazzi, G. Ghione, R. Wollrab, J. Ziegler, “Modeling photocurrent spectra of one-color and dual-band HgCdTe photodetectors: is 3D simulation unavoidable?”, J. Electron. Mater., vol. 43, n. 8, p. 3070-3076, 2014, doi: 10.1007/s11664-014-3252-9

- M. Calciati, M. Goano, F. Bertazzi, M. Vallone, X. Zhou, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, E. Bellotti, G. Verzellesi, D. Zhu, C. Humphreys, "Correlating electroluminescence characterization and physics-based models of InGaN/GaN LEDs: Pitfalls and open issues." Aip Advances, vol. 4; p. 067118-1-067118-23, 2014, doi: 10.1063/1.4882176

- M. Vallone, F. Bertazzi, M. Goano, G. Ghione, “Model for carrier capture time through phonon emission in InGaN/GaN quantum wells”, Physica Status Solidi B, vol. 252, n. 5, p. 971-976, 2015, doi: 10.1002/pssb.201451580

- M. Mandurrino, G. Verzellesi, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, “Physics-based modeling and experimental implications of trap-assisted tunneling in InGaN/GaN light-emitting diodes”, Physica Status Solidi B, vol. 212 n. 5, pp. 947-953, 2015, doi:10.1002/pssa.201431743

- M. Mandurrino, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, G. Verzellesi, M. Meneghini, G. Meneghesso, E. Zanoni, “Semiclassical simulation of trap-assisted tunneling in GaN-based light-emitting diodes”, J. Comp. Electron., vol. 14, n. 2, p. 444-455, 2015, doi: 10.1007/s10825-015-0675-3

- M. Vallone, M. Mandurrino, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, H. Figgemeier, “Numerical modeling of SRH and tunneling mechanisms in High Operating Temperature MWIR HgCdTe photodetectors”, J. Electron. Mater., vol. 44, n. 9, p. 3056-3063, 2015, doi: 10.1007/s11664-015-3767-8

- M. Vallone, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, H. Figgemeier, “Comparing FDTD and ray tracing models in the numerical simulation of HgCdTe LWIR photodetectors ”, accepted for publication in the J. Electron. Mater. (forthcoming 2016).

Contributions presented in International Conferences

- M. Vallone, M. Goano, F. Bertazzi, G. Ghione, R. Wollrab, J. Ziegler, "Modeling photocurrent spectra of one-color and dual-band HgCdTe photodetectors: is 3D simulation unavoidable?", II-VI US-Workshop, October 1-3, 2013, Chicago, IL, U.S.A. (oral presentation)

- M. Vallone, F. Bertazzi, M. Goano, G. Ghione, "Quantum model for carrier capture time through phonon emission in InGaN/GaN LEDs", 14th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD), 1-4 Sept. 2014, Palma de Mallorca (Spain), doi: 10.1109/NUSOD.2014.6935325 (poster)

- M. Mandurrino, G. Verzellesi, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, "Trap-assisted tunneling in InGaN/GaN LEDs: experiments and physicsbased simulation", 14th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD), 1-4 Sept. 2014, Palma de Mallorca (Spain), doi: 10.1109/NUSOD.2014.6935325 (poster)

- M. Vallone, F. Bertazzi, M. Goano, G. Ghione, "Quantum model for carrier capture time through phonon emission in InGaN/GaN LEDs", The International Workshop on Nitride semiconductors (IWN), 24-29 Sept. 2014, Wroclaw (Poland) (poster)

- M. Mandurrino, G. Verzellesi, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, "Trap-assisted tunneling in InGaN/GaN LEDs: experiments and physicsbased simulation", The International Workshop on Nitride semiconductors (IWN), 24-29 Sept. 2014, Wroclaw (Poland) (poster)

- M. Vallone, M. Mandurrino, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, H. Figgemeier, "Numerical modeling of SRH and tunneling mechanisms in High Operating Temperature MWIR HgCdTe photodetectors", II-VI US-Workshop, October 21 - 23, 2014, Baltimore, MD, U.S.A. (oral presentation)

- M. Vallone, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, H. Figgemeier, "Comparing FDTD and ray tracing models in the numerical simulation of HgCdTe LWIR photodetectors ", II-VI US-Workshop, October 5 - 8, 2015, Chicago, IL, U.S.A. (oral presentation)

Contents

List of Tables

List of Figures

List of acronyms and symbols

Acronyms and abbreviations

1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
BLIP	Background Limited Infrared Photodetector
BTBT	Band-To-Band-Tunneling
FDTD	Finite Difference Time Domain
FEM	Finite Element Method
FPA	Focal Plane Array
GR	Generation-Recombination
HOT	High Operating Temperature (detector)
I.Ion.	Impact Ionization
IQE	Internal Quantum Efficiency
IR	Infrared
LED	Light Emitting Diode
LO	longitudinal optical, referred to phonons mode
LPE	Liquid Phase Epitaxy
LWIR	Long-Wave InfraRed
MBE	molecular beam epitaxy

MOCVD	Metal-Organic Chemical Vapor Deposition
MQW	Multi-Quantum Well
MWIR	Middle-Wave InfraRed
NEΔT	Noise-Equivalent Difference of Temperature
NEP	Noise Equivalent Power
PDE	partial differential equation
QW	Quantum Well
RCWA	Rigorous Coupled-wave analysis
RPA	Random Phase Approximation
SRH	Shockley-Read-Hall
SWIR	Short-Wave InfraRed
TAT	Trap-Assisted Tunneling
TCAD	Technology Computer Aided Design
TMM	Transfer Matrix Method
VLWIR	Very Long-Wave InfraRed

Constants, parameters and variables

α_0	fine structure constant
ϵ_0	vacuum dielectric permittivity
ϵ_∞	high frequency (dynamic) dielectric constant
ϵ_s	low frequency (static) dielectric constant
$\hbar$	reduced Planck's constant
λ	wavelength
$\mathcal{R}$	detector's responsivity
ω_{LO}	longitudinal optical phonon frequency
c	light velocity

D^*	normalized and area-independent detectivity
h	Planck's constant
K_ϵ	$\epsilon_\infty^{-1} - \epsilon_s^{-1}$
k_B	Boltzmann's constant
n	electron density
N_A	density of acceptors
N_A^-	density of ionized acceptors
N_D	density of donors
N_D^+	density of ionized donors
p	hole density
q, q_e	electron charge
T	absolute temperature

Part I

Narrow band gap devices

Chapter 1

Infrared detectors, materials and simulation tools

Infrared (IR) imaging is widely used for both military and civilian applications. Military applications include target acquisition, surveillance, night-vision and tracking to guide defense interceptor seekers. Civilian applications include thermography, short-range wireless communication, spectroscopy, weather forecasting and infrared astronomy. The related technology has been significantly improving over the years and is finding new areas of applications as the cost for high sensitivity sensors keeps decreasing. But because of the absorption by earth's atmospheric CO_2 and H_2O, not all wavelengths can be used for free space optical data transmission in the infrared region. Looking at the transmission of

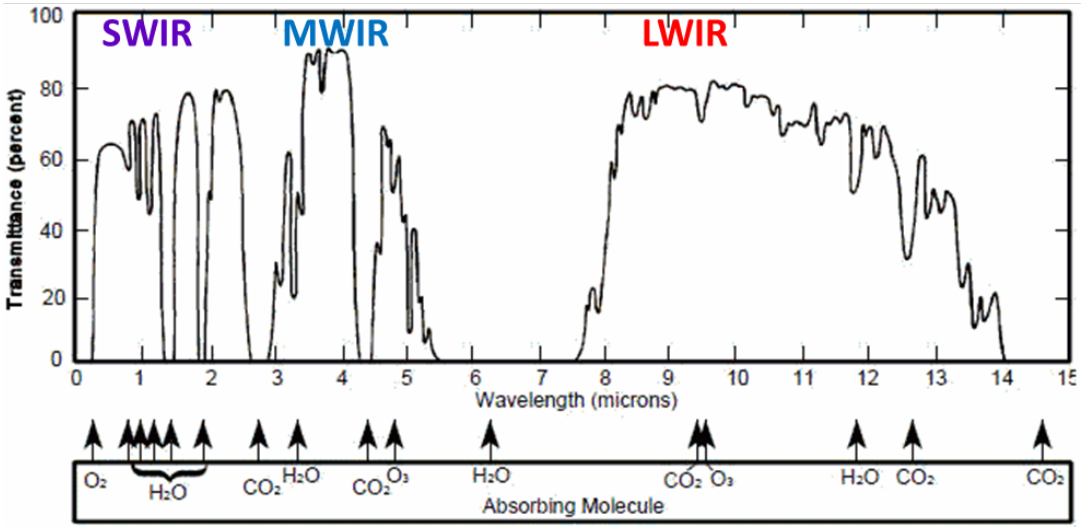

Figure 1.1. Atmospheric transmittance of electromagnetic radiation vs. wavelength.

air as a function wavelength (in Fig. 1.1), it is evident that only a few wavelength windows are available for such purpose. Therefore, interest is centered mainly on the wavelengths (λ) of the atmospheric windows "short-wave infrared" (SWIR, $\lambda = 1 \div 3\,\mu$m), "mid-wave infrared" (MWIR, $\lambda = 3 \div 5\,\mu$m), "long-wave infrared" (LWIR, $\lambda = 8 \div 14\,\mu$m) and "very long-wave infrared" (VLWIR, $\lambda = 14 \div 30\,\mu$m, not shown in the figure).

In this chapter the most important types of IR detectors and materials are reviewed (Sec. 1.1), among which a special position is kept by two dimensional (2D) focal plane arrays (FPAs) based on photodiodes (Sec. 1.2). This is followed by the presentation of the simulation method employed in this thesis (Sec. 1.3). Since the HgCdTe (mercury-cadmium-telluride) alloy, among all the most suitable materials for IR detectors, occupies a central position, in Chapter 2 its main properties are reviewed and the development of a composition- and temperature-dependent HgCdTe software library to be integrated into a commercial simulator is described. With this preliminary, but unavoidable tool (not provided in the general-purpose simulator software package), many investigations about HgCdTe-based photodetectors became possible, as will be shown in Chapters 2-4.

1.1 Types of infrared detectors

In general, each application requires a sensor capable of detecting the IR radiation in one or several of the windows described previously. In order to understand how they work, what are their limitations and what could be the future of IR technology, it is meaningful to examine the road already traveled.

First IR detectors were focused on the simple detection of IR radiation, without pretending to obtain any image of the scene in IR wavelengths. Many materials have been investigated in the IR field exploiting several physical phenomena, and P. R. Norton stated [7] that *"all physical phenomena in the range of about 0.1-1 eV can be proposed for IR detectors"*. According to the involved physical phenomena, IR detectors can be firstly divided in two large classes:

- thermal detectors
- photodetectors.

1.1.1 Thermal detectors

J. Seebeck in 1821 realized the first **thermocouple**: he discovered that a difference of potential at the junction of two dissimilar conductors could be generated by a change in temperature. Using this effect, Melloni produced the first Bi-Cu thermocouple detector in 1833, to investigate the IR spectrum. Connecting several thermocouples in series, Nobili in 1829 generated a higher and therefore measurable voltage.

Another widely used thermal detector is the **bolometer**. The bolometer is a resistive element in which the incident radiation is absorbed to change the *material temperature*, causing a change in electrical resistivity. The device is operated by passing an accurately controlled bias current through the detector and monitoring the output voltage. In the case of a bolometer, radiant power produces a temperature increase within the material, which in turn produces the resistance change. There is no direct photon-electron interaction, as for photoconductors, examined in Section 1.1.2.1. The first bolometer was designed in 1880 by American astronomer S. P. Langley for solar observations. Langley was able to make bolometers that were more sensitive than the thermocouples available at that time. Although other thermal devices have been developed since that time, the bolometer remains one of the most used infrared detectors.

In **pyroelectric detectors**, materials show a temperature dependent spontaneous electrical polarization. A thermal detector can also be considered a forward-biased diode [8], since the output current is affected by temperature, hence by a IR radiation impinging on it.

As a general remark, in thermal IR detectors the incident radiation is absorbed to change the *material temperature*, causing a change in *some physical property*, for example the electrical resistance (bolometers) or potential difference across the detector, exploiting the pyroelectric or thermoelectric effects, but there is not a photon-electron interaction or electron-hole photogeneration, as in photodetectors (see Section 1.1.2).

1.1.2 Photodetectors

More interesting is the second class, the **photodetectors**, that constitutes the modern class of IR detectors. They started to be developed in Germany prior to World War II (WWII) making use of PbS and later PbSe, PbTe and InSb [8–10], exploiting the **photoconductive effect** and measuring a variation of resistance induced by the shining radiation (as remarked, this effect must not be confused with the principle of operation of the bolometer).

1.1.2.1 Photoconductive IR detectors

Materials that, exploiting the photoconductive effect, are suitable to build a photodetector can be divided in two classes:

- **extrinsic photoconductors**: in a doped semiconductor, the majority carriers located in extrinsic impurity levels located in the bandgap (Fig. 1.2) can be excited by the radiation to the majority carrier band. Hence, in this type of photoconductors, the incoming radiation promotes an impurity-to-band transition, involving only one type of carrier. The most employed materials were Si:Ga, Si:As, Ge:Cu and Ge:Hg.

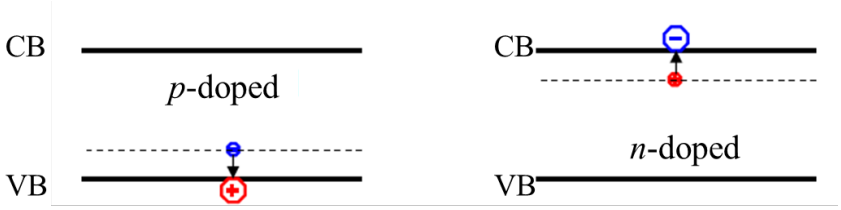

Figure 1.2. Extrinsic p-doped (left) and n-doped (right) photoconductors. CB and VB stand respectively for conduction and valence band. The incoming radiation promotes an impurity-to-band transition, involving only one type of carrier.

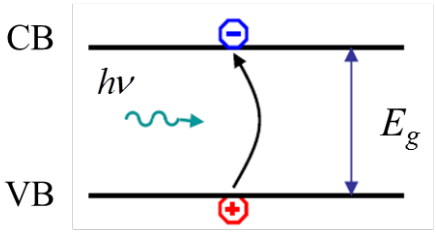

Figure 1.3. Intrinsic photoconductors: the incoming radiation of energy $h\nu$ promotes an electron across the bandgap, hence involving two types of carrier.

Absorption coefficients are relative small, typically in the order of 10 cm^{-1}, the operating temperature is around 40 K and the available IR spectral bands for detection depend only on the energy levels of impurities with respect to the majority carrier band. They constitute the earliest form of photoconductors, developed since 1940s until the early 1960s as linear FPAs, and exploit the photoconductive effect, since the net-effect of the radiation is an increase of the semiconductor conductivity (or a reduction of its resistivity), due to the increase of the majority carrier density.

- **intrinsic photoconductors**: considering intrinsic semiconductor with energy gap E_g, the radiation with energy $h\nu \geq E_g$ can be absorbed producing electron-hole pairs. Hence, differently from the previous case, the photoconduction takes place through minority carriers photogenerated current (Fig. 1.3). For alloys with direct bandgap at the Γ point, the absorption coefficients are quite large, typically in the order of 1000 cm^{-1}, and the spectral response of the photoconductor is determined primarily by the bandgap of the semiconductor.

During WWII and thereafter, the development of the modern IR detection and imaging techniques started, aiming at obtaining true images of the scene in IR bands. First, a simple row of detectors were employed: an image was generated by scanning the scene across the strip using, as a rule, a mechanical scanner. This constituted a 1D FPA and these kind of detectors, still exploiting the photoconductive effect, constituted the so-called 1$^{\mathbf{st}}$**-generation** FPA [11].

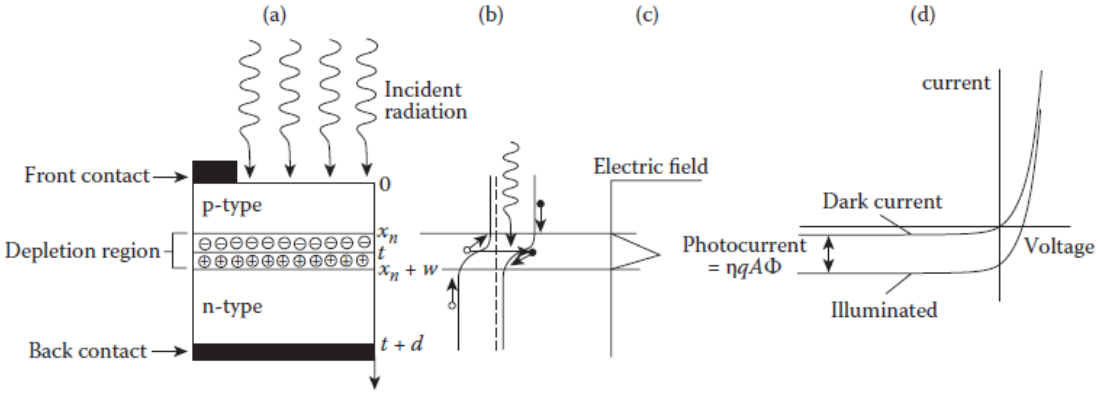

Figure 1.4. $p - n$ junction photodiode: (a) structure of abrupt junction, (b) energy band diagram, (c) electric field, and (d) current-voltage characteristics for the illuminated and nonilluminated photodiode. The figure is from [8, Ch. 9].

The discovery of new narrow gap materials in the Fifties and Sixties allowed the development of more versatile detectors. In 1959, research by Lawson and co-workers [12] triggered development of variable bandgap $Hg_{1-x}Cd_xTe$ alloys, providing an unprecedented degree of freedom in infrared detector design: in this II-VI pseudo-binary semiconductor alloy, the Γ point direct bandgap can be tuned from that of the semimetal HgTe (-0.3 eV) to that of the CdTe (1.5 eV) simply varying the mole fraction x of the Cd in the composition.

1.1.2.2 Photovoltaic IR detectors

Since Eighties new photodetectors started to exploit the **photovoltaic effect** in semiconductors: the difference with the photoconductive detectors consists in the fact that the IR radiation produces a (photo)current, instead of a variation of their resistance. The greatest part of IR detectors exploiting the photovoltaic effects are **narrow-gap semiconductor** $p - n$ **photodiodes**, but also Schottky barriers, metal-insulator-semiconductor photocapacitors and many types of heterostructure may be designed to build an IR photodetector.

Anyway, considering the most common example of a photovoltaic detector, i.e. the abrupt $p - n$ junction constituting the **photodiode**, its operating principle is illustrated in Fig. 1.4. Photons with energy greater than the energy gap, incident on the front surface of the device, create electron-hole pairs in the material on both sides of the junction. By diffusion, the electrons and holes generated within a diffusion length from the junction reach the space-charge region. Then electron-hole pairs are separated by the strong electric field; minority carriers are readily accelerated to become majority carriers on the other side. This way a photocurrent is generated, shifting the current-voltage characteristic in

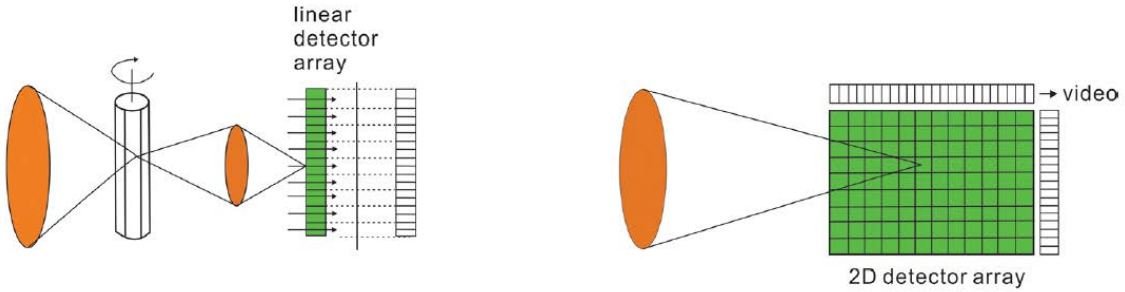

Figure 1.5. First (left) and second (right) generation of FPAs.

the direction of reverse current (see Fig. 1.4(d)). The photocurrent I_{ph} is determined by

$$I_{ph} = \eta q A \Phi \tag{1.1}$$

where η is the quantum efficiency (the number of electron-hole pairs generated per incident photon), Φ is the optical photon flux (photons per unit area and time), A and q are respectively the detector illuminated area and the elementary charge. In Sec. 1.2.3 more details will be given.

The outstanding properties of HgCdTe alloys (see Chapter 2 for an extensive description) triggered the development of detectors based on large 2D FPAs, constituting the core of modern IR vision and imaging systems. This kind of detectors are designed to detect radiation only in one of the IR bands and constitute one of the so called $2^{\mathbf{nd}}$**-generation** IR detectors (Fig. 1.5). Each pixel can be considered an independent photodetector with its size in the order of a few microns, capable to convert the optical flux in an electric signal, and the FPA is constituted by a $N \times N$ array of pixels (Fig. 1.5), with N in the order of thousands. Pixels are scanned electronically by readout integrated circuits (ROICS) that are hybrid packaged with the arrays often by indium bumps (see Fig. 1.6).

Detector arrays are usually illuminated from the back side with photons passing through the transparent detector array substrate, often a CdZnTe layer transparent in IR bands, on which a few microns of HgCdTe (the absorber) are epitaxially grown (see Section 2.1). The Cd molar fraction of the HgCdTe layer is chosen conveniently, in order to obtain an energy gap suitable for SWIR, MWIR or LWIR bands. The $p-n$ junction can be obtained by ion-implantation or other suitable techniques.

Multicolour capabilities are highly desirable for advanced IR detectors [11]. Systems that gather data in separate IR spectral bands (see Fig. 1.7) can discriminate both absolute temperature and unique signatures of objects in the scene. By providing this new dimension of contrast, multiband detection also enables advanced colour processing algorithms to further improve sensitivity above that of single-colour devices: all these and other features identify the $3^{\mathbf{rd}}$**-generation** FPAs. Such multispectral detection permits rapid and

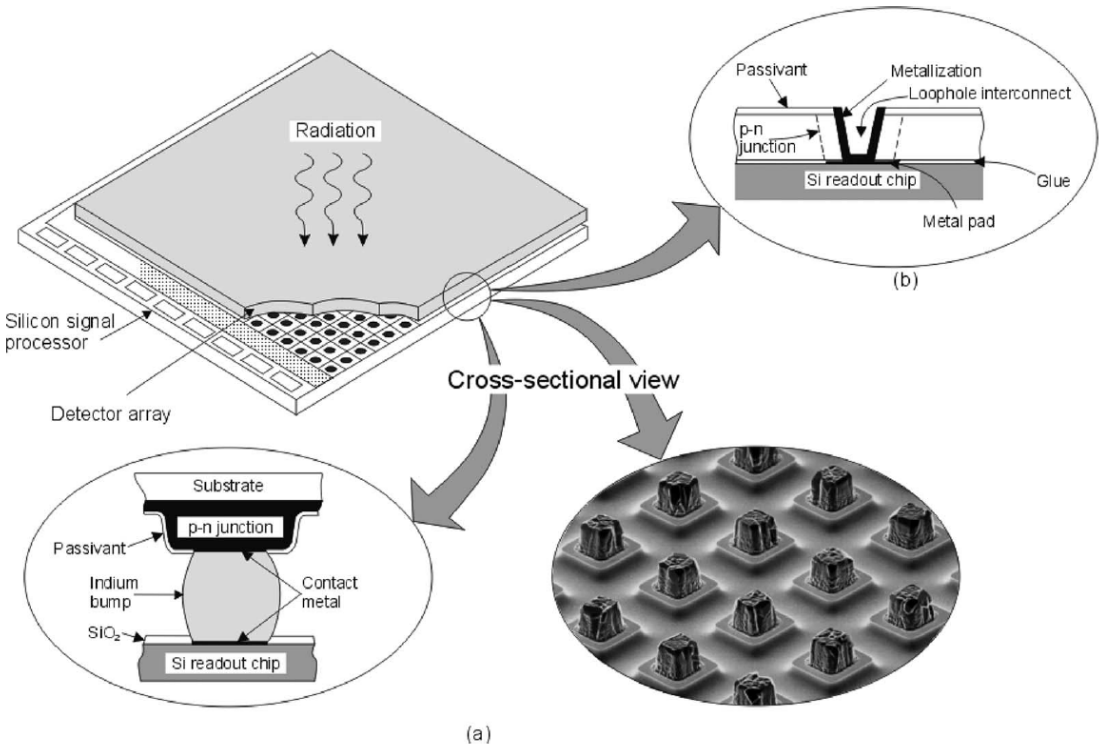

Figure 1.6. Hybrid IR FPA with independently optimized signal detection and readout: (a) indium bump technique; (b) loophole technique. The figure is from [11].

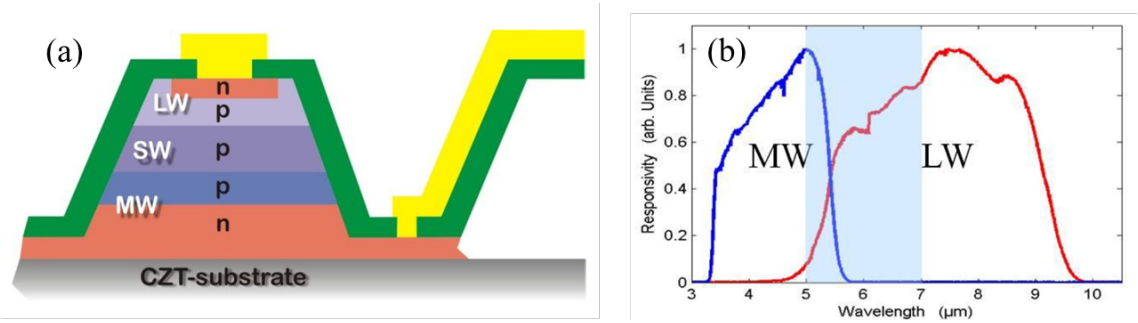

Figure 1.7. (a) Schematic of dual-band MWIR / LWIR detector pixel; (b) its experimental normalized spectral response spectrum at about 60 K. The cut-off wavelengths of the two bands were 5.4 μm for the MWIR and 9.1 μm for the LWIR band. The figure is from [13].

efficient understanding of the scene in a variety of ways. In particular, two-colour IR FPAs can be especially beneficial for threat-warning applications. By using two IR wavebands, spurious information, such as background clutter and sun glint, may be subtracted from an IR image, leaving only the objects of interest. Multispectral IR FPAs can also play many important roles in Earth and planetary remote sensing, astronomy, etc. For these

reasons, the effective signal-to-noise ratio of two-colour IR FPAs greatly exceeds that of single colour IR FPAs for specific applications.

It should be remarked that this third generation of 2D FPA is still developing in the present days, and also its definition is not particularly well established yet. In the common understanding, third-generation IR systems provide enhanced capabilities such as larger number of pixels, higher frame rates, better thermal resolution, as well as multicolor functionality and other on-chip signal-processing functions [14].

1.2 Photodiode-based detectors: principle of operations

Let us consider the simplest macroscopic description of a photodetection system exploiting the photovoltaic effect, that is a reverse biased *n*-*p*-diode (the positive bias voltage is at the *n*-contact, whereas *p*-doped side is connected to ground), schematically shown in Fig. 1.4, in dark conditions and without applied bias. Due to the built-in potential, at the junction between the two doped regions the excess of holes in valence band of the *p*-side tends to compensate the excess of free electrons in the conduction band of the *n*-side generating a two-fold free carriers *diffusion current*: electrons diffuse into *p*-type region and holes into *n*-type region up to an equilibrium condition. The resultant of this diffusion is the formation at the junction of a free carrier depletion region with remaining ionized atoms. This ionization, in turn, generates an electrical potential (and field $\vec{\mathcal{E}}$) acting against the further diffusion through a *drift current* mechanism according to which any free carrier in the depleted region is drifted into its appropriate side (electrons are swept in *n*- and holes in *p*-type region). The reverse bias acts as a further driving force in the drift mechanism, enhancing it with respect to the diffusion. This leads to an increasing of the potential, through a so-called *space charge region*, and also of the width of the depleted region.

If now the illumination is considered, intrinsic absorption processes can occur by the production of electron-hole pairs. Electrons generated in the *p*-type region diffuse in this region and here partly recombine with holes, while holes generated in *n*-type region diffuse in *n*-region and there partly recombine with electrons. Pairs generated in the depleted region and outside it, but at a distance from the depletion region edge *less than the diffusion length*, instead, are separated: thanks to the electric field these new carriers are drifted into their respective sides, resulting in a small detectable reverse current – the *photocurrent* – proportional to the radiation intensity.

1.2.1 Short remarks on the drift-diffusion model

Considering a $n - p$-diode, the *drift-diffusion* approximation of the electrical transport equations yields a mathematical model well suited for the present investigation. In short, under the assumption of Fermi-Dirac statistics, a drift-diffusion model [15] is defined by

a set of three coupled differential equations, namely, the Poisson equation

$$\vec{\nabla} \cdot \epsilon \left(\vec{\nabla} \psi \right) = -q \left(p - n + N_D^+ - N_A^- \right) - \rho_{\text{trap}} \tag{1.2}$$

and the continuity equations for electrons and holes

$$\vec{\nabla} \cdot \vec{J_n} - q(R_n - G_n) = q \frac{\partial n}{\partial t} \tag{1.3}$$

$$-\vec{\nabla} \cdot \vec{J_p} - q(R_p - G_p) = q \frac{\partial p}{\partial t} \tag{1.4}$$

where the current densities are given by

$$\vec{J_n} = -nq\mu_n \vec{\nabla} \phi_n, \qquad \vec{J_p} = -pq\mu_p \vec{\nabla} \phi_p. \tag{1.5}$$

Here ϵ is the electrical permittivity, q is the elementary electron charge, n and p are the electron and hole densities, N_D^+ is the density of ionized donors, N_A^- is the density of ionized acceptors, $R_{n,p} - G_{n,p}$ are the net electron and hole recombination rates, $\vec{J_n}$ is the electron current density, $\vec{J_p}$ is the hole current density, μ_n and μ_p are the electron and hole mobilities, ϕ_n and ϕ_p are the electron and hole quasi-Fermi potentials, and ρ_{trap} is the charge density contributed by traps and fixed charges. In the construction of the model it has been assumed the validity of the Einstein relation between the diffusion coefficients $D_{n,p}$ for electrons and holes and the mobilities

$$D_{n,p} = \frac{k_B T}{q} \mu_{n,p} \tag{1.6}$$

where k_B is the Boltzmann's constant and T the absolute temperature. Electron and hole densities n and p can be computed from the electron and hole quasi-Fermi potentials, and vice versa, under the assumption of Fermi-Dirac statistics:

$$n = N_C \mathcal{F}_{1/2} \left(\frac{E_{F_n} - E_C}{k_B T} \right), \qquad p = N_V \mathcal{F}_{1/2} \left(\frac{E_V - E_{F_p}}{k_B T} \right) \tag{1.7}$$

where $\mathcal{F}_{1/2}$ is the Fermi-Dirac integral of order 1/2, N_C and N_V are the effective densities of states in the conduction and valence bands, E_{F_n} and E_{F_p} are the quasi-Fermi energies for electrons and holes, and E_C and E_V are the conduction and valence band edges.

1.2.2 Dark current

A critical characteristic of IR detector is the current produced by a reverse biased photodiode in absence of any IR signal. This current is known as the **dark current** $J_{\text{dark}}(V)$,

where V is the applied voltage. In dark conditions, the Shockley ideal diode equation or the diode law

$$J_{\text{dark}}(V) = J_0 \left(e^{\frac{q\,V}{k_B T}} - 1 \right), \tag{1.8}$$

defines a saturation current density J_0, representing at first approximation (neglecting the drift component in quasi-neutral regions) a diffusion current of electrons into the *p*-type region and holes into the *n*-type region [8]:

$$\begin{aligned} J_0 \;=\; & \left(\frac{q\,D_p\,p_n}{L_{\text{diff},p}} \right) \frac{\left(\frac{v_1\,L_{\text{diff},p}}{D_p} \right) \cosh\left[\frac{x_n}{L_{\text{diff},p}} \right] + \sinh\left[\frac{x_n}{L_{\text{diff},p}} \right]}{\left(\frac{v_1\,L_{\text{diff},p}}{D_p} \right) \sinh\left[\frac{x_n}{L_{\text{diff},p}} \right] + \cosh\left[\frac{x_n}{L_{\text{diff},p}} \right]} + \\ & + \left(\frac{q\,D_n\,n_p}{L_{\text{diff},n}} \right) \frac{\left(\frac{v_2\,L_{\text{diff},n}}{D_n} \right) \cosh\left[\frac{t+d-x_n-w}{L_{\text{diff},n}} \right] + \sinh\left[\frac{t+d-x_n-w}{L_{\text{diff},n}} \right]}{\left(\frac{v_2\,L_{\text{diff},n}}{D_n} \right) \sinh\left[\frac{t+d-x_n-w}{L_{\text{diff},n}} \right] + \cosh\left[\frac{t+d-x_n-w}{L_{\text{diff},n}} \right]}. \end{aligned} \tag{1.9}$$

Here

$$w\,(V) = \sqrt{\frac{2\epsilon_s \left(\frac{k_B T}{q} \ln\left[\frac{N_A N_D}{n_i^2} \right] \pm V \right)}{q N_A N_D\,(N_A + N_D)}}, \tag{1.10}$$

N_A and N_D are respectively the acceptor and donor concentrations, p_n is the hole carrier concentration in *n*-side and n_p the vice-versa, $D_n = \mu_n k_B T/q$ and $D_p = \mu_p k_B T/q$ are the minority carrier diffusion coefficients according to the Einstein relations, $\mu_{n,p}$ are the electron and hole mobilities, $L_{\text{diff},n} = \sqrt{D_n\,\tau_n}$ and $L_{\text{diff},p} = \sqrt{D_p\,\tau_p}$ the minority carrier diffusion length, $\tau_{n,p}$ their recombination lifetimes (determined by the recombination mechanisms), $v_{1,2}$ the mean recombination velocity in the p- and n-doped regions, and the geometrical parameters x_n, w, t and d are defined in Fig. 1.4(a).

In the so-called "short" diode approximation, $L_{\text{diff}} \gg x_n$ and $L_{\text{diff}} \gg t + d - x_n - w$, and in quasi-neutral regions (in which the electric field $\vec{\mathcal{E}}$ is negligible), it is found [16]

$$J_0 \simeq \frac{q n_i^2 t}{n \tau_n} + \frac{q n_i^2 d}{p \tau_p} \tag{1.11}$$

in which $v_1 \approx t/\tau_n$, $v_2 \approx d/\tau_p$ have been assumed, and the two addenda respectively refer to the p- and n-doped regions.

The dark current sets a sort of background current, somehow a noise current. Any IR signal that produces in the detector a current due to photogenerated carriers (see Sec. 1.2.3) that is less than the noise produced by *the dark current plus the background* is washed out and cannot be detected. As a consequence, it is important to reduce dark current as much as possible, in order to obtain a detector ideally limited only by the background (BLIP, Background Limited Infrared Photodetector).

Regarding the mechanisms responsible for the dark current generation, we remind that minority carrier generation-recombination (GR) lifetimes in high-quality direct-bandgap semiconductors are limited primarily by the intrinsic radiative and Auger GR mechanisms within the quasi-neutral narrow-bandgap absorber, while in poorer quality material, extrinsic mechanisms like Shockley-Read-Hall (SRH) and trap-assisted tunneling (TAT) occurring within depleted regions tend to dominate the dark current. The first two mechanisms depend primarily by the band structure and doping concentration of the material, whereas SRH and TAT are determined by defects or impurities. Discarding TAT, that can also be described as an enhancement of SRH process [17], the minority carrier lifetime τ is given by

$$\tau = \left(\tau_{\text{Auger}}^{-1} + \tau_{\text{radiative}}^{-1} + \tau_{\text{SRH}}^{-1}\right)^{-1}. \tag{1.12}$$

In HgCdTe, Auger, radiative and SRH lifetime expressions are described and discussed in Sections 2.2.7, 2.2.8 and 2.2.9. Other important sources of dark current are the band-to-band-tunneling (BTBT) and the Impact Ionization (I.Ion.) process. Detectors fabricated with high quality material should be only intrinsically limited: this means that only intrinsic, unavoidable dark-current generation mechanisms should be at play. The Auger process is the classical example: it can be reduced in several ways, but not eliminated. On the contrary, the SRH and TAT processes, since they are connected to defects, in principle could be minimized, whereas there is no agreement about the importance of radiative processes, and some authors assume that a mechanism of photon-recycling assures that its net-effect is virtually null [18, Sec. 4.1.1], [19]. BTBT and I.Ion. should be at play only in case of high reverse voltage operating conditions, that are very unusual. In summary, ultimate detectors are often referred to as Auger-limited, exactly for this motivation.

In Chapter 3 these dark current generation mechanisms will be addressed and their expressions will be employed to reproduce experimental dark currents.

1.2.3 Photocurrent

In order to illustrate the concept of photocurrent and its relation with the illuminating optical flux, is is convenient to consider, for simplicity a 1D semiconductor with length L and uniform composition, whose complex refractive index is $\hat{n}(\lambda) = n_0(\lambda) + i\kappa(\lambda)$, since the principle of operation for 2D or three dimensional (3D) structures is similar. The absorption coefficient at a given wavelength λ is given by $\alpha(\lambda) = 4\pi\kappa(\lambda)/\lambda$, usually determined experimentally at a given temperature T and for a given semiconductor composition (in the case of HgCdTe, the x Cd molar fraction).

If P_0 is the optical power flux (W cm^{-2}) and R_λ is the reflection coefficient of the illuminated photodetector face, the photon flux (photons cm^{-2} s^{-1}) after the light travelled a distance z into the material is

$$\phi(z,\lambda) = (1 - R_\lambda)\frac{\lambda}{hc}P_0\, e^{-\alpha(\lambda)z} \tag{1.13}$$

where h and c are respectively the Planck's constant and the light velocity, the **absorbed photon density** A_0 (photons $\mathrm{cm}^{-3}\,\mathrm{s}^{-1}$) is given by

$$A_0(z,\lambda) = \phi(z,\lambda)\alpha(\lambda) = (1-R_\lambda)\frac{\lambda}{hc}P_0\, e^{-\alpha(\lambda)z}\alpha(\lambda). \tag{1.14}$$

The **optical generation rate**, that is the generated carrier per unit volume and time is $G(z,\lambda) = \eta A_0(z,\lambda)$, and η represents the elementary quantum efficiency, that is the number of electron-hole (e-h) couples produced by each absorbed photons, usually taken as unity for simplicity.

The **photocurrent** density J_{ph} is found integrating G on z,

$$J_{ph}(\lambda) = \int_0^L qG(z,\lambda)dz = (1-R_\lambda)\eta P_0\frac{q\lambda}{hc}\int_0^L e^{-\alpha(\lambda)z}\alpha(\lambda)dz. \tag{1.15}$$

where q is the elementary charge, and in the given units results expressed in $\mathrm{A\,cm}^{-2}$. Nothing of substantial changes if α is function of z, for example because the material changes along the device: in this case, $\alpha(\lambda)$ must be substituted by $\alpha(\lambda,z)$, and $\phi(z,\lambda)$ is given by

$$\phi(z,\lambda) = (1-R_\lambda)\frac{\lambda}{hc}P_0\, e^{-\int_0^z \alpha(\lambda,z')z'}dz' \tag{1.16}$$

It must be stressed that in realized devices the current density $J_{\text{meas.}}$ measured illuminating a photodiode is not the experimental photocurrent, neither in ideal conditions, because it also includes the dark current $J_{\text{dark,meas.}}$. Hence the experimental photocurrent to be compared with the theoretical one J_{ph} is the difference

$$J_{ph,\text{meas.}} = J_{\text{meas.}} - J_{\text{dark,meas.}}, \tag{1.17}$$

where $J_{\text{dark,meas.}}$ must be measured at dark, with the same applied voltage employed under illumination.

if N_e is the number of photogenerated electrons (numbers of electrons generatred per unit time) in the detector, the photocurrent I is given by

$$I = qN_e. \tag{1.18}$$

The number of photons entering the detector's illuminated area S per unit time is, if $R = 0$,

$$N_{ph} = \frac{P_0 S}{\frac{hc}{\lambda}} \tag{1.19}$$

and in the simplest and ideal case $R = 0$ and zero dark current, the current I must be proportional to N_{ph} according to

$$I = \eta g N_{ph} = \eta g\frac{P_0 S\lambda}{hc} \tag{1.20}$$

where g is the gain of the detector, that depends on the detector length, the absorption coefficient $\alpha(\lambda, z)$, and other detector's characteristics. Dividing by S, we obtain

$$J_{ph} = \eta g \frac{P_0 \lambda}{hc}\,, \tag{1.21}$$

and this result will be used in the subsequent sections, where some useful figures of merit are shortly recalled.

1.2.4 Detectors figures of merit

In order to compare detectors performances, several figures of merit have been defined, among which we can mention the *Responsivity*, the *Noise Equivalent Power* and the *Detectivity*.

The **Responsivity** $\mathcal{R}$ of an infrared detector is defined as the ratio of the root mean square (rms) value of the fundamental component of the electrical output signal of the detector to the rms value of the fundamental component of the input radiation power. If the output signal is the photocurrent density J_{ph} (A cm^{-2}) produced when the detector is illuminated by the monochromatic optical power flux P_0 (W cm^{-2}), the units of responsivity are Ampers per Watt (A/W), and R is given by

$$\mathcal{R} = \frac{J_{ph}}{P_0}. \tag{1.22}$$

Recalling Eq. 1.21, we can express the responsivity as

$$\mathcal{R} = \frac{\eta g \lambda}{hc}. \tag{1.23}$$

as given in Ref. [8, Ch. 9].

The **Noise Equivalent Power** NEP is the optical power plux P_0 incident on the detector that generates a signal output (e.g. a current density J_n) equal to the rms noise output. Stated another way, the NEP is the signal level that produce a signal-to-noise ratio (SNR) of 1. It can be written in terms of R as:

$$\text{NEP} = \frac{J_n}{\mathcal{R}} \tag{1.24}$$

and the unit of NEP is W.

A more directly applicable figure of merit, suitable for imaging applications like thermography, is the Noise-Equivalent Difference of Temperature $NE\Delta T$ which describes the noise of an image as a minimum resolvable temperature. Experimentally, the $NE\Delta T$

is measured by imaging a blackbody target of temperature T_t in front of a background of known temperature T_b. The difference between the target and background temperatures, normalized to the SNR of the detector system gives the $NE\Delta T$: as:

$$NE\Delta T = \frac{T_t - T_b}{\text{SNR}} \tag{1.25}$$

The **Detectivity** D is the reciprocal of NEP, $D = 1/\text{NEP}$. Since it was found that normally NEP is proportional to the square root of the detector area A, a normalized and area-independent detectivity D^* can be defined as

$$D^* = D\sqrt{A} = \frac{\sqrt{A}}{\text{NEP}} \tag{1.26}$$

and both NEP and D^* may be defined for a monochromatic radiation of for a given optical band, defining a spectral optical power flux and integrating over λ as given for the responsivity $\mathcal{R}$.

The **internal quantum efficiency** (IQE) can be defined as the ratio between the number of photogenerated electrons and the number of photons entering the detector's illuminated area S per unit time N_{ph},

$$\text{IQE} = \frac{N_e}{N_{\text{ph}}}\,. \tag{1.27}$$

The latter, if the illuminated detector's area reflectivity is R, is given by

$$N_{ph} = (1 - R)\frac{P_0 S\lambda}{hc}. \tag{1.28}$$

therefore, since $N_e = (I - I_{\text{dark}})/q$, the IQE can be expressed as

$$\text{IQE} = \frac{hc}{(1 - R)q\lambda}\frac{(I - I_{\text{dark}})}{P_0 S}\,. \tag{1.29}$$

In this definition it is implicit the assumption that all the absorbed photons entering the detector are converted into electron-hole couples. One can also define an *external quantum efficiency* EQE, in which all the impinging photons are considered. In this case the factor $1 - R$ in the denominator of Eq. 1.29 is not present.

1.2.5 Detectors limitations: the problem of cooling

A detector limitation states a maximum performance that can be obtained considering a certain feature. For example, it is obvious that apart form the radiation flux P_s coming from the IR source, a detector reveal also the radiation flux P_b coming from the background. Therefore the *signal*, in order to be revealed, must be greater than the noise

originating from the background, and the background flux is a limitation. Nevertheless, it is *not* an intrinsic limitation, because the background is related to the environment, not to the detector.

Supposing that the *signal* is stronger that the *background noise*, the latter is not a limitation. Nevertheless *thermal-generated* carriers may be in the order of *signal-generated* carriers, and this is the motivation that often requires to decrease the operating temperature well below room temperature.

To be more precise, in addition to the carriers generated by P_s and P_b, also the thermal generation contributes to the carrier density in the detector, that can be considered as *good* detector only if thermal-generated carrier density n_{th} is much lesser than the background-generated carrier density n_b: only in this case the limiting factor is still the background.

If the detector thickness is t and carrier lifetime is τ, n_b is roughly given by $n_b \approx \eta P_b \tau / t$. Therefore, a detector is *background limited* (BLIP, Background Limited Infrared Photodetector) if

$$\frac{\eta P_b \tau}{t} > n_{th}. \tag{1.30}$$

We can also say that the photon generation rate per unit area ηP_b must be greater than the thermal generation rate per unit area,

$$\eta P_b > \frac{n_{th} t}{\eta \tau}, \tag{1.31}$$

This condition defines a BLIP. The detector quantum efficiency η can be roughly expressed as $\eta \approx \alpha t$, where α is the absorption coefficient (of course, at a given wavelength). The BLIP condition becomes independent from t,

$$\eta P_b > \frac{n_{th}}{\alpha \tau}, \tag{1.32}$$

and defines the minimum background flux P_b (or the maximum thermal generation rate per unit area $G_{th} = n_{th}/(\alpha\tau)$) required to obtain a BLIP. This definition for G_{th} can be used to predict the ultimate performance of any IR material, and to compare the relative performance of different IR materials, as a function of temperature. The only requirement is a knowledge of the dependence of n_{th} and τ on temperature. It must also be noticed that the thermal-generated dark current density (that is only one of the sources of dark current) is given by

$$J_{d,th} = G_{th} q \,. \tag{1.33}$$

Ref. [20] discusses in depth these and other intrinsic limitations of photodetectors, but among all possible limitations, the most severe one is precisely given by the thermal generation rate. Cryogenic cooling of IR detectors has always been the burden of sensitive IR systems, and an universal goal for IR photon detection systems is to increase their operating temperature without sacrificing performance. Auger generation typically dominates the dark current at elevated temperatures, and standard $p-n$ junction photodiodes

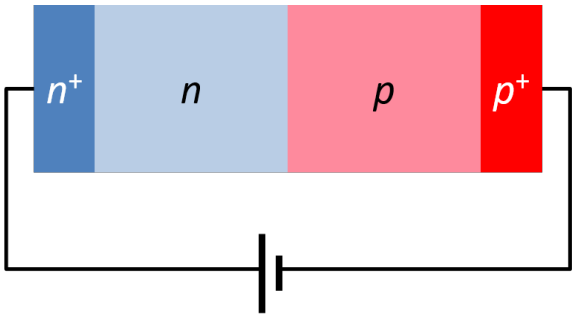

Figure 1.8. Schematic of HOT detector.

therefore become very noisy when operated near room temperature. In addition, the detector itself, its package, the circuitry, etc. are sources of thermal IR radiation, especially in LWIR, that contribute to the dark current. Particularly in the past, most FPAs have operated at a temperature of 77 K or below, in order to minimize thermal generation and the resulting the dark current. At present days novel device structures based on HgCdTe are object of intense study and development to achieve high sensitivity infrared imaging at high temperatures. In fact, there is a considerable system advantage if the operating temperature of the FPA can be increased: the resulting cooler package into which the FPA is integrated will be considerably lighter, smaller, and thus cheaper (also the power required to cool the FPA can be decreased).

A detector concept that enables this elevated temperature operation is the high operating temperature (HOT) detector [21]. The HOT detector is simply a reverse-biased photodiode with minority and majority carrier contacts, illustrated in Fig. 1.8. The absorber volume, which is less than a diffusion volume, is in non-equilibrium due to the reverse bias, and the intrinsic thermally generated minority carriers are fully extracted through the minority carrier contacts. HOT narrow-gap semiconductor devices operate in nonequilibrium mode such that the carrier densities are held below their equilibrium, near-intrinsic, levels. The reduction in majority and minority carrier concentrations in the active volume of the HOT device results in a significant reduction of the Auger generation processes, responsible for a large part of the dark current. In principle, such technologies have the potential for BLIP performance at room temperature [22]. In view of reducing the cryogenic cooling, **barrier-layer photodetectors** were first proposed by A. M. White [24], who postulated an n-type heterostructure with a narrow gap absorber region coupled to a thin wide bandgap layer, followed by a narrow bandgap n-type contact region [23, 25] (n-B-n detector). The barrier layer selectively blocks majority carriers (dark current) and collects photogenerated minority carriers (photocurrent). The concept assumes almost zero valence band offset approximation throughout the heterostructure, allowing flow of only minority carriers in a photoconductor. Little or no valence band offset was difficult to

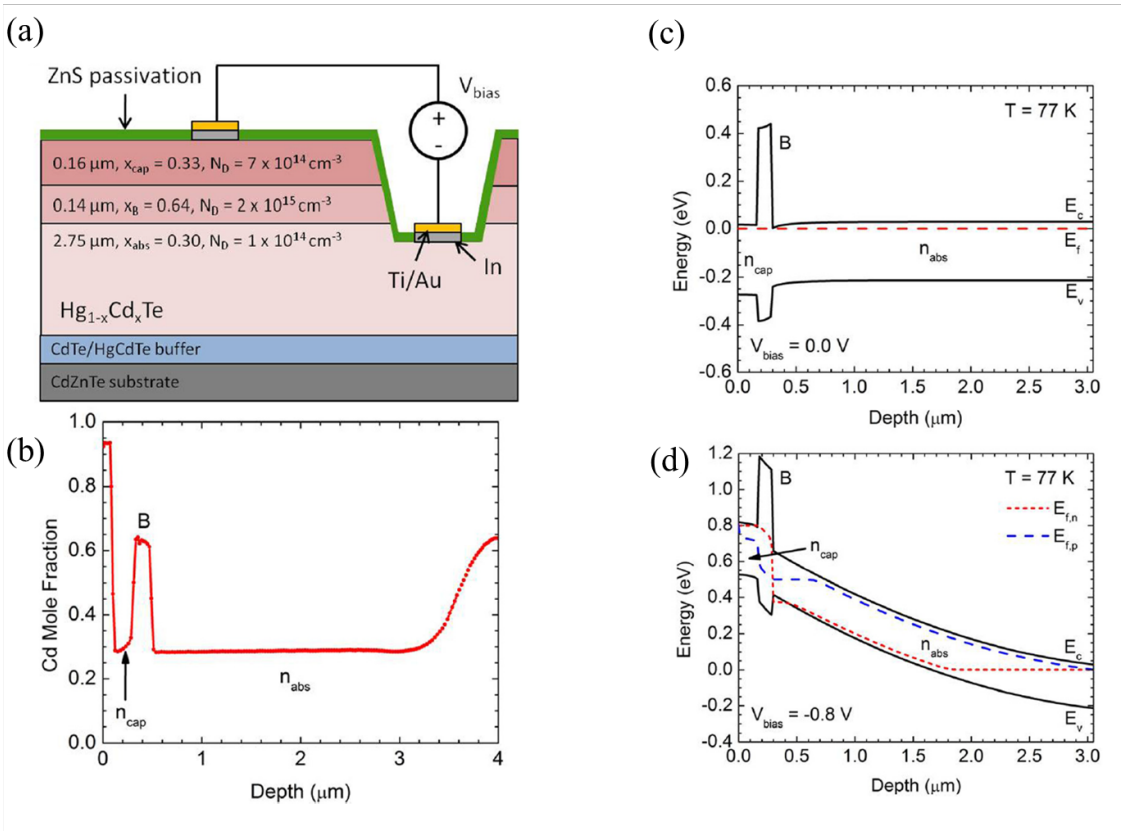

Figure 1.9. n-B-n junction photodetector: (a) overall view of a possible implementation, with dopant concentrations; (b) composition profile; (c) band energy diagrams at equilibrium and (d) under reverse bias. The figure is from [23].

realize using standard infrared detector materials such as InSb and HgCdTe. In general, unipolar barriers are used to implement the barrier detector architecture for increasing the collection efficiency of photogenerated carriers and reducing dark current generation without inhibiting photocurrent flow. A possible implementation making use of HgCdTe, with composition, doping and band diagram profiles is shown in Fig. 1.9.

1.3 TCAD-based simulation methods

Technology Computer Aided Design (TCAD) represents the modern workbench for scientists who want to investigate advanced device performances. Nevertheless modeling is a complex issue both in mathematical and physical terms, and the solution of the *electric transport problem* represented by a set of PDEs (partial differential equations) is often impossible to be found without approximations.

Many commercial and powerful simulators are available, each with specific advantages and drawbacks. Whatever the simulator one can choose, the numerical solution of the PDEs system is always evaluated within a discretized representation of the model itself, through one of the several techniques available. The spatial approximation of the PDEs system by means of a discretization scheme yields, in turn, a discrete system of algebraic equations, often linearized under specific physical simplifications.

In this work, the Synopsys® TCAD Sentaurus suite [1] has been extensively exploited, in particular its implementation of the drift-diffusion model.

1.3.1 Solving the electrical problem

In this tool, the involved partial differential equations are discretized according to the *Finite Boxes* (FB) method [26], a numerical method belonging to the class of the finite differences method: the device is divided into great number (often some thousands) of *boxes*, constituting the 1D, 2D or 3D grid, called the *mesh*. This allows a locally averaged calculation of the unknowns onto the nodes of the calculation grid. Through this procedure, the set of coupled partial differential equations composing the model become ordinary linear algebraic equations, reducing much the complexity of the mathematical problem.

Just to make an example, the Poisson equation Eq. 1.2 is converted in [27]

$$\begin{aligned} \frac{q}{\epsilon}\left(N_{D,i} - N_{A,i} + p_i - n_i\right) S_i &= \sum_j l_{ij} \left\langle \mathcal{E}^{\perp} \right\rangle_{ij} \\ &\approx \sum_j l_{ij} \frac{\phi_i - \phi_j}{d_{ij}} \end{aligned} \tag{1.34}$$

in which d_{ij} is the distance between nodes i and j, l_{ij} is a distance defined in Fig. 1.10 and where the second line concerns the first-order finite difference approximation of the electric field $\mathcal{E}^{\perp}$ normal to the considered box surface (ϕ is the electrostatic potential), averaged around the box sides. Similar conversions hold for all other equations. In Fig. 1.10 the discretization scheme is shown and the involved geometrical quantities are defined.

1.3.2 Solving the optical problem

The j-th mesh element, of volume V_j, absorbs an amount of radiation $A_{\text{opt},j}$, where the spatial distribution $A_{\text{opt}}(x, y, z) = A_{\text{opt},j}/V_j$ is computed according to the selected optical model (for example, the ray tracing or the finite difference time domain methods) and x, y, z are the coordinates of the j-th element barycenter. In its turn, $A_{\text{opt}}(x, y, z)$ enters the continuity equations for the electron and hole current densities $J_{n,p}$ (where $R_{n,p} - G_{n,p}$

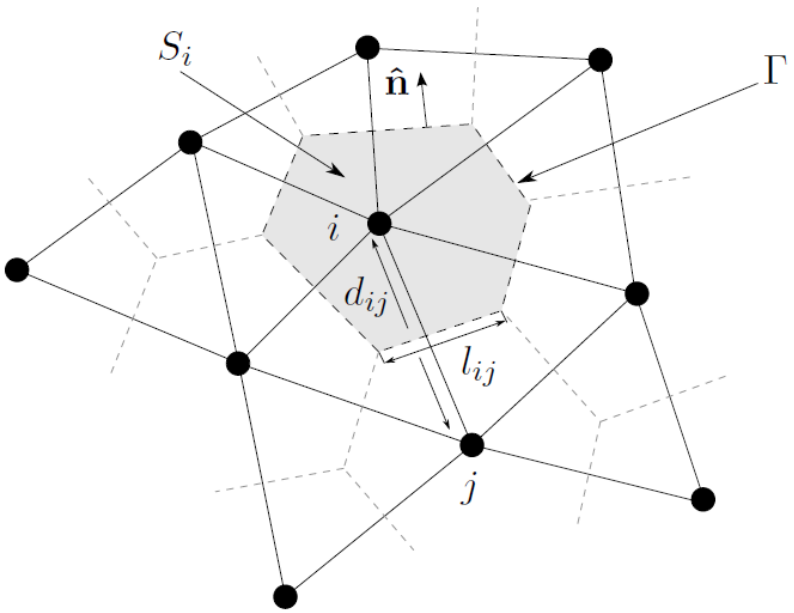

Figure 1.10. Schematic representation of a "box-method" grid. The dashed shapes are the edges of the finite boxes and are calculated through the bisectors of the triangle sides converging in the corresponding *node*. The grey area represents the finite box surrounding the i-th node.

is the net recombination rate in absence of carrier photogeneration and q is the elementary charge)

$$\begin{aligned} \nabla \cdot J_n &= q(R_n - G_n - G_{\text{opt}}) + q\frac{\partial n}{\partial t} \\ -\nabla \cdot J_p &= q(R_p - G_p - G_{\text{opt}}) + q\frac{\partial p}{\partial t} \end{aligned} \tag{1.35}$$

through the optical generation rate $G_{\text{opt}} = \eta A_{\text{opt}}$ due to interband optical absorption, which depends also on the model for the quantum yield η, defined as the fraction of absorbed photons which are converted to photogenerated electron-hole pairs.

Concerning the solution of the *optical problem* (e.g. for the evaluation of photocurrent spectra), the central point is the evaluation of A_{opt}. We adopted for the largest part of this work the so-called *ray tracing* method: despite it ignores the wavelike nature of the light, many important and reliable simulations can be obtained. Instead, when effects induced by light interference and diffraction effects are at play, the direct solution of the Maxwell's equations must be considered, e.g. within the Finite Difference Time Domain (FDTD) method.

In Chapter 4 we will present a comparison between simulations obtained by ray tracing and by the solution of the Maxwell's equation in the FDTD scheme: for clarity, this method will be described only at that point, since in Chapters 2-3 all simulations have been done according to the ray tracing method.

According to the **ray tracing** method, rays are used to represent the propagation of electromagnetic waves according to geometrical optics. Fresnel's and Snell's laws of reflection and refraction determine the paths travelled by rays as they interact with different materials. Usually, a large number of rays is traced within an optical system in order to simulate its behavior; in the present simulations, we considered ≈ 2500 rays uniformly spaced in the xy plane of the detector's illuminated face, entering the pixel with initial wavevector parallel to the z-axis. Each ray i transmitted into the j-th mesh element corresponds to an optical rate intensity $P_{j,i}$ (number of photons entering the element per second). The ensuing absorbed photon density in V_j is

$$A_{\text{opt},j,i} = \frac{P_{j,i}}{V_j}\left(1 - e^{-\alpha_j L_{j,i}}\right) \tag{1.36}$$

where $L_{j,i}$ is the path length travelled by ray i into the j-th element, α_j is the absorption coefficient α averaged on the j-th element, and α is related to the imaginary part κ of the complex refractive index $\hat{n} = n + i\kappa$ as $\alpha = 4\pi\kappa/\lambda$. Summation over all the ray histories produces the mesh-discretized distribution $A_{\text{opt},j}(x, y, z)$, and hence the optical generation rate $G_{\text{opt}}(x, y, z)$ entering the continuity equations (1.35). This method manages all the internal back-reflections, but without keeping into account interference effects.

1.3.3 Simulation steps in TCAD Sentaurus

The simulation itself consists of several steps, or logical "blocks":

- a first tool (Sentaurus Device Editor, SDE) reads a **first user-defined script**, building-up the 1D, 2D or 3D device geometry, specifying materials, doping and composition profiles "region-wise", electrical contacts and meshing criteria; the tool generates the input files necessary to the subsequent logical block (i.e. the so-called TDR boundary file and the mesh command file), written in Synopsys specific formats;
- a second tool (Sentaurus Mesh, SNMESH) reads the two files produced by the previous block, generating the TDR grid and data file for the device structure; this file contains the material identifier, the doping and composition profiles and all other information needed at each node of the computational grid. This is the input file for the simulation engine, the Sentaurus Device (SDEVICE) tool;
- the SDEVICE tool solves the numerical problem, on the basis of a **second user-defined script**: it contains specific instruction to locate and read a) the TDR file

produced by SNMESH, b) the parameter file containing the materials library. Finally, c) it defines what are the models and the simulation steps to consider. For example, a set of keywords and user-defined procedures specify what is the carrier statistic to adopt (Boltzmann or Fermi), what is the semiconductor model to consider (drift-diffusion, hydrodynamic model, Monte Carlo, etc.), and what are the simulation steps to follow (for example a bias voltage slow ramping).

As already specified, the drift-diffusion model was considered, selecting the Fermi statistics and considering incomplete ionization of dopants (having defined in the parameter file the donor and acceptor ionization energy). Just to make an example, the simulation steps usually followed to obtain a current voltage $I(V)$ characteristic curve in dark conditions (dark current) can be:

- find the solution at equilibrium;
- drive slowly the device out of equilibrium, increasing the bias voltage;
- save the complete solution (carrier density, electric field, band energies, etc.) at one or more values of bias voltage, and the characteristic $I(V)$.

These steps may follow instructions to describe the illumination parameters (the illuminating window, the radiation wavelength, the wavevector direction, the optical power density, the polarization, etc.) and file names that will contain e.g. the photocurrent spectrum, the $I(V)$, the photogenerated carrier profile inside the detector, etc.

All these and many other simulations are possible only if a complete and accurate material parameter library is available, consisting in a **user-defined material(s) parameter file(s)** to be read by the SDEVICE block. Regarding the HgCdTe alloy, the material library was not present in the Sentaurus suite and the initial deal of effort put in this thesis consisted in the development of a custom software library for HgCdTe material properties, function of composition and lattice temperature, exploiting state-of-the-art experimental and theoretical information about HgCdTe. In Sec. 2.2 details are given about its definition.

1.3.4 Electric and optical boundary conditions

In order to solve the carrier transport equations, it is necessary to specify the appropriate boundary conditions at the edges of the device. For the simulation of single pixel, all the electric contacts will be assumed to be Ohmic, with a condition of charge neutrality at equilibrium given by

$$n_0 - p_0 = N_D^+ - N_A^- \tag{1.37}$$

$$n_0 p_0 = n_i^2 \tag{1.38}$$

where n_0 and p_0 are the electron and hole equilibrium concentrations. By default, the conditions $n = n_0$, $p = p_0$ are applied at the ohmic contacts so that the excess carriers are removed at the contact interface. All the other boundaries are treated with ideal Neumann boundary conditions

$$\vec{\mathcal{E}} \cdot \hat{n} = 0 \tag{1.39}$$

$$\vec{J_n} \cdot \hat{n} = 0 \tag{1.40}$$

$$\vec{J_p} \cdot \hat{n} = 0 \tag{1.41}$$

where $\vec{\mathcal{E}}$ is the electric field and $\hat{n}$ is a versor normal to the boundary.

Optical boundary conditions consist in specifying the reflectivity of the computational box boundary, for which a good choice can be to simulate the detector as posed into an infinitely extended medium, imposing $R = 0$ and $T = 1$ (where R and T are the reflection and transmission coefficients) for all the box boundaries. The TCAD Sentaurus simulator imposes by default these conditions, and they will be left unchanged. The detector is treated as in vacuum, and Fresnel reflection/refraction laws apply; nevertheless, it is also possible to specify a fixed value for the reflectivity of a particular face of the detector (e.g. $R = 0$ for the illuminated face, if a perfect anti-reflection coating is supposed to be deposited on it).

Chapter 2

The material system HgCdTe

Since first synthesized, $Hg_{1-x}Cd_xTe$, or simply MCT has been used in the development of photodetectors in the entire infrared spectrum. Its unique position as a quasi-ideal IR detector material system is based on three key features [8]:

- composition-dependent tailorable energy gap over the entire wavelength range of interest, $\lambda = 1 \div 30\,\mu$m,
- large optical coefficients, enabling high quantum efficiency,
- favorable inherent recombination mechanisms, leading to long carrier lifetime and high operating temperature.

These properties are a direct consequence of the electronic structure of this semiconductor. HgCdTe has a zinc-blende structure with two interpenetrating face-centered cubic lattices offset by $(1/4, 1/4, 1/4)a_0$ in the primitive cell (being a_0 the lattice constant). The mole fraction x of Cd in the $Hg_{1-x}Cd_xTe$ alloy (the alloy composition) can be chosen so as to tune the optical absorption of the material to the desired infrared wavelength. CdTe is a semiconductor with a bandgap of approximately 1.5 eV at room temperature. HgTe is a semimetal, hence its bandgap energy is zero. An alloy obtained from these two compounds allows to obtain any bandgap between 0 and 1.5 eV. This fundamental property is illustrated in Fig. 2.1, where the energy gap E_g, the *cut-off* wavelength (here defined as $\lambda_c = 1.24/E_g$, energy is in eV, λ_c in μm) and the lattice constant are reported as functions of the Cd mole fraction. It is well evident that HgCdTe – at least as far as the λ_c tunability and the invariance of the lattice constant with x are concerned – is a compound well suitable for IR detection. Other properties more involving the potentially very low dark current will be analyzed in the following chapters of this thesis, but it can be here anticipated that it is very difficult to find a real competitor of this alloy in the market of the materials for IR detectors industry (see a discussion about this point in Ref. [28]).

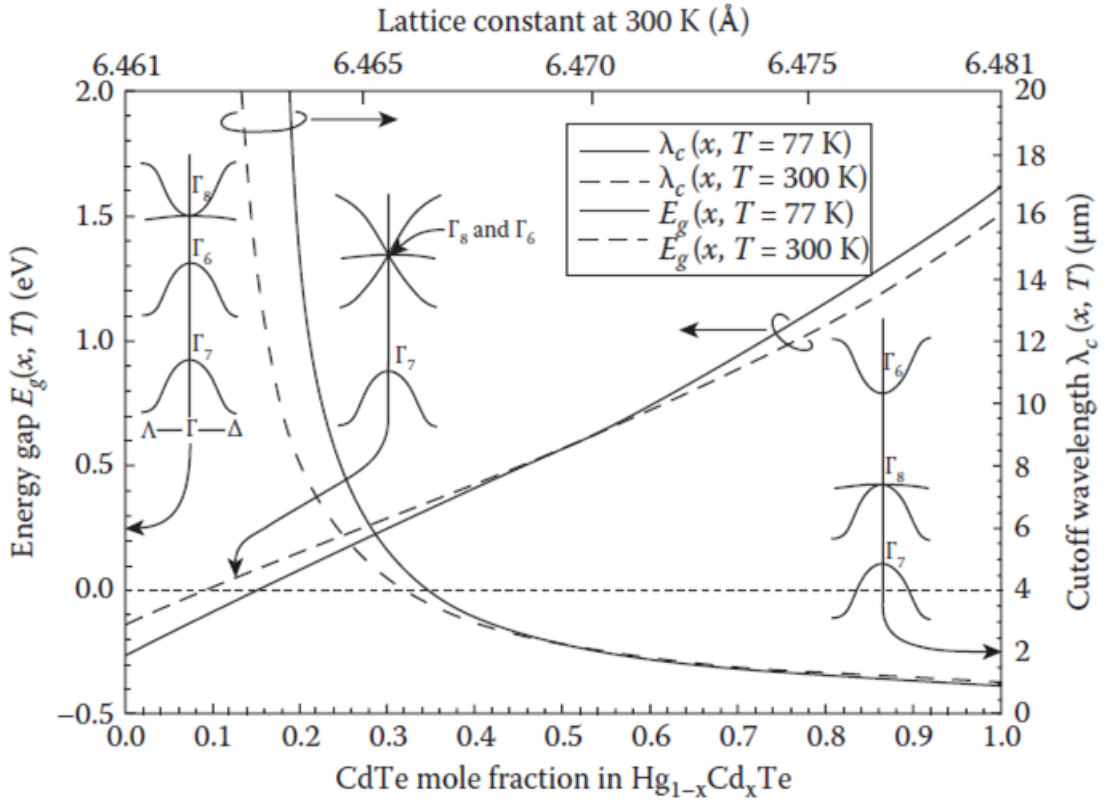

Figure 2.1. The bandgap structure of $Hg_{1-x}Cd_xTe$ near the Γ-point for three different values of the forbidden energy gap. The energy bandgap is defined at the difference between the Γ_6 and Γ_8 band extrema at $\Gamma = 0$ (figure from [8, Ch. 14]).

Additional specific advantages of HgCdTe are the ability to obtain both low and high carrier concentrations, high mobility of electrons, and low dielectric constant. The extremely small change in lattice constant with composition makes it possible to grow high-quality layered and graded-gap structures. As a result, HgCdTe can be used for detectors operated in various modes, as photoconductors, photodiodes, or metal-insulator-semiconductor detectors.

2.1 Crystal growth techniques

Historically, HgCdTe crystal growth has been a major problem mainly because a relatively high Hg pressure is present during growth, which makes it difficult to control the stoichiometry and composition of the grown material. The wide separation between the

liquidus and solidus, leading to marked segregation between CdTe and HgTe, was instrumental in slowing the development of all the bulk growth techniques to this system. In addition to solidus-liquidus separation, high Hg partial pressures are also influential both during growth and postgrowth heat treatments.

Epitaxial techniques (see e.g. [14] for an exhaustive historical review of main techniques applied to obtain high quality HgCdTe epitaxial layers) offer the possibility of growing large area epilayers and fabrication of sophisticated device structures with good lateral homogeneity and abrupt and complex composition and doping profiles, which can be configured to improve the performance of photodetectors. Among the various epitaxial techniques, liquid phase epitaxy (LPE) is the most technologically mature method. LPE is a single crystal growth process in which growth from a cooling solution occurs onto a substrate. Another technique, vapor phase epitaxial growth of HgCdTe is typically carried out by nonequilibrium methods which also apply to metal-organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), and their derivatives. The great potential benefit of MBE and MOCVD over equilibrium methods is the ability to modify the growth conditions dynamically during growth to tailor band gaps, add and remove dopants, prepare surfaces and interfaces, add passivations, perform anneals, and even grow on selected areas of a substrate. The growth control is exercised with great precision to obtain basic material properties comparable to those routinely obtained from equilibrium growth.

Epitaxial growth of HgCdTe layers requires a suitable substrate. CdTe was used initially, since it was available from commercial sources in reasonably large sizes. The main drawback to CdTe is that it has a few percent lattice mismatch with LWIR and MWIR HgCdTe. By the mid-1980s it was demonstrated that the addition of a few percent of ZnTe to CdTe (typically 4%) could create a lattice matched substrate.

2.2 HgCdTe properties: building a software material library

3D physical device simulation capabilities are well established and documented for the most widely used semiconductors such as silicon. While the set of equations in the numerical model are the same for any material, HgCdTe poses an additional layer of difficulty due to the strong molar fraction and temperature dependance of the material properties. In addition, most of the commercial software packages include material models for the description of the electrical properties that are not compatible with HgCdTe and, at first sight, make them not suited for HgCdTe.

Part of the initial deal of effort put in this thesis consisted in the development of a custom software library for HgCdTe material properties, function of composition and lattice temperature, exploiting state-of-the-art experimental and theoretical information

about HgCdTe. In fact, it is evident that an accurate description of the physical material parameters is essential for an accurate prediction of the real device performance. Hence, HgCdTe properties must be carefully described both in terms of the operating temperature T and Cd mole fraction x and provided as an input for the software simulation tool.

As far as a mole fraction dependent material is concerned, as for $Hg_{1-x}Cd_xTe$, it is necessary to build an ASCII file for each value of T under consideration, with the format expected by Sentaurus SDEVICE tool. In particular it must contain, for each considered material property having a functional expression $f(x)$, a set of third-order polynomial coefficients interpolating $f(x)$ onto the interval $x = [0, 1]$. It is evident that, if $f(x)$ has a complicated, highly nonlinear behavior, a third-order polynomial cannot interpolate $f(x)$ with the desired accuracy. In this case, multi-interval may be provided, with a different polynomial coefficient set for each subinterval.

For this scope, using a Matlab-based home-made set of software scripts, the material properties have been represented in the required format, and the ASCII file was named "HgCdTe_Txx_par.par", where "xx" stands for the operating temperature. In the user-defined command file for the Sentaurus simulator block SDEVICE, each employed material has been given a label that must point to a library file "***_par.par". Each point of the mesh belongs to a given material, and SDEVICE assigns to them the correct material properties described in the library files. If a given material already exists in the default Sentaurus library, the user is not required to specify the material file name. On the contrary, since HgCdTe does not exist in the default Sentaurus library, the user must link the library with statements like:

```
Material = "HgCdTe" {
   #if @T@ == 77
   #include  "HgCdTe_77_par.par"
   #endif
}
```

In this example, a molar-fraction-dependent HgCdTe library designed for T = 77 K is assigned to each node of the mesh belonging to the material HgCdTe.

SDEVICE solves the linear system arising in the drift-diffusion model and computes the desired quantities.

Differently from all the other properties (energy gap, effective masses, mobility, recombination coefficients, etc.), the TCAD Sentaurus syntax does not allow to employ this method for the absorption coefficient, if the material mole fraction varies along the device (e.g. if a compositional grading is present). Instead, it is possible to write a C++ routine, linked at run-time to Sentaurus Device, that allows in addition more degrees of freedom with respect to the other method. Therefore this is the way we followed to implement the complex refractive index formulas in Sentaurus.

This Section provides the values or the functional expressions implemented in the simulation code for all required material parameters of $Hg_{1-x}Cd_xTe$, along with relevant

literature references supporting them. Unless otherwise stated, in the present Section, apart from the variables x and T already defined,

- energies are given in eV
- recombination/generation times are given in seconds
- electron and hole effective masses are given in free electron mass units
- mobilities μ are in $\mathrm{cm}^2/\mathrm{V/s}$.

In Appendix A the custom interface [1] to *Sentaurus* is described, with details about the translation of each functional form representing a given HgCdTe property into the required format.

2.2.1 Energy gap

The band gap E_g is the difference between the bottom of the conduction band and the top of the valence band. In a $p-n$ photodetector, photons with energy greater than E_g, incident on one of the surfaces of the device, may create electron–hole pairs in the material on either side of the junction. By diffusion, the electrons and holes generated within a diffusion length from the junction reach the space–charge region. As explained in Sec. 1.2, electron–hole pairs are separated by the strong electric field and minority carriers are readily accelerated to become majority carriers on the other side: the generated photocurrent is revealed, constituting the signal. Hence, the value of x determines the maximum value of incident radiation λ_c that can be revealed. Its measure is often the only way to obtain and tune the alloy composition, that is the x parameter; for this reason, an accurate formula must be chosen relating E_g, x and T. In literature many formulations have been adopted, aiming to fit as close as possible the experimental results and/or the results coming from fundamental models.

The adopted energy gap formula, proposed by Seiler and Laurenti [29], is both composition- and temperature-dependent:

$$E_g(x,T) = E_{g_1}(x,T) + E_{g_2}(x,T) \tag{2.1}$$

$$E_{g_1}(x,T) = -0.302 + 1.93x - 0.81x^2 + 0.832x^3 \tag{2.2}$$

$$E_{g_2}(x,T) = 5.35\times10^{-4}\left(\frac{T^3-1822}{T^2-255.2}\right)(1-2x) \tag{2.3}$$

A plot of E_g as a function of x according to (2.1) is presented in Fig. 2.2. This formula has the advantage to well reproduce the nonlinear $T-$dependence below 100 K, differently from e.g. the Hansen's one [30], one of the more widely used [1].

[1] The review of Capper [31] presents several formulations for E_g: in addition to the two cited in the present document, he gives also one by Legros and Triboulet [32], although valid only for 0.7<x<1 and still linear in T, and another by Tong [33], a variant of Seiler's formula that could also be considered.

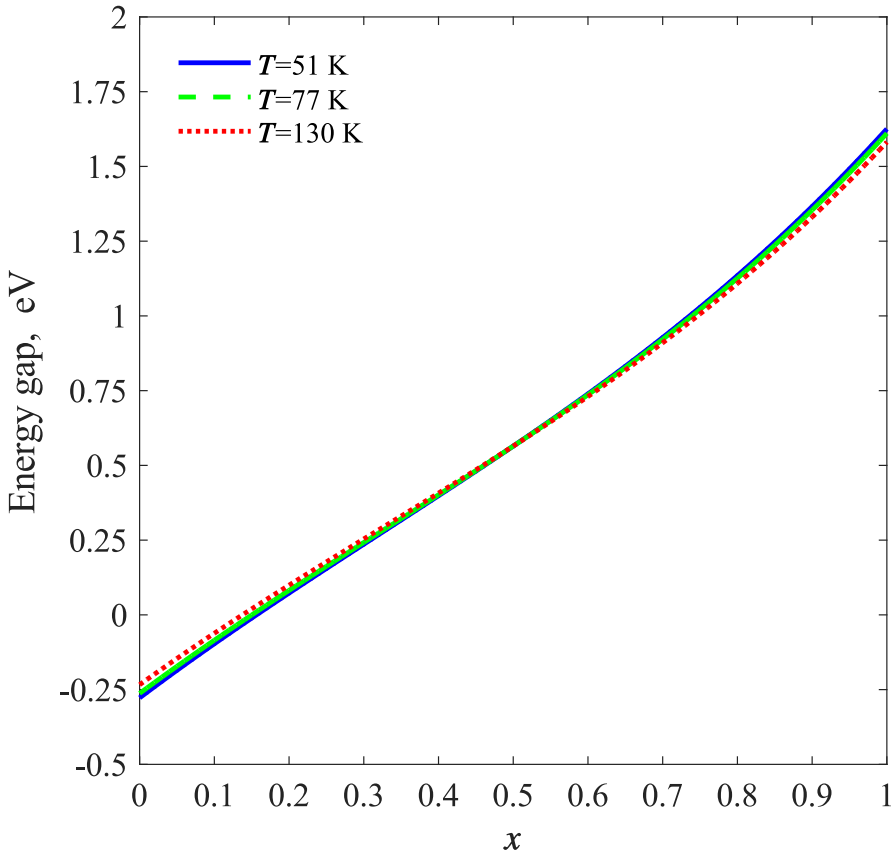

Figure 2.2. Energy gap as a function of Cd molar fraction x.

2.2.2 Electron and hole effective mass

Optical and transport properties are evaluated in the effective mass approximation, considering parabolic bands. The electron m_e and light hole m_{lh} effective masses in the narrow-gap mercury compounds are close and they can be established according to the Kane band model. Following Rogalski [8], Weiler's expression [34], valid for them both [2]. The result is reported in Fig. 2.3:

[2]Other authors, e.g. Ref. [35], adopt the same formula was employed, but with their own E_g and using a parameter E_p variable with x. Nevertheless, the value of this parameter does not affect much m_e and, using the Weiler's value for it, an uncertainty of 9% includes all the values used in literature. Instead, the E_g variation with x and T is the most critical in determining the m_e behavior [31].

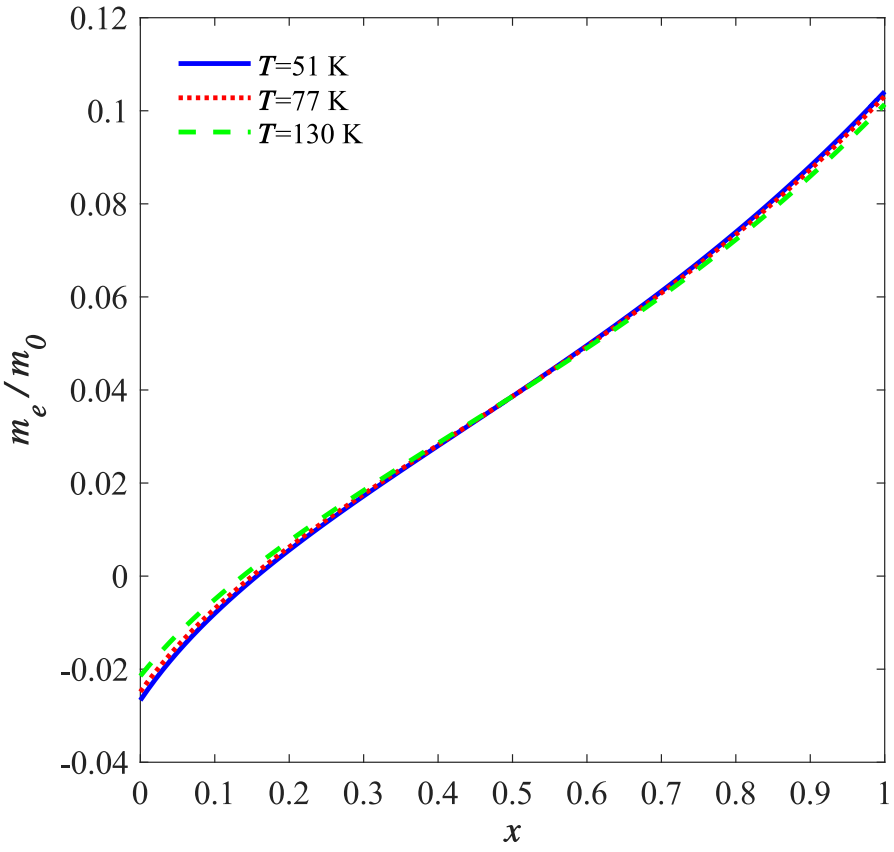

Figure 2.3. Electron effective mass as a function of Cd molar fraction x.

$$\frac{m_{e,lh}(x,T)}{m_0} = \frac{1}{1+2F+\frac{1}{3}E_p\left(\frac{2}{E_g(x,T)}+\frac{1}{E_g(x,T)+\delta}\right)} \tag{2.4}$$

where m_0 is the free electron mass and other parameters are defined as:

$$E_p = 19\,\text{eV}, \qquad \delta = 1\,\text{eV}, \qquad F = -0.8.$$

The heavy hole effective mass is high and Rogalsky [8] proposes the frequently used value of it

$$m_{hh} = 0.55 \tag{2.5}$$

for every composition and temperature, since measured values range between 0.3–0.7 m_0.

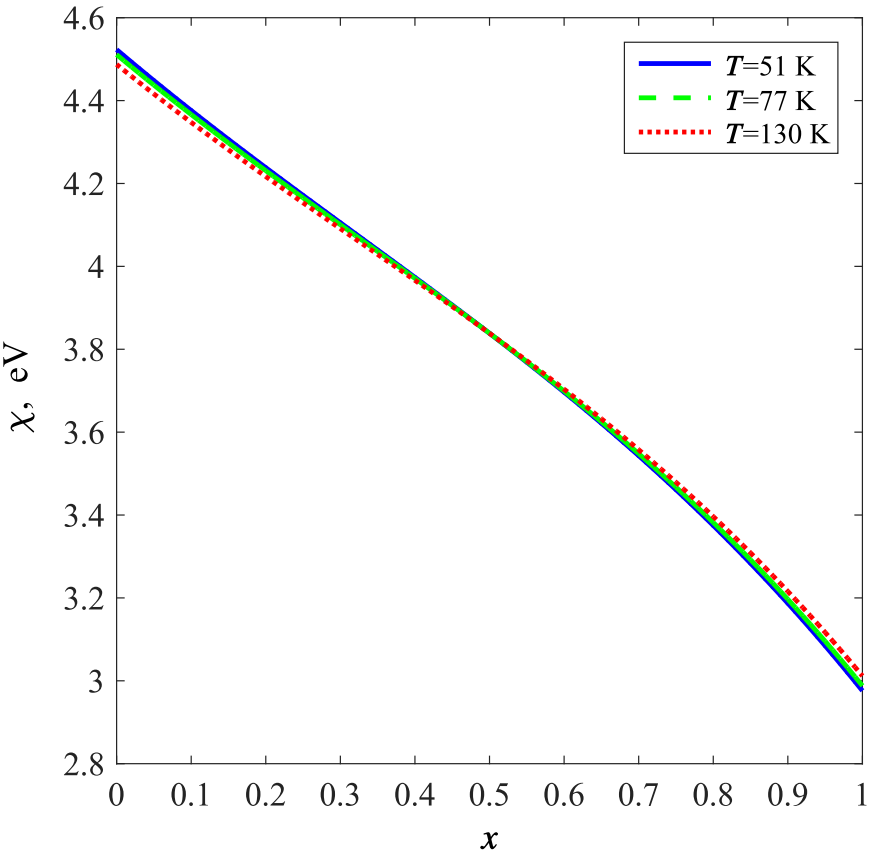

Figure 2.4. Electron affinity as a function of Cd molar fraction x.

2.2.3 Electron affinity

The electron affinity χ is the difference between the vacuum level and the bottom of the conduction band. Affinity is very important in heterostructures, because it affects how the band offset splits between conduction and valence band.

Regarding the affinity χ, it has been adopted the model of Ref. [36], defined by the authors through the investigation of band offsets for several graded and abrupt heterojunctions:

$$\chi(x,T) = 4.23 - 0.813\,(E_g(x,T) - 0.083) \tag{2.6}$$

According to (2.6), most (81%) of the energy gap discontinuity at a heterojunction is located in the conduction band. In Fig. 2.4 its plot as a function of x is reported.

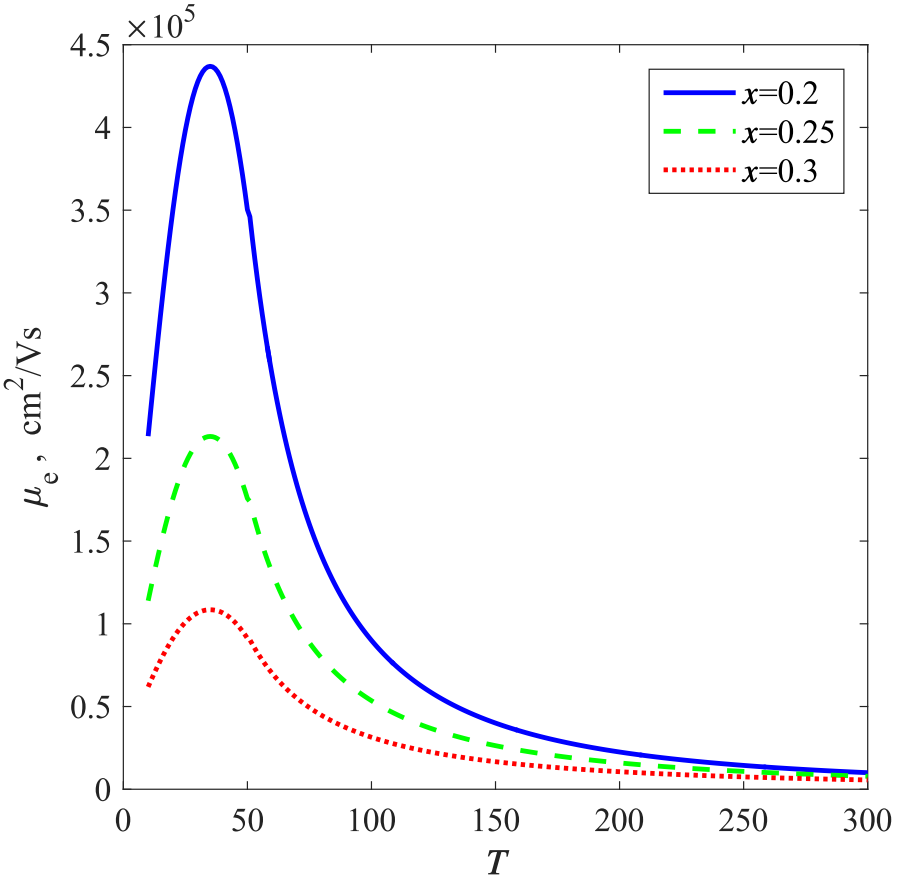

Figure 2.5. Electron mobility as a function of temperature.

2.2.4 Electron and hole mobility

Due to the small electron effective mass, the electron mobility in HgCdTe is remarkably high, while heavy–hole mobility is two orders of magnitude lower. A number of scattering mechanisms dominates electron mobility (see [8] and references therein). The x–dependence of the mobility (see Fig. 2.5) results primarily from the x–dependence of the energy gap, while the temperature dependence derives from the competition among different temperature-dependent scattering mechanisms.

According to [8] and [37], electron mobility is well approximated by the model [3]:

$$\mu_e(x,T) = 9 \times 10^8 \left(\frac{0.2}{x}\right)^{7.5} \frac{1}{Z^{2\left(\frac{0.2}{x}\right)^{0.6}}} \qquad \mathrm{cm}^2/\,\mathrm{V/s} \qquad (2.7)$$

[3] The Capper review [31] cites a work of Higgins [38], where he gives an empirical formula, valid for $0.18 < x < 0.25$, for the variation of μ_e with x, but only for the fixed temperature of 300 K.

where

$$Z = \begin{cases} T & \text{if } T > 50, \\ \frac{1.18\times10^5}{2600-|T-35|^{2.07}} & \text{if } T \leq 50. \end{cases} \tag{2.8}$$

This expression derives from experimental measurements for the interval $0.2 \leq x \leq 0.6$, which includes all compositions of practical interest.

The same references report a widely used expression for the much lower heavy–hole mobility:

$$\mu_{hh}(x,T) = \frac{\mu_e(x,T)}{100}. \tag{2.9}$$

2.2.5 Low– and high–frequency dielectric constant

The high frequency dielectric constant, ϵ_∞, and the static (low–frequency) constant, ϵ_s, are usually derived from reflectivity data in evaluating the real and imaginary parts of ϵ. The dielectric constants are not linear functions of x, and their temperature dependence is usually neglected.

Among others, the following widely used expressions were chosen (plotted in Fig. 2.6 as functions of x) [11, 36]:

$$\epsilon_s(x) = 20.5 - 15.5x + 5.7x^2 \tag{2.10}$$

$$\epsilon_\infty(x) = 15.2 - 13.7x + 6.4x^2 \tag{2.11}$$

whereas slightly different coefficients can be found in Refs. [8] and [39].

2.2.6 Intrinsic carrier density

The intrinsic carrier density n_i of $Hg_{1-x}Cd_xTe$ enters in the Auger recombination rate (whose functional x and T dependence is discussed later in this document). The expression given in Ref. [40] has been chosen, because it has been derived consistently with the energy gap (see Ref. [29]). As discussed in Ref. [40], n_i is not directly measured: previous measurements showed considerably scatter, due to the much lower mobility for holes than for electrons. However n_i can be calculated from bands parameters, but among them, the energy gap E_g affects very much n_i and, consequently, it should be determined with great accuracy, especially for low x, where E_g assumes low values. The assumed expression for E_g, determined from magnetoabsorption experimental data, allowed Lowney [40] to accurately determine n_i, applying a Kane 3–bands model for non–parabolicity effects

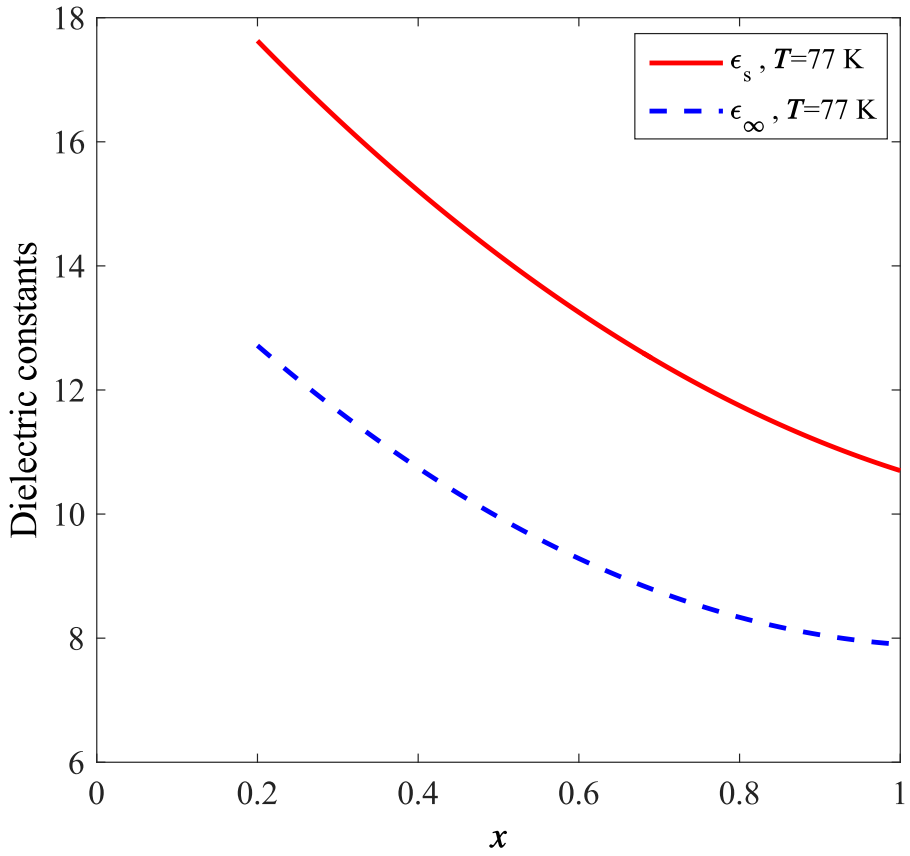

Figure 2.6. Dielectric constants as a function of Cd molar fraction x, at T=77 K.

and full Fermi statistics. The obtained expression is [4]:

$$n_i(x,T) = 10^{14}\, n_{i0}(x,T)\, E_g(x,T)^{\frac{3}{4}}\, T^{3/2}\, \exp\left(-\frac{E_g(x,T)}{2k_BT}\right) \tag{2.12}$$

where $k_B = 8.617 \times 10^{-5}$ is the Boltzmann's constant in [eV/K] and

$$\begin{aligned} n_{i0}(x,T) &= 5.24256 - 3.5729x - 4.74019 \times 10^{-4}T \\ &\quad + 1.25942 \times 10^{-2}xT - 5.77046x^2 - 4.24123 \times 10^{-6}T^2 \end{aligned} \tag{2.13}$$

[4] Several expressions are reported by Capper [31], among which also the adopted expression (the other proposed formulas can be found in Refs. [38], [41] and [42])

2.2.7 Electron and hole Auger recombination coefficients

The expressions here considered is that presented in Ref. [43], considering Auger–1 (A1) process. This consists in the direct band–to–band recombination of a conduction band electron with a heavy hole and the excitation of another electron in the conduction band, believed to be the dominant Auger process in n–type material. For p–type material, the Auger–7 (A7) process, involving holes in a very similar way, is important and has been included it as well.

In the device simulator, the Auger generation-recombination rate G_A is defined through the Auger coefficients C_n and C_p, respectively, for electrons and holes:

$$G_A = (C_n n + C_p p)\left(np - \Gamma n_i^2\right). \tag{2.14}$$

Γ is a dimensionless factor derived from Fermi statistics [15, Sec. 1.4.3], and in Boltmann's statistics $\Gamma = 1$.

The link with Auger intrinsic lifetimes τ_{A1}^i and τ_{A7}^i for A1 and A7 processes is given by [8, Sec. 14.3.4.3]

$$G_A = \left(\frac{n}{(1+an)\,\tau_{A1}^i} + \frac{p}{\tau_{A7}^i}\right)\frac{(np - \Gamma n_i^2)}{2n_i^2}, \tag{2.15}$$

where, for HgCdTe $a = 5.26\,10^{-18}\,\mathrm{cm}^3$ [8, Sec. 14.3.4.3]. Comparing the two expressions, if $an \ll 1$ (valid for the most practical cases, except for very high doped semiconductor) it follows:

$$C_n = \frac{1}{2n_i^2\,\tau_{A1}^i}$$
$$C_p = \frac{1}{2n_i^2\,\tau_{A7}^i}$$

that is the form implemented in the simulator.

Concerning the lifetime formulas available in literature, for the A1 process the expression presented by Lopes *et al.* [43], originally derived by Beattie and Landsberg [44] and Blakemore [45] (BLB formula) was considered:

$$\tau_{A1}^i(x,T) = 3.8\ \times\ 10^{-18}\,\epsilon_\infty(x)^2\sqrt{1+\frac{m_e(x,T)}{m_h}}\left(1+2\frac{m_e(x,T)}{m_h}\right)\frac{\exp(A)}{B(x,T)} \tag{2.16}$$

The parameters appearing in the equations are defined as

$$A(x,T) = \left(\frac{1+2\frac{m_e(x,T)}{m_h}}{1+\frac{m_e(x,T)}{m_h}}\right)\frac{E_g(x,T)}{k_B T}, \tag{2.17}$$

$$B(x,T) = m_e(x,T)|F_{12}|^2 \left(\frac{k_B T}{E_g(x,T)}\right)^{1.5}, \tag{2.18}$$

and $F_{12} = 0.2$ is the overlap integral of the Bloch functions (the cited Reference reports values ranging between 0.1 and 0.3). it For p–type material, the A7 process is the dominant one, and has been expressed considering the relation between C_p and C_n

$$C_p(x,T) = \frac{C_n(x,T)}{6}\left(\frac{1-\frac{3E_g(x,T)}{2k_BT}}{1-\frac{5E_g(x,T)}{4k_BT}}\right). \tag{2.19}$$

In Fig. 2.7 the electrons C_n and holes C_p recombination coefficients is shown as function of x.

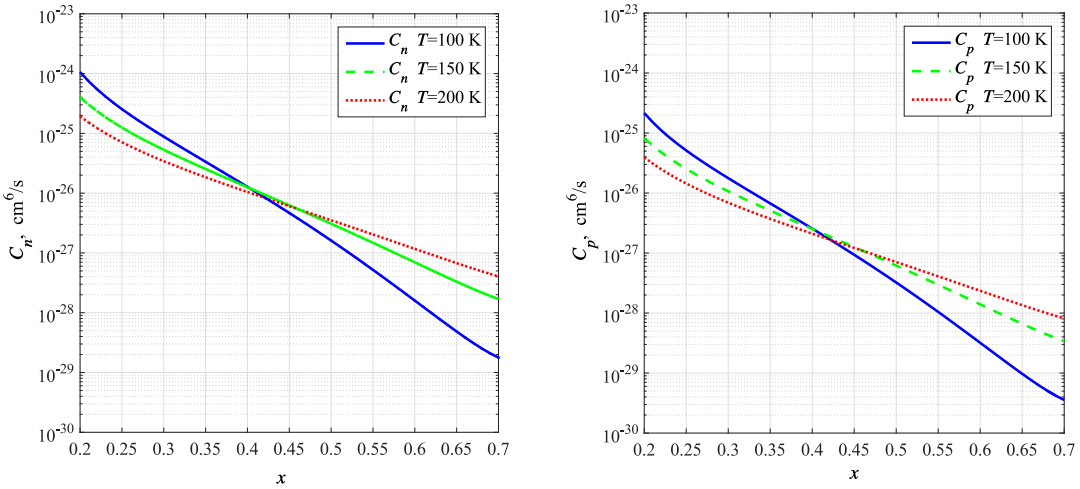

Figure 2.7. Auger electrons recombination coefficients as functions of Cd molar fraction x.

As an alternative, it could be possible to consider also the semi-empirical fitting formula for intrinsic A1 lifetime proposed by Kinch in 2005 [46] and deduced in a more complete previous paper [47] of 1973:

$$\tau^i_{A1}(x,T) = \frac{2.12\ 10^{-14}\, E_g(x,T)^{\frac{1}{2}}\ \exp\left(\frac{E_g(x,T)}{K_BT}\right)}{|F_{12}|^2\,(k_BT)^{1.5}}, \tag{2.20}$$

where F_{12} is the wavefunctions overlap integral. The expression in Ref. [47] by Kinch (1973) substantially concides with that in Ref. [43] by Lopes *et al.* and now implemented

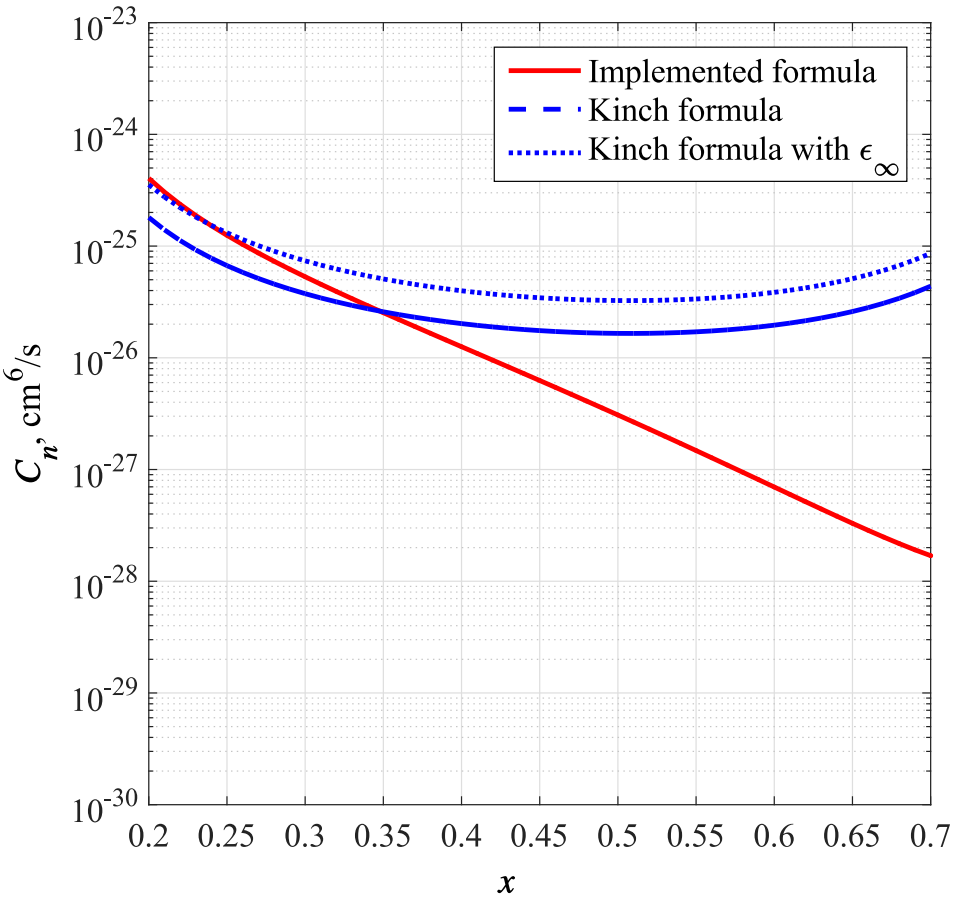

Figure 2.8. Auger electron recombination coefficient as a function of Cd mole fraction x: the formula implemented in the device simulator is compared with that proposed by Kinch [46].

(that cites it in its references). The simplified formula can be obtained in the limits:

$$\frac{m_e}{m_h} \ll 1$$

$$\frac{m_e}{m_0} = \frac{1}{1 + 2F + \frac{1}{3}E_p\left(\frac{2}{E_g(x,T)} + \frac{1}{E_g+\delta}\right)} \simeq \frac{3E_g}{2E_p}$$

where m_0 is the free electron mass, $E_p = 19\,\text{eV}$, $\delta = 1\,\text{eV}$ and $F = -0.8$; in obtaining his formula, Kinch builds the prefactor making use of a fixed (not x- and T-dependent) value of the dielectric constant. Furthermore it uses the *static* dielectric constant ϵ_s instead of the *dynamic* one ϵ_∞. In Fig. 2.8 a comparison is shown for $T = 150\,\text{K}$ between:

- C_n calculated with Eq. 2.16, the implemented formula;

- the original Kinch C_n calculated with Eq. 2.20;
- Kinch's C_n calculated with ϵ_∞ instead of ϵ_s in the prefactor.

(To build the C_n described in the third item, it is sufficient to multiply Kinch's coefficients by the squared ratio between the two dielectric constants, almost independent of the composition). It is possible to observe that Kinch's formula is comparable with Lopes' one, if the former uses ϵ_∞ instead of ϵ_s. Of course, increasing the Cd content, the approximated Kinch's formula progressively loses its accuracy.

About F_{12} uncertainties [48], Rogalski [8, Sec. 14.3.4.3] states that:

> the overlap integrals cause the biggest uncertainly in the Auger 1 lifetime. Values ranging from 0.1 to 0.3 have been obtained by various authors. In practice it is taken as a constant equal to anywhere between 0.1 and 0.3 leading to changes by almost an order of magnitude in the lifetime.

Unfortunately there is no agreement among authors about the ratio $\gamma = \tau^i_{A7}/\tau^i_{A1}$, which has been reported in the wide range $3 \leq \gamma \leq 60$ [48–51].

2.2.8 Radiative recombination rate

Radiative processes in HgCdTe involve the direct (band-to-band) generation-recombination of conduction band electrons with heavy holes and the corresponding absorption-emission of photons. The radiative GR rate G_R is given by:

$$G_R = B\left(np - \Gamma n_i^2\right), \tag{2.21}$$

where n, p and n_i are electron, hole and intrinsic density, Γ derives from Fermi statistics [15, Sec. 1.4.3] (in Boltmann's statistics $\Gamma = 1$), and B is the recombination coefficient ($\mathrm{cm^3 s^{-1}}$), related to the radiative lifetime τ_R by the relation [43]

$$\tau_R = \frac{1}{B\ (n+p)}. \tag{2.22}$$

Various expressions have been proposed for B, and among them, it has been chosen the formulation adopted in [43], plotted in Fig. 2.9 as a function of x, first obtained by Schachman and Finkman [52]:

$$\begin{aligned} B(x,T) = 5.8 \times 10^{-13}\ \epsilon_\infty(x)^{1/2} &\left(\frac{1}{m_e(x)+m_h}\right)^{3/2} \\ &\times \left(1 + \frac{1}{m_e(x)} + \frac{1}{m_h}\right)\ \left(\frac{300}{T}\right)^{3/2} \\ &\times \left(E_g(x,T)^2 + 3k_BTE_g(x,T) + 3.75(k_BT)^2\right). \end{aligned} \tag{2.23}$$

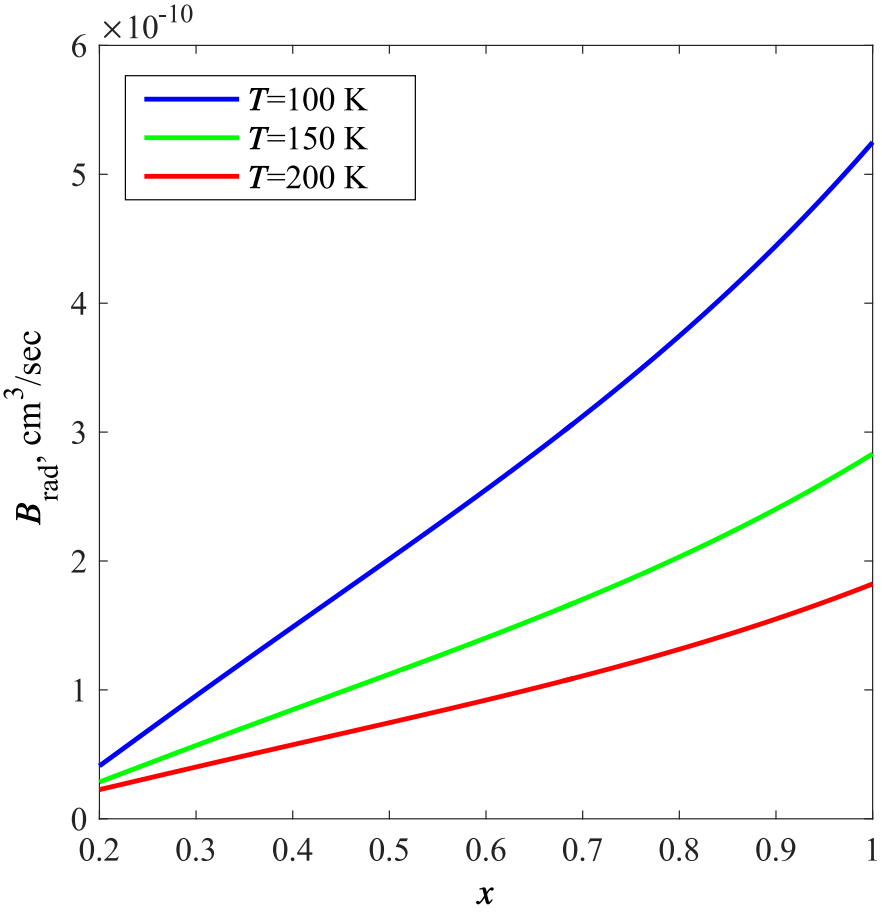

Figure 2.9. Radiative electrons recombination coefficient B as a function of Cd mole fraction x.

2.2.9 SRH recombination rate

The SRH model was introduced in 1952 [53,54] to describe the statistics of recombination and generation of holes and electrons in semiconductors occurring through the mechanism of trapping. The presence of trap levels within the forbidden band caused by crystal impurities facilitates the recombination process, since the jump can be split into two parts, each of them cheaper in terms of energy.

Models for this process involve equations for the densities of electrons in the conduction band, holes in the valence band, and trapped electrons. Basic for the SRH model are the drift–diffusion assumption for the transport of electrons and holes, the assumption of one trap level in the forbidden band, and the assumption that the dynamics of the trapped

electrons is quasi–stationary, which can be motivated by the smallness of the density of trapped states compared to typical carrier densities.

The individual characteristic properties of generation-recombination centers depend strongly on the technology. Therefore, they are usually lumped together in quantities like the effective electron and hole lifetimes, ending up with one effective single-level SRH mechanism. A doping dependent lifetime τ is empirically modeled by the so–called Scharfetter relation:

$$\tau = \tau_{min} + \frac{\tau_{max} - \tau_{min}}{1 + \left(\frac{N_A + N_D}{N_{ref}}\right)^{\gamma}} \tag{2.24}$$

where N_D and N_A are donors and acceptors densities, whereas γ, τ_{min}, τ_{max} and the density N_{ref} can be regarded as fitting parameters. This expression can be basically considered as a fit formula to account for experimental facts which strongly depend on process technology.

2.2.10 Absorption coefficient and refractive index

Direct bandgap semiconductors, such as HgCdTe, have a sharp onset of optical absorption as the photon energy increases above E_g. Strong optical absorption allows detector structures to capture very high percentage of the incoming signal even with thin layers of HgCdTe (usually within the range 10–20 μm). The results presented in this thesis make use of the Hougen formula [55] to describe the HgCdTe absorption coefficient.

The absorption in the exponential region, that is $E_{ph} < E_T$, can be modeled as:

$$\alpha(x) = \alpha_0(x)\ \exp\left(\frac{\sigma(x)\ (E_{ph} - E_0)}{T - T_0}\right) \tag{2.25}$$

whereas in the high absorbing region ($E_{ph} \geq E_T$) can be modeled as

$$\alpha(x) = \alpha_T(x)\sqrt{\frac{2\,\sigma(x)}{T - T_0}\left(E_{ph} - E_0 - \frac{T + T_0}{\sigma(x)} - \left(\ln\left(\frac{\alpha_T(x)}{\alpha_0(x)}\right) - 0.5\right)\right)} \tag{2.26}$$

The parameters introduced in the absorption equations have the following expressions and values:

$$\begin{aligned}
\alpha_0(x) &= \exp\left(-18.88 + 53.61x\right) \\
\alpha_T(x) &= 100 + 5000x \\
\sigma(x) &= 3.267 \times 10^4\ (1 + x) \\
E_0(x) &= 1.838x - 0.3424 \\
E_T(x) &= E_0 + \left(\frac{T + T_0}{\sigma(x)}\ \log\left(\frac{\alpha_T(x)}{\alpha_0(x)}\right)\right) \\
T_0 &= 81.9
\end{aligned}$$

and in Fig. 2.10 the absorption coefficient at 80 K are shown, for typical MWIR ($x = 0.3$) and LWIR ($x = 0.233$) HgCdTe compositions.

The real part of the refractive index n is implemented according to a T-independent expression described in the same Ref. [55], based on an experimental work by E. Finkman and S. E. Schacham [56], giving $n(x, T)$ as:

$$\begin{aligned} A(x) &= 16.4135 - 22.1914x + 11.081x^2 \\ B(x) &= 0.037514 + 1.060482x + 0.876032x^2 \\ C(x) &= 0.5694x^{-1.5355} \\ D(x) &= -1.4917 \times 10^{-3} + 2.1144 \times 10^{-3}x - 1.0415 \times 10^{-3}x^2 \\ E(x) &= -1.63 \times 10^{-7} \\ n(x,T) &= \left(A(x) + \frac{B(x)}{1 - \left(\frac{C(x)}{\lambda}\right)^2} + D(x)\lambda^2 + E(x)\lambda^4 \right)^{1/2} \end{aligned} \tag{2.27}$$

P. Capper and J. W. Garland [57] proposed another expression, based on experimental fitting procedure, that includes a dependence on T:

$$\begin{aligned} A(x) &= 13.173 - 9.852x + 2.909x^2 + 10^{-3}(300 - T) \\ B(x) &= 0.83 - 0.246x - 0.0961x^2 + 8 \times 10^{-4}(300 - T) \\ C(x) &= 6.706 - 14.437x + 8.531x^2 + 7 \times 10^{-4}(300 - T) \\ D(x) &= 1.953 \times 10^{-4} - 0.00128x + 1.853 \times 10^{-4}x^2 \\ n(x,T) &= \left(A(x) + \frac{B(x)}{1 - \left(\frac{C(x)}{\lambda}\right)^2} + D(x)\lambda^2 \right)^{1/2} \end{aligned} \tag{2.28}$$

In Fig. 2.11 the refractive index is reported, for typical MWIR ($x = 0.3$) HgCdTe composition, comparing the two formulations. Despite their differences, the results of optical simulations have been found very similar and the first one was adopted throughout.

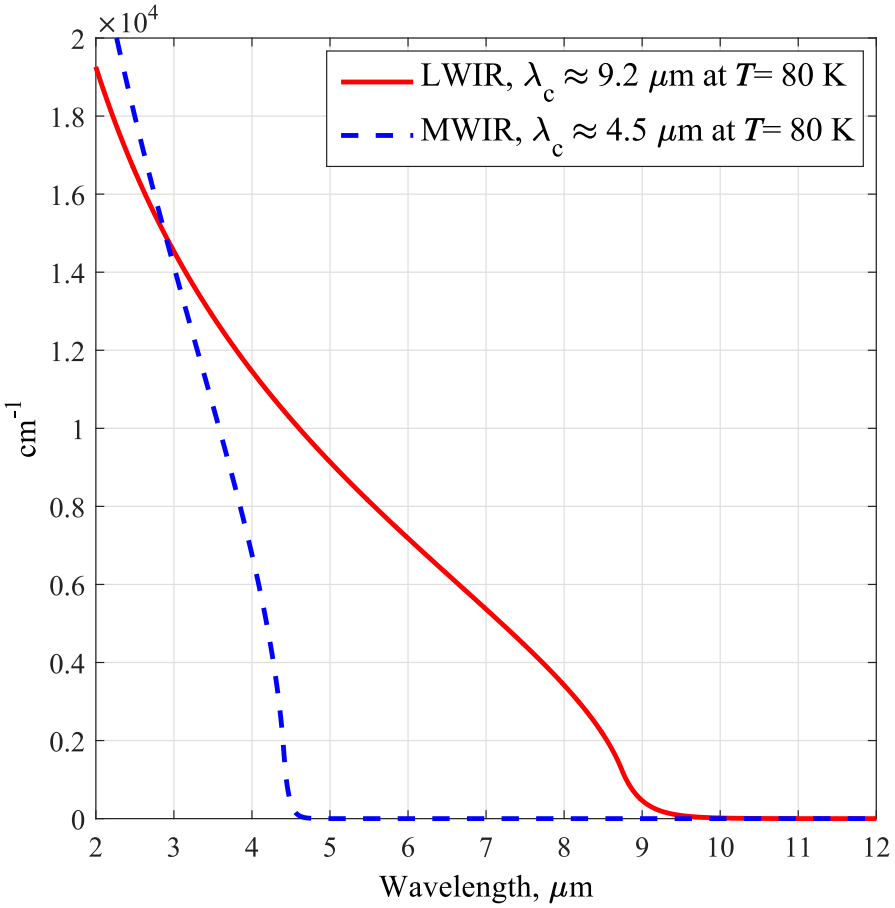

Figure 2.10. Absorption coefficient for typical LWIR ($x = 0.233$) and MWIR ($x = 0.3$) HgCdTe material.

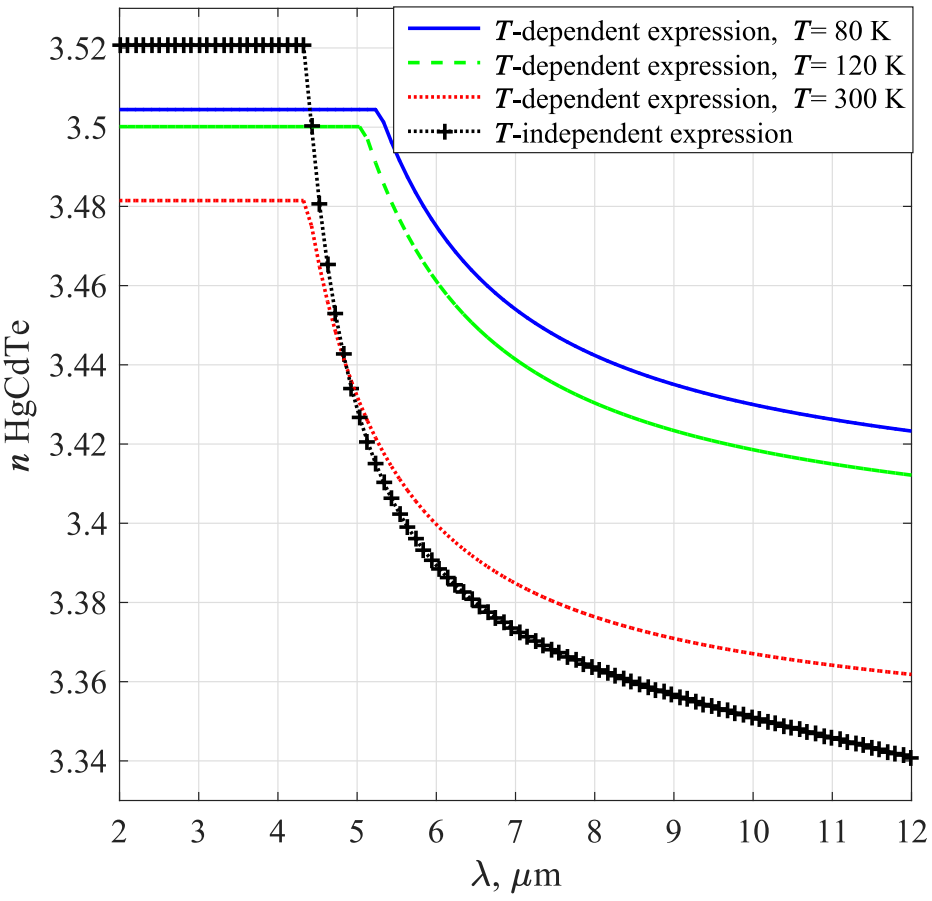

Figure 2.11. Refractive index for typical MWIR ($x = 0.3$) HgCdTe material.

2.2.11 Band to band tunneling

The classical expression of electron BTBT is due to Kane [58]

$$R_{\text{BTBT, Kane}} = A\mathcal{E}^2 \exp\left(\frac{B}{\mathcal{E}}\right) \tag{2.29}$$

where, for parabolic barriers, the A and B coefficients are [59, 60]

$$A = -\frac{q^2\sqrt{2m_e}}{4\pi^3\hbar^2\sqrt{E_g}}, \quad B = \frac{\pi\sqrt{m_e E_g^3}}{2\sqrt{2}\,q\hbar}, \tag{2.30}$$

$\hbar$ being the reduced Planck's constant. These expressions, as well as closely related ones derived for triangular barriers (see e.g. [60] for a review), apply rigorously to the case of constant electric field $\mathcal{E}$ across the junction.

Following the approach described in detail in Sec. 4.4, we obtained a new formulation, overcoming the ideality of the Kane's formulation:

$$R_{\text{BTBT}} = A\,\delta^{D-1}\mathcal{E}^{\frac{D+1}{2}} \exp\left(-\frac{B}{\delta\sqrt{\mathcal{E}}}\right). \tag{2.31}$$

It depends on the parameters D and $\delta \approx \sqrt{2qn_iL/\epsilon}$, where ϵ is the average dielectric constant, L is the total device length. In the simulations δ and D have been treated as fitting parameters expressing the non-ideality of barriers.

Chapter 3

1D, 2D, 3D simulations of HgCdTe based photodetectors

Three-dimensional numerical simulation of HgCdTe single-color and dual-band IR FPA photodetectors [13, 61] is obviously essential when investigating electrical and optical inter-pixel cross-talk [62], single pixel electromagnetic properties [63] or photon-trapping effects [64]. Nevertheless, when the single pixel has a simple axially-symmetric geometry, its dark current characteristics and photocurrent spectra are studied also with simpler 2D simulations [23, 36, 61, 65] in a symmetry plane containing all electric contacts. Furthermore, even simpler one-dimensional (1D) models along the main symmetry axis are also adopted, especially for dark current calculations [66–68].

For such symmetric devices, one could expect that simple relations exist between results computed with simulations of different dimensionality. In Sec. 3.1 a comparison is presented and discussed between dark currents and photocurrent spectra determined with 1D, 2D and 3D simulations of a single-color n-on-p single pixel of a FPA photodetector. The results demonstrate that, in general, 1D simulations are not applicable, especially under illumination, and some caution has to be used also in the interpretation of 2D results. Conversely, Sec. 3.2 discusses a dual-band MWIR/LWIR detector whose photoresponse can be approximated also with a 1D description, thanks to the device composition and doping profile, while 2D or 3D simulations can be reserved for validation/calibration of the 1D analysis and for the study of pixel-shape-dependent issues and inter-pixel cross-talk [69, 70]. Limitations of 1D and 2D models with respect to 3D descriptions are also discussed for dual-band devices, and some final remarks are given in Sec. 3.3.

This Chapter closely follows a manuscript published on a special issue of the J. Electron. Mater., vol. 43, n. 8, pp. 3070-3076, 2014 [2], and a summary of the results of this investigation has also been presented at the U.S. Workshop on the Physics and Chemistry of II-VI Materials, Chicago, Illinois (USA), October 1-3, 2013.

All the simulations have been done employing the HgCdTe software library developed

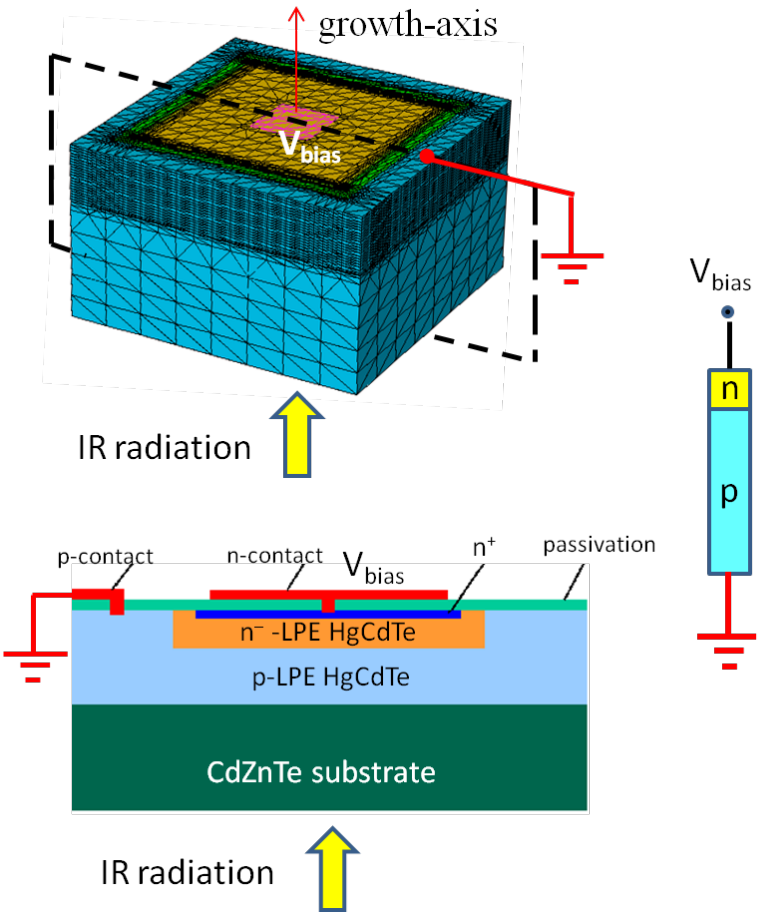

Figure 3.1. Single-color, single-pixel detector: 3D structure (above) showing the computational grid, its schematic 2D cross-section (below) and a 1D representation along the main symmetry axis (on the right).

and discussed in Section 2.2.

3.1 Single-color structures

The most widely manufactured second-generation FPAs are two-dimensional arrays of photovoltaic single-color detectors [8], usually operating in the mid-wavelength or long-wavelength infrared (MWIR, LWIR) bands. A possible architecture consists of a $N \times M$ pixel matrix realized in n-on-p planar technology with uniform composition and cutoff wavelength $\lambda_c \approx 5\,\mu$m (MWIR) or $\lambda_c \approx 9\,\mu$m (LWIR). Each pixel is a n-on-p HgCdTe diode fabricated on a CdZnTe substrate. For the present case study a square $15\,\mu$m-wide single pixel is considered, where the p-type HgCdTe layer has $\lambda_c = 5.3\,\mu$m (MWIR) and

acceptor density $N_A \approx 10^{16}\,\mathrm{cm}^{-3}$. The junction is obtained through ion implantation, yielding a $n - n^+$-doped region [71] whose doping profile is well approximated by an error function with maximum donor density $N_D = 10^{18}\,\mathrm{cm}^{-3}$. The detector is back-illuminated, and the n-contact is placed at the center of the n^+-doped region on the top surface and biased at a voltage V_{bias} with respect to the reference potential of the p-doped layer contact, making the photodiode to operate in reverse bias. In the following, electric contacts have been considered ideally ohmic, and all electrical and optical cross-talk effects have been neglected, considering only the isolated pixel properties.

Due to the axial symmetry of the device, it is possible in principle to compare numerical simulations based on the complete 3D structure, on a 2D cross-section, and even on a simple 1D description along the main vertical axis (see Fig. 3.1). Our analysis is based on the robust 1D/2D/3D numerical simulation model described in Section 1.3 which takes into account the composition, doping and temperature dependence of the HgCdTe alloy. The optical parameters depend on both molar fraction and temperature, and the drift-diffusion equations are solved using the finite box method. The simulated structures are discretized with a highly customizable meshing process, generating a denser mesh in regions where high gradients of current density, electric field, free charge density and material composition are expected. Fermi-Dirac statistics and incomplete dopant ionization are taken into account, with activation energies for HgCdTe alloys according to [8, 57].

Due to the high computational cost of 3D simulations, investigating to what extent 2D or 1D analyses are representative of the full 3D structure is of relevant practical interest. Thermodynamic equilibrium (no applied bias and no incident radiation) is the only condition where 1D, 2D and 3D simulations are equivalent, and the exact location of the p-contact has no practical effect, except possibly in a region very close to it. The first discrepancies between different dimensionality descriptions arise in dark current simulations. Fig. 3.2 compares simulated dark current densities for 1D, 2D and 3D pixels. The 2D and 3D dark currents are similar, but the 1D results are significantly overestimated, especially at higher temperatures. A possible explanation is the different p-contact position: in fact, both in 2D and 3D structures, that contact can be placed laterally, occupying a small area at a convenient distance from the pixel, as shown for example in [36] or in Fig. 3.1. On the contrary, in a 1D analysis the p-contact can only be placed on the illuminated face. Hence, a 1D structure is equivalent to a 2D or 3D device where the p-contact covers the whole illuminated face and the n-contact the whole top surface, enhancing the chance of excess minority carriers to be collected by contacts (which act as recombination centers), therefore producing higher dark currents. On the contrary, a 2D analysis can be regarded as a valid alternative, because of the higher similarity between 2D and 3D structures, thanks to their symmetry: therefore a 2D model can lead to simulation results quite representative of the 3D device, as observed in Fig. 3.2. If the pixels are less symmetric (e.g. dual-band structures with more complex geometry [72]) a simple proportionality between 2D and 3D results cannot be established, and 2D simulations are not directly applicable.

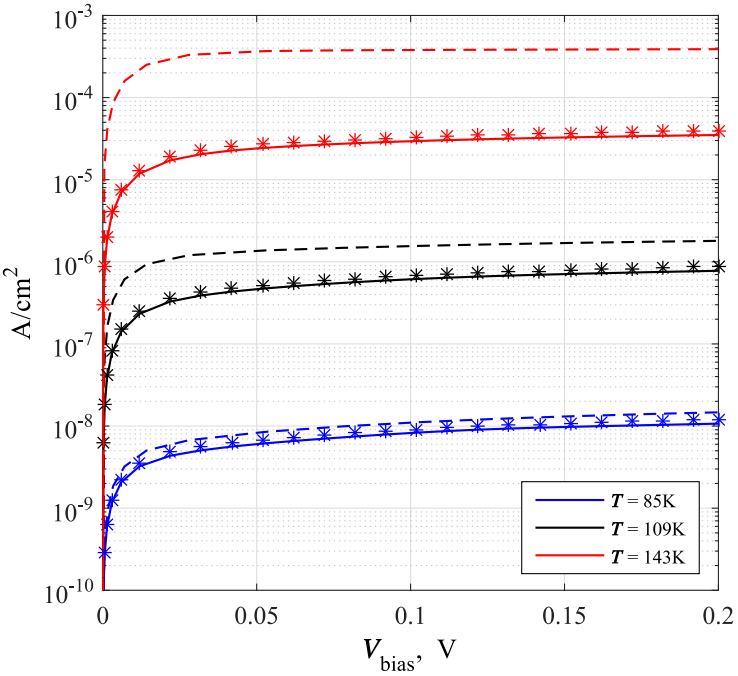

Figure 3.2. Single-color detector: 1D (dashed lines), 2D (symbols) and 3D (solid lines) simulated dark currents.

The simulation of 3D, 2D and 1D structures when illumination from the bottom face is considered must also be treated with special caution. In Fig. 3.3 an overall view of the photocurrent spectra is shown, calculated for all three cases with the same reverse bias ($V_{\text{bias}} = 0.2$ V) and temperature ($T = 85$ K, 109 K, 143 K), varying the wavelength λ of a monochromatic source of 1 mW/cm^2 (plane wave with propagation vector normal to the pixel bottom face). The 2D and 3D results are very similar and it is possible to consider the faster 2D simulation well representative of the 3D one. Instead, for the 1D simulation the difference is remarkable: in particular, 1D photocurrent spectra have a nonlinear shape on their short-λ side. The reason of the 1D behavior lies in the very different minority carrier density profile in the p region, due to the contact position: in 1D analysis, contacts are point-like, and the boundary conditions imposed by the p-contact correspond to a 2D or 3D structure where the contact extends unrealistically on the whole illuminated face. This particular placement of the p-contact, unavoidable in 1D simulations, forces the electron (minority carrier) density n on the p-contact to the thermal equilibrium value $n_0 = n_i^2/N_A^-$ determined by the continuity equation boundary conditions, where n_i and N_A^- are the intrinsic carrier density and the ionized acceptor concentration, respectively (partial compensation effects here are neglected). The minority carrier lifetime along the

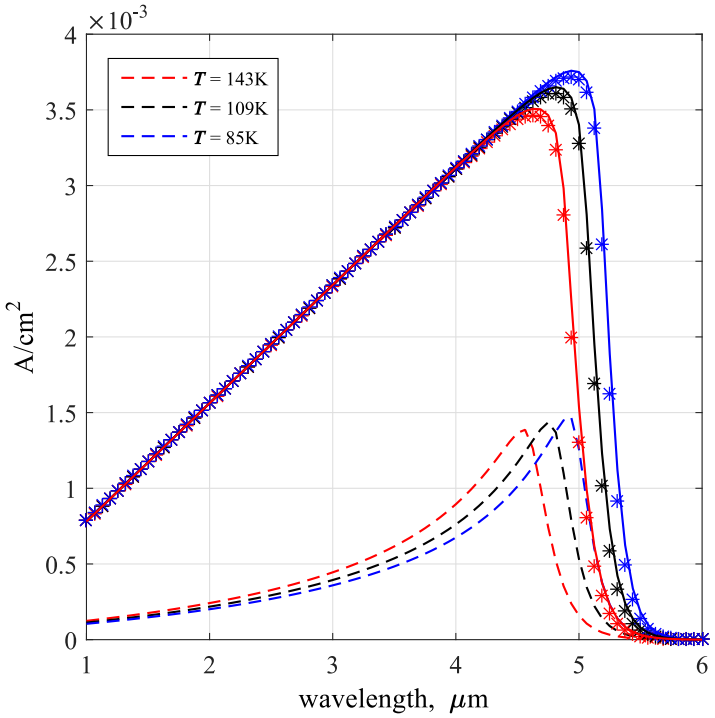

Figure 3.3. Single-color detector: 1D (dashed lines), 2D (symbols) and 3D (solid lines) simulated photocurrents.

absorption region is given by

$$\tau_e = \frac{n}{R} \tag{3.1}$$

where R is the recombination rate per unit volume, calculated self-consistently by the simulator considering all the contributions (radiative, Auger and SHR). Fig. 3.4 reports the τ_e profile along the pixel main symmetry axis according to Eq. (3.1) in the 1D, 2D and 3D structures. In all cases a monochromatic radiation of 1 mW/cm^2 at $\lambda = 5\,\mu$m illuminates the device (for better clarity, only the first $0.1\,\mu$m of the region near the illuminated face are shown). The lifetime τ_e has practically a uniform value everywhere in both 2D and 3D structures, whereas along the 1D structure it varies over several orders of magnitude in a few nanometers due to the boundary conditions. In the literature, 1D models suitable to calculate photocurrent spectra have been proposed in [73–76] (see also the review in [8]). In particular, considering a z-dependent lifetime $\tau(z)$ and absorption coefficient $\alpha(\lambda, z)$, an analytic form of the photoresponse has been derived [74–76]:

$$I_{\text{ph}}(\lambda) = P_0 \frac{q\lambda}{hc} \int_0^L \frac{\tau(z)}{\tau_0}\, \alpha(\lambda, z) e^{-\alpha(\lambda,z)z}\, dz \tag{3.2}$$

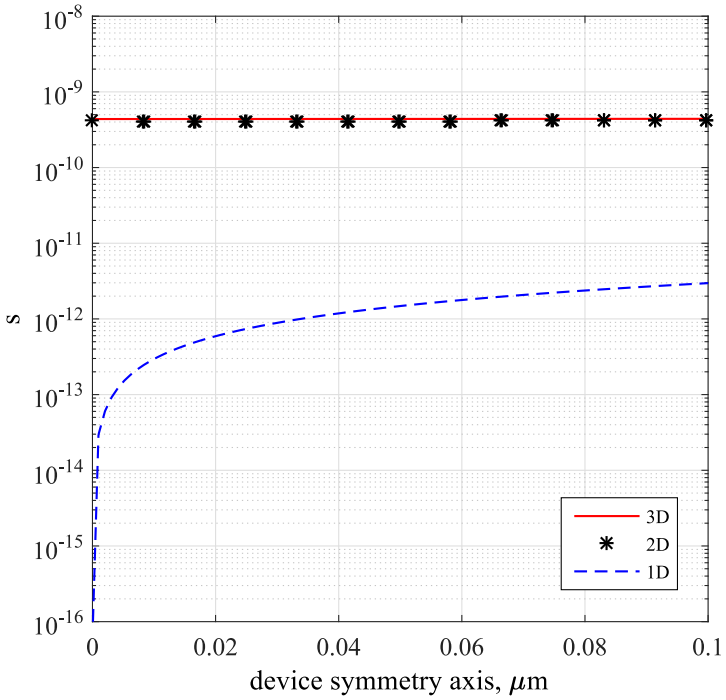

Figure 3.4. Single-color detector: 3D, 2D and 1D simulation of the minority carrier lifetime near the illuminated face at $T = 85\,\mathrm{K}$ and $V_{\mathrm{bias}} = 0.2\,\mathrm{V}$, with an illumination of $1\,\mathrm{mW/cm^2}$ at $\lambda = 5\,\mu\mathrm{m}$.

where the integration runs along the device axis (having a length L), L_0 is a fitting parameter determining the steepness of the exponential, P_0 is the illuminating optical intensity, τ_0 is a reference value for the lifetime in the absorption region, and

$$\tau(z) = \tau_0 \left[1 - \left(1 - \frac{\tau_s}{\tau_0}\right) \exp\left(-z/L_0\right)\right]. \tag{3.3}$$

Making use of Eq. (3.2), it is easily possible to reproduce both 3D-like and 1D-like photocurrent spectra by considering, respectively, a uniform or exponential τ profile along the pixel symmetry axis, as shown in Fig. 3.5. As an additional check, it was performed a dedicated 2D simulation of a pixel where the p contact, instead of being placed laterally, covers the whole surface of the lower face as for the 1D structure (the CdZnTe substrate was not included): the corresponding photocurrent spectrum is identical to the 1D result, demonstrating how the spectrum shape is determined by the unphysical location, unavoidable in 1D, of the p-contact.

Finally, Fig. 3.6 reports the 3D and 2D IQE whereas Fig. 3.7 shows the ratio between the 3D and 2D photocurrents $J_{\mathrm{ph\,3D}}/J_{\mathrm{ph\,2D}}$. The ratio, very close to unity (within 10% for all wavelengths), demonstrates the good consistency between 3D and 2D simulations.

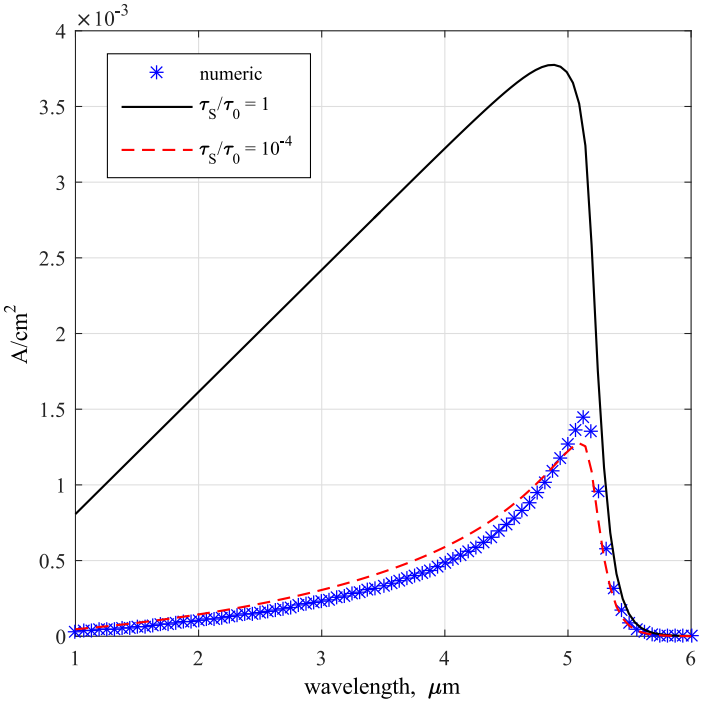

Figure 3.5. 1D analysis of the photocurrent spectrum for the single-color MWIR HgCdTe photodetector at $T = 85\,\mathrm{K}$, $V_{\mathrm{bias}} = 0.2\,\mathrm{V}$, with an illumination of $1\,\mathrm{mW/cm^2}$: numerical simulation (blue symbols), closed-form calculations with uniform (black solid) and exponential (red dashed) lifetime law, with $L_0 = L/1.5$.

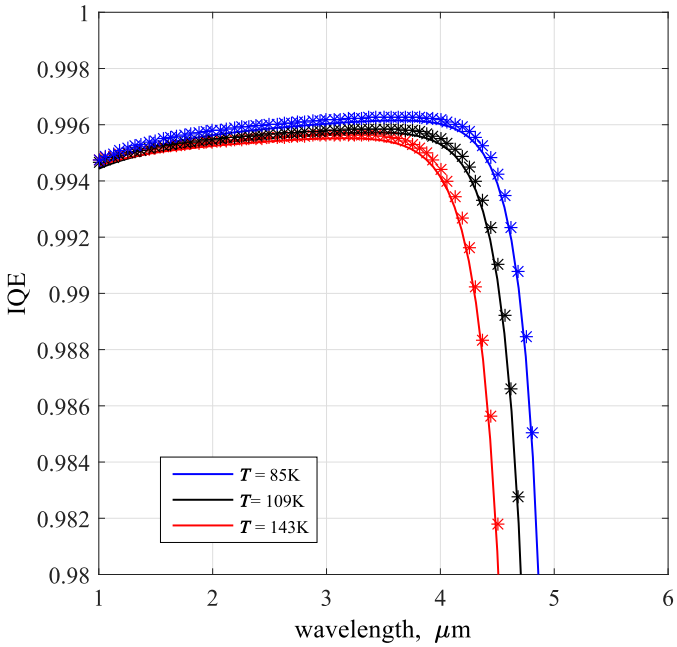

Figure 3.6. Single-color detector: 3D (solid lines) and 2D (symbols) simulated IQE.

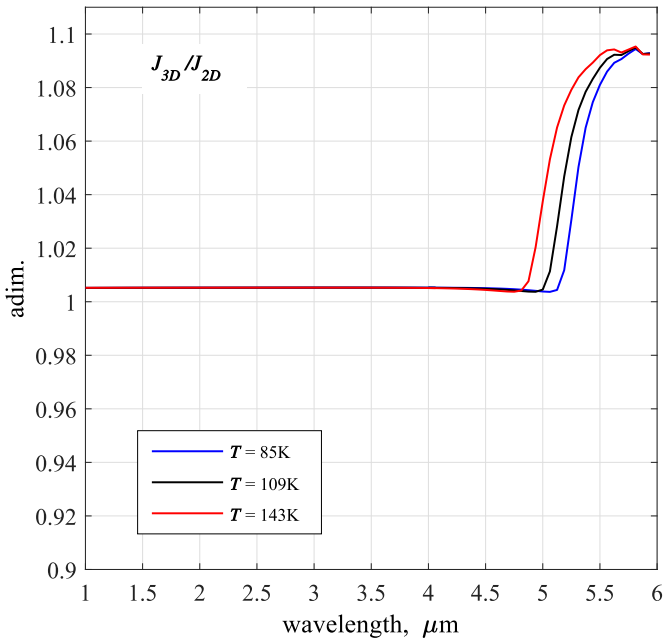

Figure 3.7. Single-color detector: ratio between the 3D and 2D simulated photocurrents.

3.2 Dual-band structures

In recent years there has been considerable interest in third-generation FPA detectors which incorporate enhanced device architectures. An example considered in this work is a dual-band bias-switchable back-illuminated $N \times M$ pixel matrix operating in LWIR and MWIR.

Above a CdZnTe substrate, a MWIR-SWIR-LWIR structure is grown ($\lambda_c = 5.3\,\mu$m, $3\,\mu$m, $8.9\,\mu$m, respectively) where the p-doped SWIR region acts as a barrier. The doping profile across the layer stack is $n-p-P-p-n$, where the capital P refers to the wide-gap SWIR barrier layer. The driving contact, biased at a voltage V_{bias}, can be placed on the top of the LWIR n-region; the reference potential contact, connected to the MWIR n-region, can be placed laterally in 3D and 2D structures, not interfering with the incoming IR radiation. (In all present dual-band simulations, the p-type central layers are floating.) The considered 3D structure and a 2D cross-section through a vertical plane containing the main symmetry axis are shown in Fig. 3.8. In the corresponding 1D structures, obtained considering just the main symmetry axis, one is forced to place the contacts at the two ends of the structure, occupying the whole top and bottom faces of the pixel. For this reason, two variants have been considered, shown in Fig. 3.9 (the CdZnTe substrate has not been included in the 1D structure): variant A has the ground contact connected directly to the n-MWIR region, as in the 2D and 3D structures, while variant B has a p-SWIR layer inserted between the contact and the n-MWIR region. This higher-gap layer should increase the equivalence of the 1D structure with the 2D/3D ones, helping to avoid the perturbation induced by the contact of the MWIR photodiode.

Photocurrent spectra have been simulated by illuminating the structure from below with a variable wavelength monochromatic source of 1 mW/cm^2. The calculated 3D and 2D IQE spectra, reported in Fig. 3.10 for three values of T, are practically equivalent; the small differences observed for the MWIR section are probably related to the 2D and 3D contact geometries, which cannot be regarded as strictly equivalent.

Concerning the 1D simulations reported in Fig. 3.11, both variants A and B have LWIR spectra very similar to 2D/3D ones, whereas only variant B seems to work properly for MWIR spectra. In fact, the MWIR spectra of variant A have the same nonlinear shape observed in 1D analysis of single-color photodiodes, due to the effect of directly contacting the n-MWIR region, where absorption mainly takes place.

On the contrary, the LWIR 1D spectra are not affected by the contact placement because the common p-doped section (its main absorption region) is floating. In general, a 1D cut along the symmetry axis cannot be regarded as equivalent to the corresponding 2D/3D structure, but an appropriate selection of the regions included in the simulation can minimize the perturbations due to the placement of a contact on the illuminated face.

Regarding the 2D/3D structures, the minority carrier (holes) density near the reference potential contact drops to very low values due to the electric boundary conditions ($p = n_i^2/N_D^+$), but this does not affect the overall device behavior, unlike in the 1D analysis

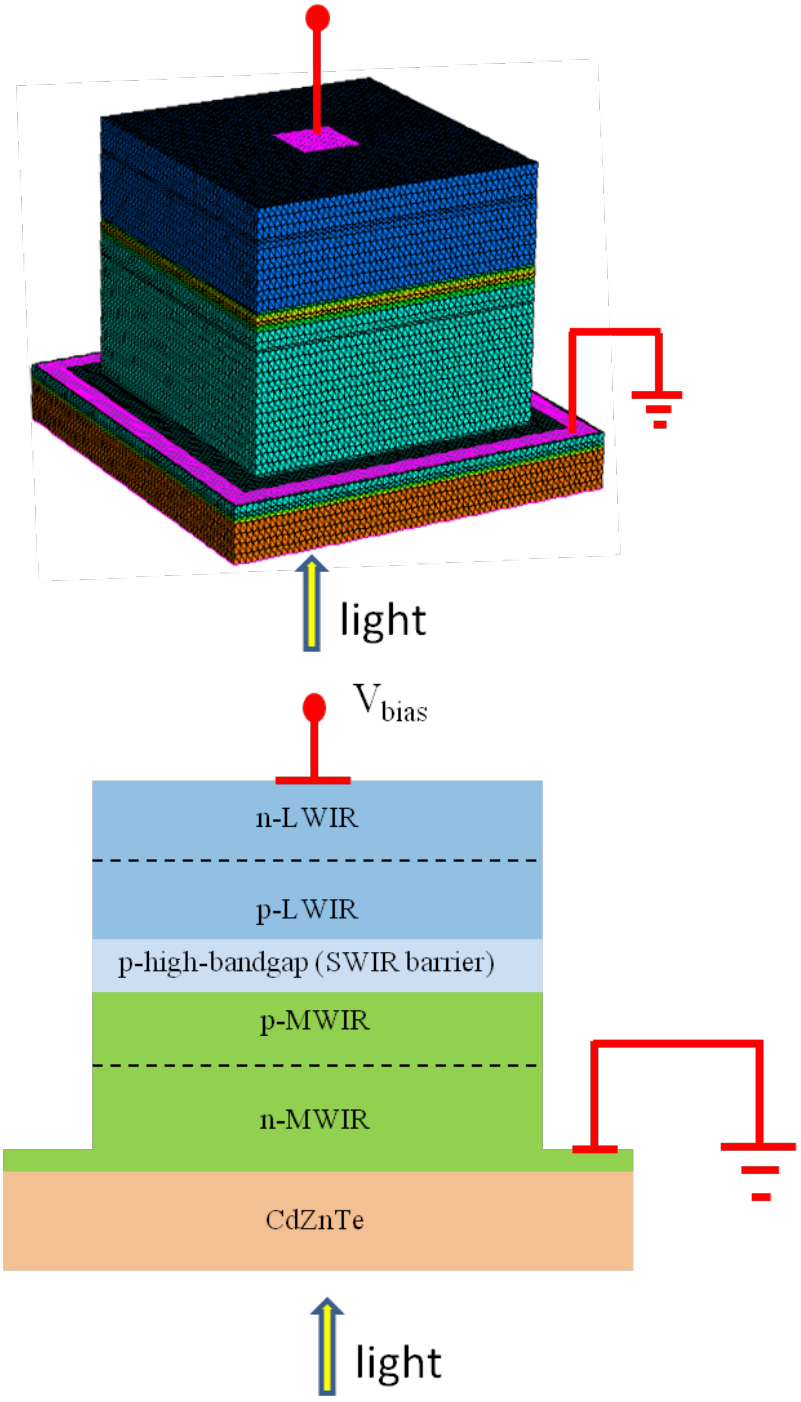

Figure 3.8. 3D (above) and 2D (below) structure of the dual-band single-pixel detector.

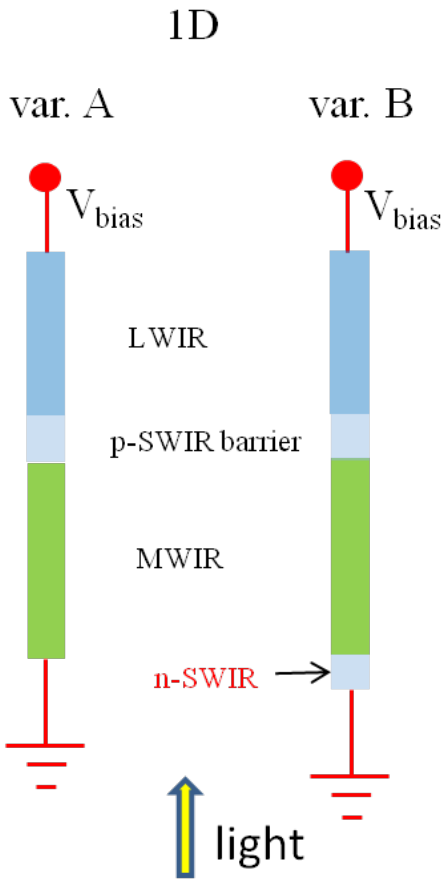

Figure 3.9. 1D structure, variants A (left) and B (right), of the dual-band single-pixel detector.

of dual-band variant A and of the single-color detector. In fact, looking at the carrier density distribution under illumination reported in Fig. 3.12, it is possible to observe that the perturbation due to the contact affects only a very small part of the n-MWIR region.

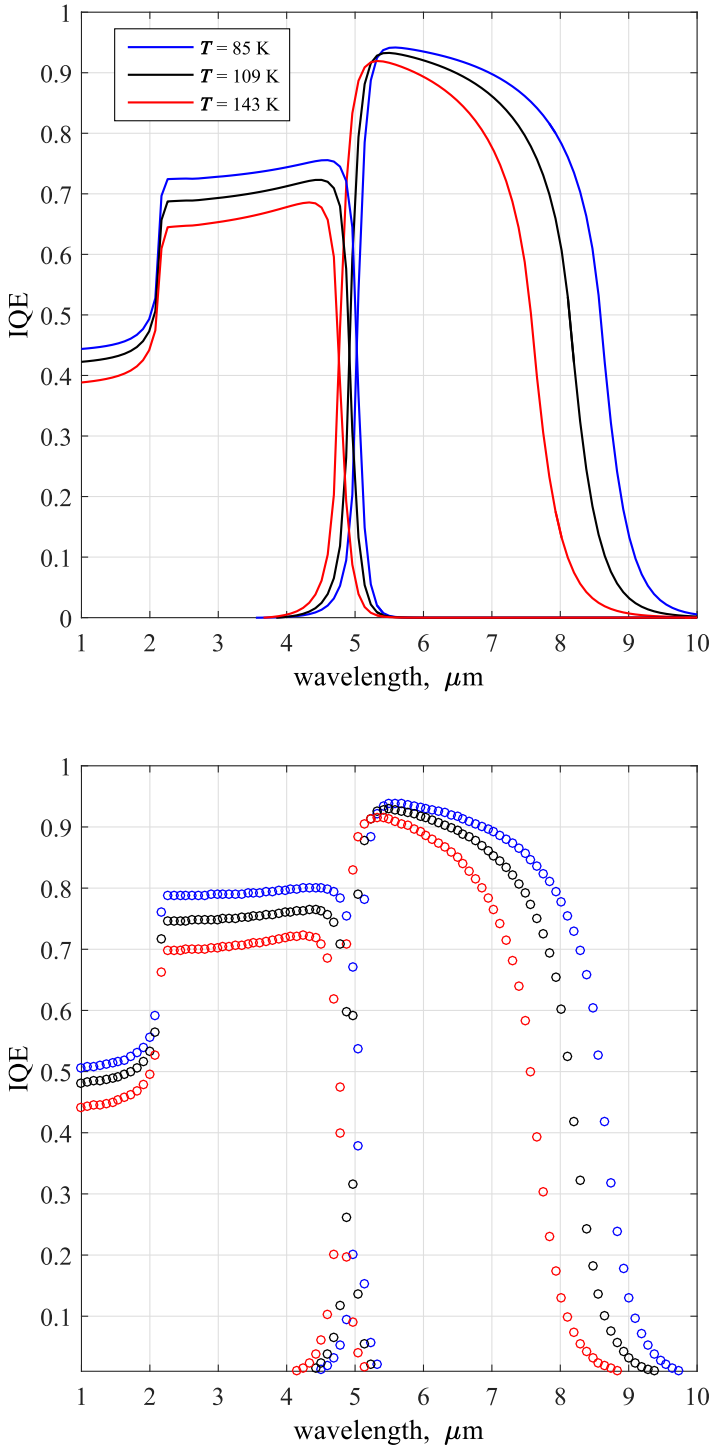

Figure 3.10. Dual-band detector: 3D (above, solid lines) and 2D (below, symbols) IQE for three values of T.

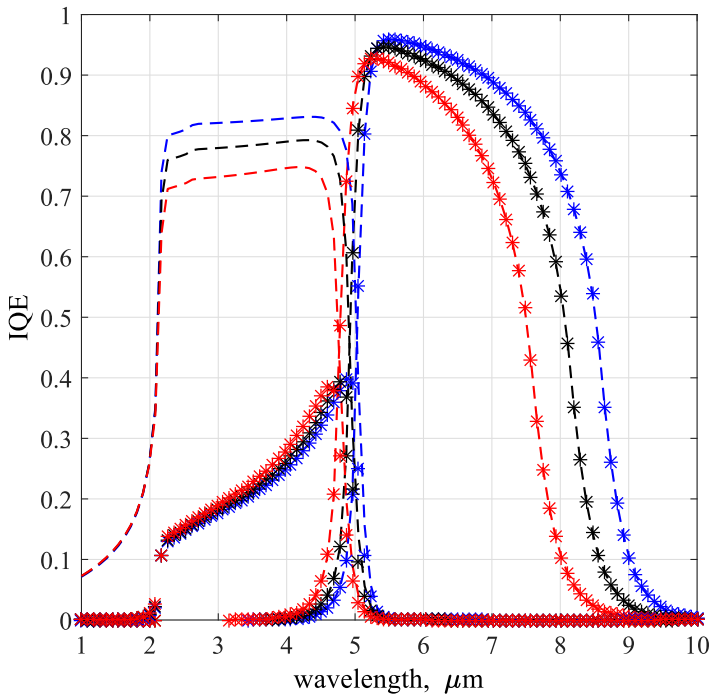

Figure 3.11. Dual-band detector: 1D simulation of IQE for three values of T, variant A (symbols) and variant B (dashed lines).

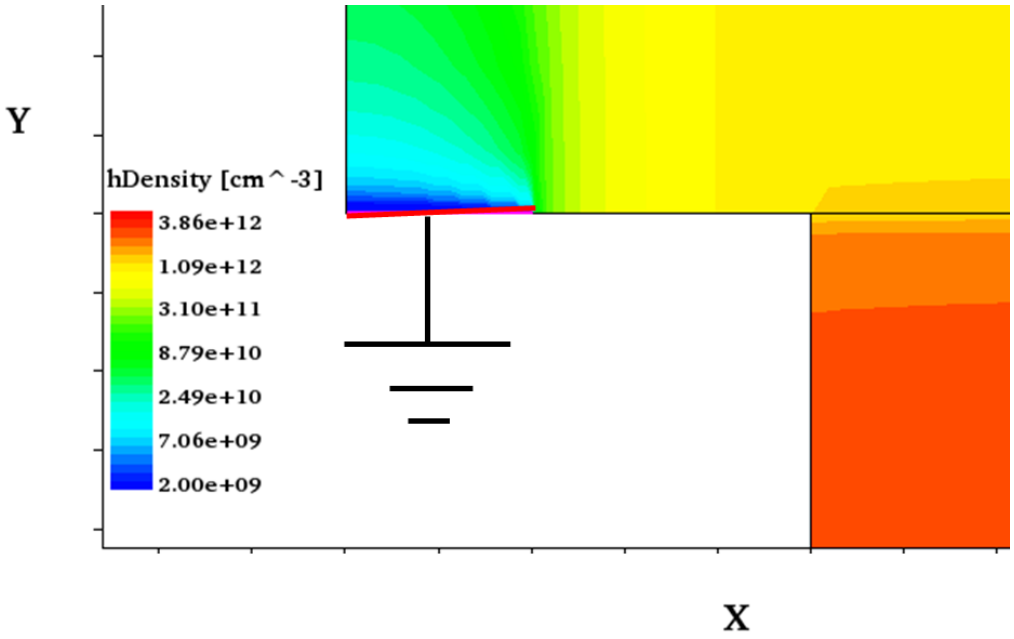

Figure 3.12. 2D simulation of hole density, for MWIR polarization. Illumination of $1\,\mathrm{mW/cm^2}$, $\lambda = 4.5\,\mu\mathrm{m}$, $T = 85\,\mathrm{K}$.

3.3 Final considerations

The comparison of 3D, 2D and 1D simulations of axially-symmetric single-color and dual-band HgCdTe pixels was the first published [2] application of the new developed HgCdTe software library, integrated in the Sentaurus simulator, although without a test against real experimental data: such further and more significant test are presented in Chapter 4.

With respect to realistic detectors, simplified structures with vertical sidewalls and without lateral passivations have been considered here, therefore the present results only underscore simulation issues that can originate considering lower dimensionality domains, without pretending to fit any experimental data. The 2D an 3D results have been found in good agreement also under illumination, and simplified 2D simulation can be regarded as valid alternative to a complete 3D analysis, at least to assess the overall performance of a structure, allowing to compare different technological solutions in a shorter time. Even simpler 1D descriptions are in general applicable only at equilibrium, where they may be useful in a preliminary design of composition and doping profiles. Especially under illumination, the applicability of 1D models depends on the particular epitaxial structure and cannot be regarded as an alternative to 2D/3D analysis.

Chapter 4

SRH and tunneling mechanisms in HgCdTe photodetectors

The present Chapter presents a combined experimental and numerical simulation study on two sets of nominally identical $Hg_{1-x}Cd_xTe$ single-color back-illuminated midwave infrared n-on-p photodetectors grown by liquid-phase epitaxy, p-doped with Hg-vacancies and with Au, respectively.

The described numerical model includes a novel formulation for *band-to-band tunneling*, which overcomes the intrinsic limitations of the classical Kane's description without introducing numerical issues typical of other approaches.

Our study confirms that adopting n-on-p architectures, avoiding metal vacancy doping, and reducing the acceptor density in the absorber region are prerequisites for obtaining HOT photodetectors.

A significant contribution to the dark current in both sets of devices is attributed to impact ionization, crucial to obtain a satisfactory explanation of the measured characteristics also at low-to-intermediate bias.

This Chapter closely follows a manuscript published on a special issue of the J. Electron. Mater., vol. 44, no. 9, pp. 3056-3063, 2015 [3], and a summary of the results of this investigation has also been presented at the U.S. Workshop on the Physics and Chemistry of II-VI Materials, Baltimore, Maryland (USA), October 21-23, 2014.

All the simulations have been done employing the HgCdTe software library developed and presented in Chapter 2.

4.1 Introduction

Next generation IR detectors require focal plane arrays [8, 11, 16, 20, 77] designed according to the HOT concept [21, 46]. HOT narrow-gap semiconductor devices operate in

nonequilibrium mode such that the carrier densities are held below their equilibrium, near-intrinsic, levels. Since HOT infrared detectors should preserve at 150 K and above the performance characteristics formerly obtained at substantially lower temperatures (80 K to 100 K), a considerable reduction of the dark current is essential to obtain high sensitivity [22, 51, 61, 78].

Two sets of devices, manufactured with the same nominal structure but different p-doping technologies, are presented in Sec. 4.2, and their experimental dark currents are compared and discussed. In Sec. 4.3, a simulation study of the dark currents demonstrate the favorable effects on SRH and Auger GR rates of avoiding metal-vacancy doping and reducing the acceptor density. Sec. 4.4 discusses the significant role of impact ionization (I.Ion.) and, in one set of examined devices, band-to-band tunneling, having applied to graded-composition HgCdTe-based detectors a novel BTBT formulation derived from a description originally conceived for Si-based devices [79, 80].

4.2 Device fabrication and experimental dark currents

In this work two sets of $Hg_{1-x}Cd_xTe$ back-illuminated MWIR photodetectors are compared; their cutoff wavelength λ_c is around $5.3\,\mu$m at $T = 80$ K, and they were grown by liquid-phase epitaxy (LPE) on a CdZnTe substrate (see Fig. 4.1). *Set A* is Hg-vacancy p-doped through a standard technology (acceptor concentration $N_A = 2 \times 10^{16}\,\text{cm}^{-3}$), while in *set B* Au is used as acceptor ($N_A = 5 \times 10^{15}$ cm^{-3}) and the number of Hg vacancies is kept low [13, 81]. Both sets were fabricated as large matrices of $15 \times 15\,\mu$m pixels, passivated with a CdTe layer. A strong interdiffusion with the underlying HgCdTe resulted in the profile of the Cd mole fraction x reported in Fig. 4.2, as determined by secondary ion mass spectroscopy (SIMS). The photodiode junctions were defined by ion implantation into the p-type layer, with maximum donor density $N_D = 10^{18}\,\text{cm}^{-3}$.

Dark current measurements were performed on HgCdTe chips hybridized to fan-out circuits, with several photodiodes connected in parallel allowing the measurement of low current levels [81]. Those diode fields were surrounded by rows of interconnected diodes around the perimeter, independently biased and thus acting as guard structures. In that way, the current generated outside the pixel field under consideration can be effectively siphoned off. All measurements were performed in a custom-made liquid nitrogen evaporation cryostat where the temperature can be varied and stabilized between 80 K and 300 K. Fan-out circuits were mounted into a closed cavity held at the detector temperature.

The experimental dark current density $J_{\text{dark}}(V;T)$ for *set A* and *set B* is reported in Fig. 4.3. For all considered temperatures, the dark current in *set A* is considerably higher than in *set B*, also for low values of reverse bias voltage V_{bias}. Moreover, a much stronger contribution of BTBT and/or I.Ion. may be noticed in *set A* with respect to *set B*. The goal of the numerical investigation presented in the following Section is to provide a detailed

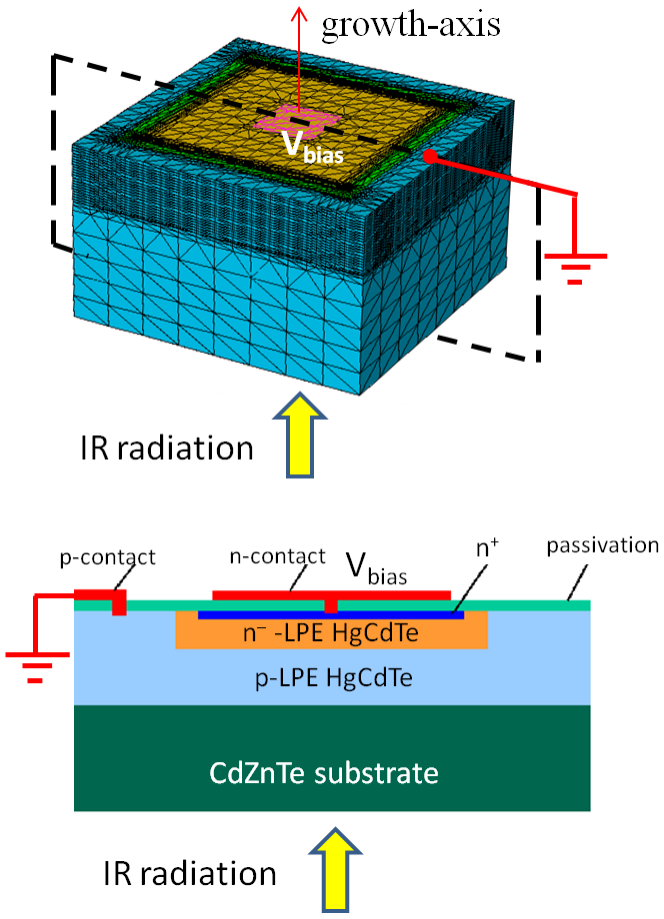

Figure 4.1. 3D structure of the single-pixel photodetectors under study (above) and 2D cross-section at the device center (below).

quantitative explanation of the observed differences in J_{dark} between *sets A* and *B*.[1]

4.3 Simulation of the low reverse bias regime

For both *set A* and *set B*, starting from the CdZnTe substrate (not included in the computational box), the simulated photodiode structures include a 7 μm-thick p-type HgCdTe region, followed by a 1.5 μm-thick, 10 μm-wide n^--n^+ HgCdTe region (see Fig. 4.1), whose Cd mole fraction closely follows the experimental profile shown in Fig. 4.2. The

[1] In all $J_{\text{dark}}(V)$ characteristics reported in the present work, voltage has been given negative values when the junctions are reverse-biased. Since the bias is applied to the cathode, a reverse bias corresponds to a positive voltage with respect to the reference ground contact, connected to the p side of the diode and located far from the junction.

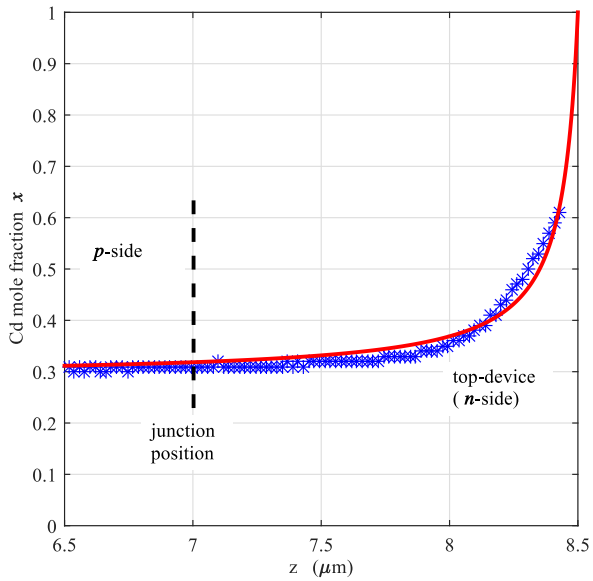

Figure 4.2. SIMS profile (blue symbols) of the Cd mole fraction x near the surface (located at $z = 8.5\,\mu$m) and least-squares fit (red solid line) used in the simulations. The estimated position of the p-n^- junction is marked with a vertical dashed line.

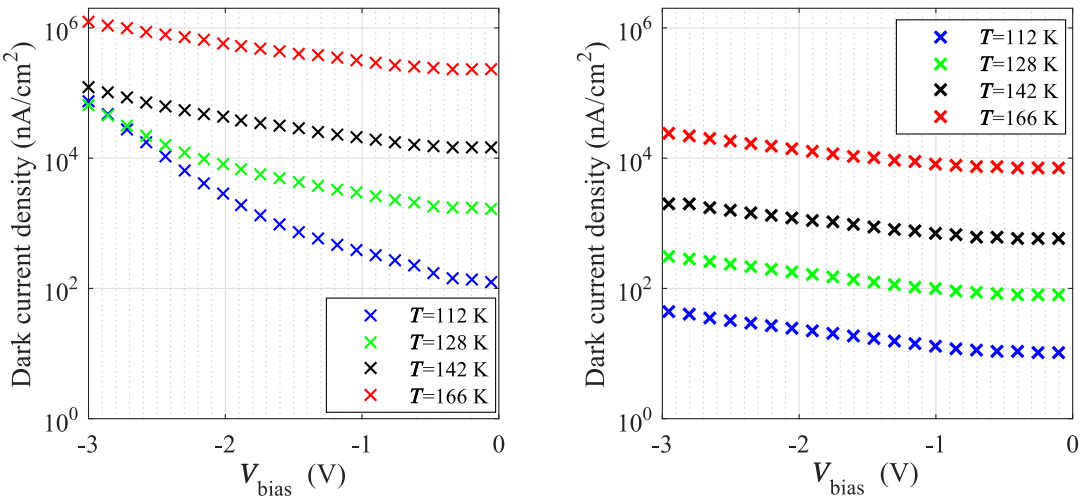

Figure 4.3. Experimental $J_{\text{dark}}(V;T)$ for *set A* (left) and *set B* (right).

n^- doping corresponds to a uniform donor density $N_D = 4 \times 10^{14}\,\text{cm}^{-3}$ while, in the topmost $0.5\,\mu$m-thick n^+ region, N_D increases from $4 \times 10^{14}\,\text{cm}^{-3}$ to $10^{18}\,\text{cm}^{-3}$ with an isotropic error-function profile. The cathode is placed at the center of the n^+-doped region

on the top surface and is biased at a voltage V_{bias} with respect to the reference potential of the p-doped layer contact, making the photodiode to operate in reverse bias. In the following, electric contacts have been considered ideally ohmic, and all electrical and optical cross-talk effects have been neglected, considering only the isolated pixel properties. The only difference between the model of the two sets of devices consists in the acceptor density N_A in the p-type absorber region, 2×10^{16} cm^{-3} for *set A*, 5×10^{15} cm^{-3} for *set B*.

It has been shown [2] in Chapter 3 that 2D simulations can reliably reproduce the experimental dark currents of the photodetectors under investigation without the computational cost of a full 3D description, the only requirement being the axial symmetry of the device. Therefore, the present analysis is based on the 2D numerical solution of the drift-diffusion (DD) equations using the finite box method, as described in Section 1.3. Composition, doping and temperature dependence of the electronic, transport and optical parameters of HgCdTe are described through the models proposed in [29, p. 1243] for the energy gap E_g, [36, p. 1331] for the electron affinity χ, [34] for electron and hole effective masses (see also [8, Eq. 14.6]), [8, Eq. 14.7] for electron and hole mobilities (derived from experimental measurements for Cd mole fraction between 0.2 and 0.6 [37]) and [43, Eqs. 5, 9] for radiative and Auger lifetimes, where it has been assumed $F_{12} = 0.2$ for the overlap integral of the Bloch functions and $\gamma = 6$ for the ratio between Auger-7 and Auger-1 intrinsic lifetimes [48, 82]. The simulated structures are discretized with a highly customizable meshing tool, generating a denser mesh in regions where high gradients of current density, electric field, free charge density and material composition are expected. Fermi-Dirac statistics and incomplete dopant ionization are taken into account, with activation energies for HgCdTe alloys according to [8, 57].

The lifetime of SRH recombination processes are modeled as [53]

$$R_{SRH} = \frac{np - n_i^2}{\tau_p \left(n + n_i e^{(E_t - E_{\text{Fi}})/(k_B T)}\right) + \tau_n \left(p + n_i e^{-(E_t - E_{\text{Fi}})/(k_B T)}\right)} \tag{4.1}$$

where n, p, n_i are the electron, hole and intrinsic density, E_t and E_{Fi} the trap energy and the intrinsic Fermi level The lifetimes τ_n, τ_p can be expressed as [83]

$$\tau_n = \frac{1}{(N_t \sigma)_n v_{th,\,n}}, \quad \tau_p = \frac{1}{(N_t \sigma)_p v_{th,\,p}}. \tag{4.2}$$

and depend on the local value of the product between trap density N_t and carrier capture cross-sections σ, and on the thermal velocity

$$v_{th,\,n} = \sqrt{\frac{3 k_B T}{m_e}}, \quad v_{th,\,p} = \sqrt{\frac{3 k_B T}{m_h}} \tag{4.3}$$

where $m_{e,h}$ are the effective electron and hole masses. According to [46], E_t lies near the intrinsic Fermi level for n-type doping and for p-type doping with As, Cu, Au, and is

locked at ≈ 30 meV below the conduction band edge E_c for vacancy-doped $Hg_{1-x}Cd_xTe$ with $0.2 < x < 0.5$.

The first step of the present simulation study has been to reproduce the experimental dark currents at low reverse bias ($V_{bias} \leq 0.5$ V) by taking into account only Auger, radiative and SRH processes (i.e., excluding both BTBT and I.Ion.). Under these assumptions, the net recombination rate R in the continuity equation becomes

$$R = R_{\text{Auger}} + R_{\text{Radiative}} + R_{\text{SRH}}. \tag{4.4}$$

The simulated dark current densities have been tuned by varying only three SRH parameters: the trap energy E_t in the n section, and the products $(N_t\sigma)_{n,p}$ in the p and n regions. In order to reduce the complexity of the tuning process, it has been assumed $E_c - E_t = 30$ meV in the p section [46] of both *set A* and *set B*, since the depleted region of the photodiode extends mainly in the n^- layer and, as a result, the sensitivity of $J_{\text{dark}}(V;T)$ to the SRH parameters in the p region is almost negligible. The resulting values for $N_t\sigma$ and E_t are reported in Table 4.1.

Table 4.1. Estimated SRH parameters for *set A* and *set B*.

region	$N_t\sigma$, cm^{-1}	$E_c - E_t$, meV
	Set A	
p	0.01	30
n^-, n^+	0.8	70
	Set B	
p	0.001	30
n^-, n^+	0.0018	120

Once determined the SRH parameters, the dark current "components" were also compared: they were determined by solving the DD model when each contribution to the total GR rate in Eq. 4.4 is considered as the only source term in the continuity equations (due to the nonlinearity of the DD equations, the total dark current is not expected to be the sum of the Auger, SRH and radiative "components". Besides these GR contributions, carrier diffusion from the quasi-neutral to the depleted regions also plays a role [28], although not easily separable from the others in a DD framework). From the simulated $J_{\text{dark}}(V;T)$ characteristics, reported in Figs. 4.4–4.5, the following comments are in order. First, since the contributions of BTBT and I.Ion. are not included, only the low-bias regime is adequately described. Second, the devices of *set A* are SRH-limited, given that the SRH

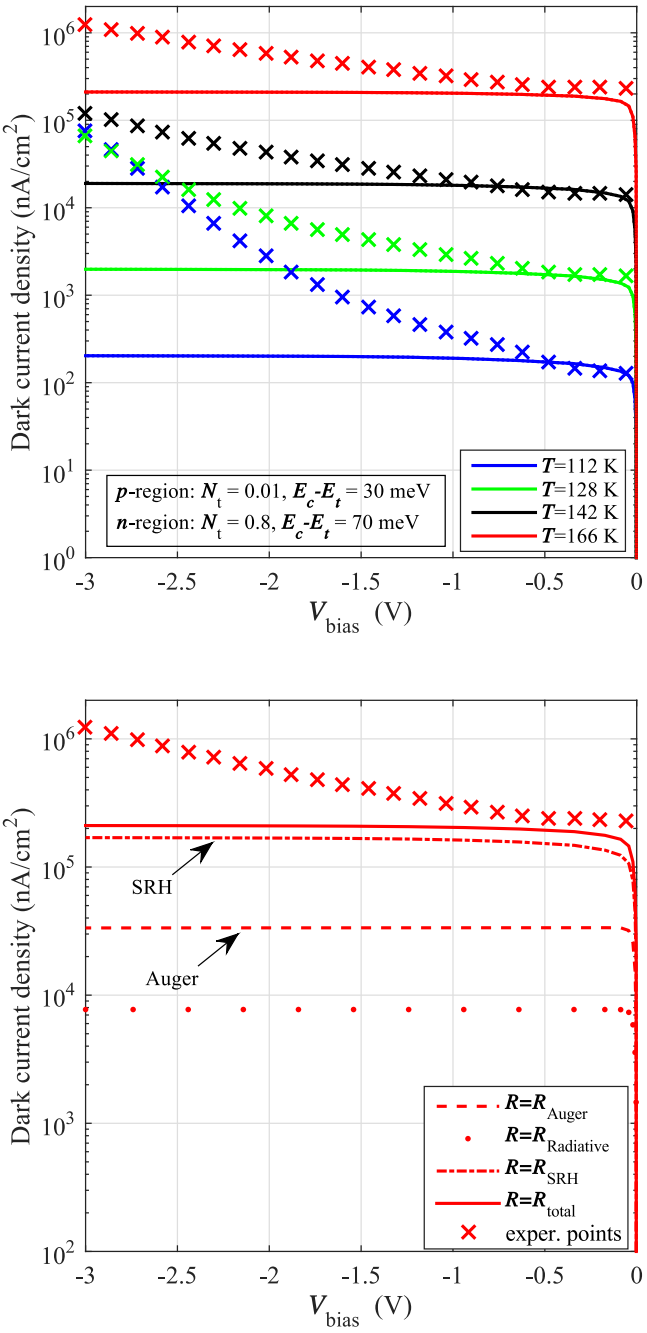

Figure 4.4. *Set A*: simulated $J_{\text{dark}}(V;T)$ (above) and corresponding components at $T = 166\,\text{K}$ (below). Crosses indicate experimental values.

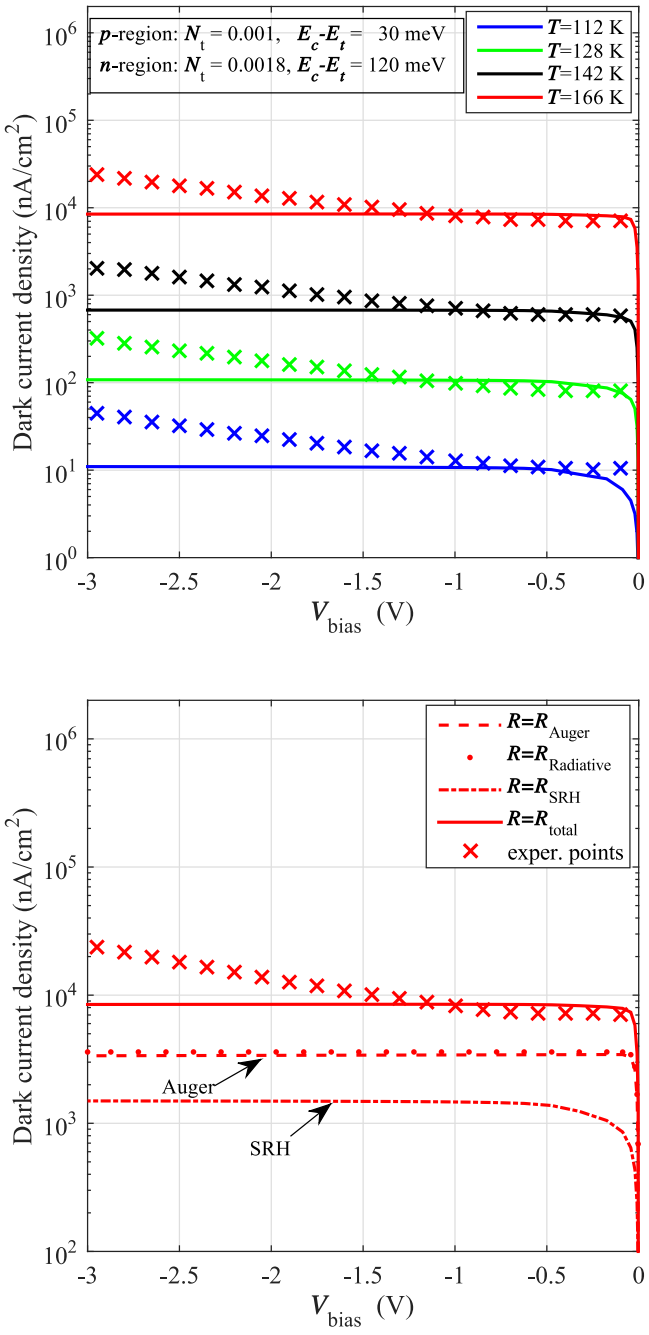

Figure 4.5. *Set B*: simulated $J_{\text{dark}}(V;T)$ (above) and corresponding components at $T = 166\,\text{K}$ (below). Crosses indicate experimental values.

current component dominates even at the highest temperature (see Fig. 4.4, lower panel). In *set A* one can expect an additional relevant contribution by BTBT and/or I.Ion., whose extent will be investigated in the next section. In *set B*, neither Auger nor SRH are clearly dominant, and all the considered GR mechanisms concur to yield the observed total dark current. Assuming that the capture cross-sections are the same in the two sets and that $\sigma_n = \sigma_p$ [84], Table 4.1 suggests that in *set B* the trap densities in the p and n sections are respectively one and two orders of magnitude lower than in *set A*, which could be correlated to the substantial reduction of metal vacancies in *set B*. The estimated N_t reductions in *set B* bring down the SRH component by more than two orders of magnitude with respect to *set A*. Last, a decrease of the Auger component by one order of magnitude is observed in *set B* with respect to *set A*, which could be ascribed to the reduction of the acceptor density in the p-doped section (see Sec. 4.2). The favorable effects of avoiding vacancy-doping and reducing the doping density in the absorber region, already discussed in [46], were experimentally demonstrated in [22] for high-density vertically-interconnected photodiodes (HDVIPs). Despite the structural differences between HDVIPs and the devices considered in the present work, the observed dark current reduction between *set A* and *set B* at 166 K is in substantial quantitative agreement with the corresponding reduction observed in [22, Fig. 19] at about the same temperature.

4.4 Simulation of the high reverse bias regime

The classical expression of electron BTBT is due to Kane [58]

$$R_{\text{BTBT, Kane}} = A\mathcal{E}^2 \exp\left(\frac{B}{\mathcal{E}}\right) \tag{4.5}$$

where, for parabolic barriers, the A and B coefficients are [59, 60]

$$A = -\frac{q^2\sqrt{2m_e}}{4\pi^3\hbar^2\sqrt{E_g}}, \qquad B = \frac{\pi\sqrt{m_e E_g^3}}{2\sqrt{2}\,q\hbar}. \tag{4.6}$$

These expressions, as well as closely related ones derived for triangular barriers (see e.g. [60] for a review), apply rigorously to the case of constant electric field $\mathcal{E}$ across the junction. Some authors have studied less idealized cases, accounting e.g. for the Fermi levels position in the neutral regions [17] or considering 2D and 3D realistic junction profiles [59, 60], but the application of the corresponding solutions in device simulators is affected by serious numerical issues. With the aim to avoid such difficulties, a different BTBT model has been recently proposed for Si-based tunnel field-effect transistors [79] and *p-i-n* diodes [80]. According to this approach, the term $\mathcal{E}^2$ in Eq. 4.5 has been replaced with the product $\mathcal{E}_{av}^{(D-1)}\mathcal{E}$ [85], where $\mathcal{E}_{av}$ is the average electric field in the depleted region and D is a dimensionless empirical parameter. Working in the WKB approximation,

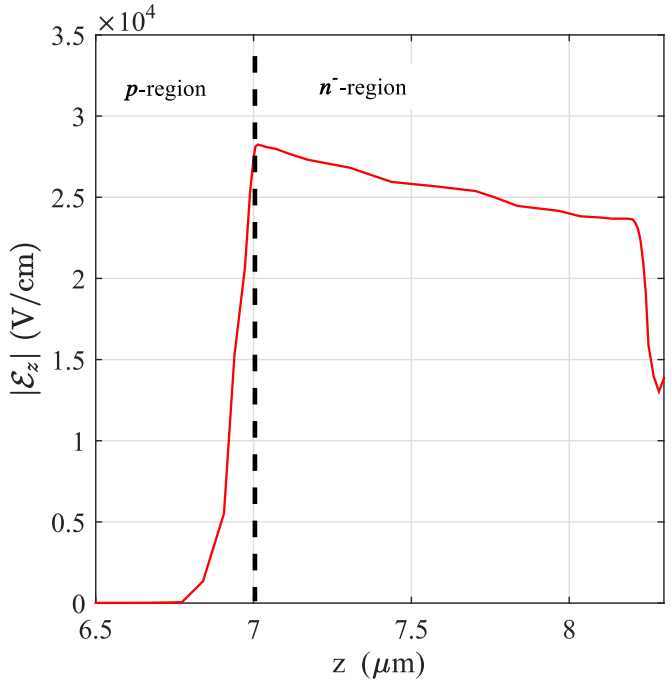

Figure 4.6. Calculated vertical component of the electric field $|\mathcal{E}_z|$ in *set A* near the p-n^- junction (marked with a vertical dashed line) at a bias $V_{\text{bias}} = -3$ V.

starting from the expression of the average field $\mathcal{E}_{av} = E_g/(q\, l_{\text{path}})$ across the junction [85] (where l_{path} is the distance between the two WKB turning points, the so-called tunneling path-length) and applying the charge neutrality equation, the expression

$$R_{\text{BTBT}} = A\, \delta^{D-1} \mathcal{E}^{\frac{D+1}{2}} \exp\left(-\frac{B}{\delta\sqrt{\mathcal{E}}}\right) \tag{4.7}$$

was derived. It depends on the parameters D and $\delta \approx \sqrt{2qn_iL/\epsilon}$, where ϵ is the average dielectric constant and L is the total device length. In the following, it has been set $D = 1$, treating δ as a fitting parameter. This formulation extends and simplifies the approach in Refs. [79, 80, 85], overcoming some limitations of Kane's solution without introducing numerical stability issues, and in the present case allows for an excellent agreement with the experimental characteristics by tuning only one parameter. The main requirement for its application is a nearly constant $\mathcal{E}$ in the depleted region. This condition is reasonably satisfied in the devices under study, as shown in Fig. 4.6, where the absolute value of the vertical component of the electric field $|\mathcal{E}_z|$ is reported at a bias $V_{\text{bias}} = -3$ V. (The p-n^- junction is located near $z = 7\,\mu$m, and the depleted region extends mainly in the n^- side of the photodiode.)

I.Ion. is customarily included in device simulations by means of a semiempirical

post-processing involving a bias-dependent gain factor $M(V)$

$$J_{\text{dark,I.Ion.}}(V) = M(V) J_{\text{dark}}(V) \tag{4.8}$$

where $J_{\text{dark}}(V)$ is simulated taking into account all other generation mechanisms. The advantage of this approximate approach is to avoid the very intensive computations involved by a self-consistent inclusion of I.Ion. in the DD model, e.g. through the Okuto-Crowell formulation [86]. One of the most popular functional forms for $M(V)$, often referred to as Beck's model, is described and theoretically justified in [87] assuming ballistic transport in the depletion layer. An alternative formulation [88, 89], successfully applied to the description of the dark current of HgCdTe APDs, corresponds to a gain factor having the form

$$M(V) = \exp\left[aV \exp\left(-\frac{V_{\text{th}}}{V}\right)\right] \tag{4.9}$$

which depends on two parameters, a and V_{th}. The authors of [88] remarked that the excess currents often attributed to BTBT at high reverse bias and to TAT at lower reverse bias, could be rather ascribed in part to I.Ion.. This comment also applies to the present study, because the reproduction of the experimental $J_{\text{dark}}(V;T)$ over the whole considered bias range (up to 3 V) by including I.Ion. through Eq. 4.9, and not by including TAT, was successful.

Here it was first assigned $V_{\text{th}} = \mathcal{E}_{\text{th}} w$, where w is the width of the depleted region ($\approx 1\,\mu$m, see also Fig. 4.6) and $\mathcal{E}_{\text{th}} \approx 30$ kV/cm, a value considered typical for HgCdTe [89]. Then a was obtained through an optimization process involving the experimental dark currents and the $J_{\text{dark}}(V)$ values computed by solving the DD model without I.Ion.. A comparison between measurements and simulations is presented in Fig. 4.7 for *set A*, where the best agreement has been observed with $\delta = 210\,\text{V}^{1/2}\text{cm}^{-1/2}$ in the BTBT model, and with the $a(T)$ values reported in the legend. The corresponding results for *set B* are shown in Fig. 4.8, where BTBT has been found to be negligible.

Concerning TAT contributions, they were included in the DD model according to [17], observing that, at least in the considered interval of V_{bias}, they decrease rapidly with increasing temperature. On the contrary, the excess current of the devices under study is fairly insensitive to temperature (see e.g. Figs. 4.7–4.8), which is considered typical of I.Ion. [8, Sec. 14.6.3]. A direct comparison between the effects of the two processes is reported for *set B* in Fig. 4.9 (similar considerations hold for *set A*), where it can be observed that at the highest temperature TAT has become negligible, whereas I.Ion.-like contributions are important.

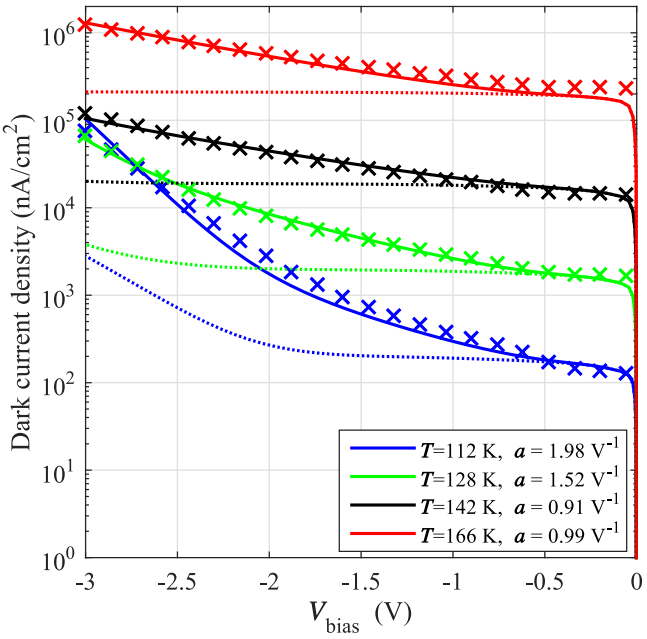

Figure 4.7. *Set A*: simulated $J_{\text{dark}}(V;T)$ considering the contributions of Auger, SRH and radiative GR and BTBT (dotted lines), and with the further inclusion of I.Ion. (solid lines). Crosses indicate experimental values.

4.5 Final remarks

The present study supports the suggestion [22, 46] that, among prerequisites to obtain HOT HgCdTe photodetectors, the reduction of Hg-vacancy density in the material is crucial due to the role of vacancies as SRH recombination centers. Also important is reducing the doping density in the absorber region and using a n-on-p architecture, due to the longer Auger-7 lifetime with respect to Auger-1. TAT processes seem to be negligible in both sets of devices under study, possibly because of the low dopant and defect densities, while I.Ion., whose contribution is sometimes neglected, has been found to be significant, becoming dominant at a reverse bias larger than 1 V. Finally, BTBT has shown a correlation with the doping density of the p-type region, since it is present only in *set A*, where N_A is 4 times higher than in *set B*. This is likely due to the different band bending corresponding to the two doping configurations which, in turn, affects the direct tunneling path.

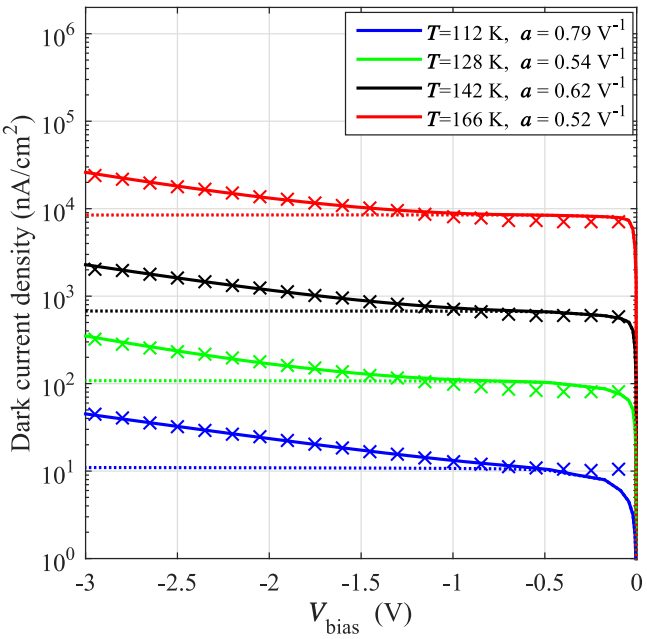

Figure 4.8. *Set B*: simulated $J_{\text{dark}}(V;T)$ considering the contributions of Auger, SRH and radiative GR (dotted lines), and with the further inclusion of I.Ion. (solid lines). Crosses indicate experimental values. BTBT is not required to obtain a satisfactory agreement with experiments.

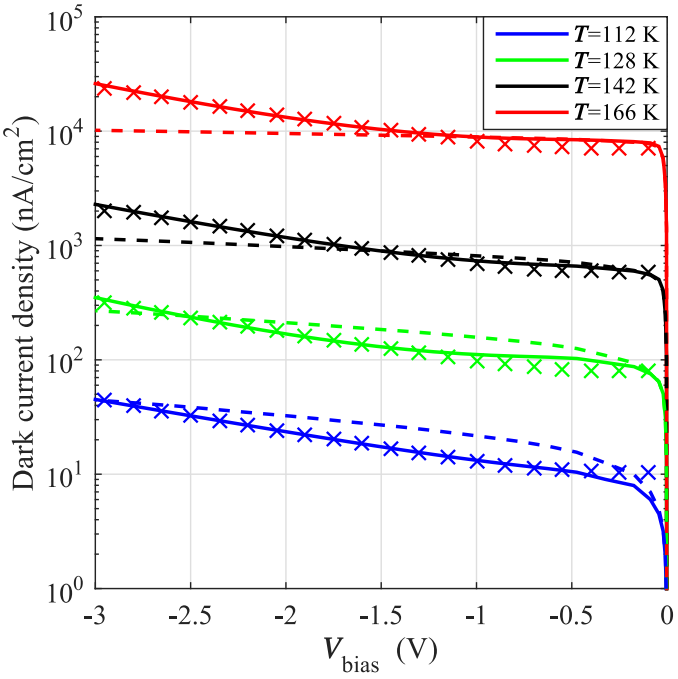

Figure 4.9. *Set B*: simulated $J_{\text{dark}}(V;T)$ considering the contributions of Auger, SRH and radiative GR, and either I.Ion. (solid lines) or TAT (dashed lines). The T-dependence of experimental characteristics (symbols) can be reproduced with I.Ion., not with TAT.

Chapter 5

Comparing FDTD and ray tracing models in HgCdTe photodetectors simulation

This Chapter presents simulations of HgCdTe-based long-wavelength infrared detectors, focusing on methodological comparisons between the FDTD and ray tracing optical models. In the present work, three-dimensional simulations have been performed to determine the absorbed photon density distribution, photocurrent and quantum efficiency (QE) spectra of isolated n-on-p uniform-composition pixels, systematically comparing the results obtained with FDTD and ray tracing.

Here we closely follow a paper we submitted on a special issue of the J. Electron. Mater. [4] (presently accepted for publication, in progress). A summary of the results of this investigation has also been presented at the U.S. Workshop on the Physics and Chemistry of II-VI Materials, Chicago, Illinois (USA), October 5-8, 2015.

5.1 Introduction

Accurate coupled electrical-optical simulations of FPA infrared photodetectors are essential to optimize detectivity and spectral response with respect to epitaxial composition and doping profiles. Several modeling approaches are available to describe light propagation, scattering and absorption in photodetectors, each involving a trade-off between computational cost and accuracy. The propagation of electromagnetic waves in any medium is completely described by the solution of Maxwell's equations, a task that can be addressed by numerical techniques including the FDTD [90–92], the finite element method (FEM) [93, 94], the transfer matrix method (TMM) [92, 95], and the rigorous coupled-Wave Analysis (RCWA) [96, 97]. A much less computationally intensive alternative is provided by ray tracing [98], which is based on classical optics. Unfortunately,

its application to photodetectors may become questionable when the wavelength λ of the illuminating radiation is comparable to the device size [99]; in fact, as the pixel pitch approaches the value of λ, resonance peaks in the quantum efficiency are observed, due to optical cavity effects [100, 101].

In the present work, three-dimensional simulations have been performed to determine the absorbed photon density distribution A_{opt}, photocurrent I_{ph} and quantum efficiency (QE) of n-on-p uniform-composition $Hg_{1-x}Cd_xTe$ back-illuminated LWIR single-pixel photodetectors, systematically comparing the results obtained with ray tracing and FDTD; the latter was chosen among the full-wave methods because of its general applicability and widespread adoption.

The considered photodetector geometry, along with its composition and doping profiles, are presented in section 5.2. The simulation approach is discussed in section 5.3, focusing on the two optical simulation methods we are considering. The simulation results are presented and discussed in section 5.4, and the main findings, advantages and limitations of the two methods are summarized in section 5.5.

5.2 The case study

We considered (Fig. 5.1) a single pixel, $20\,\mu\text{m}\times20\,\mu\text{m}$ wide and $\approx 10\,\mu\text{m}$ thick, with vertical edges. The geometry, doping and composition profiles have been inspired from the literature [102]. A HgCdTe planar layer with uniform composition was doped with an acceptor density $N_A = 2 \times 10^{16}\,\text{cm}^{-3}$, and the photodiode junction was defined by simulating an ion implantation, yielding an error-function-shaped donor density with a peak value $N_D = 10^{18}\,\text{cm}^{-3}$, as described in [3]. The bias contact was connected to a square Au metallization, partly extending over a CdTe passivation layer that covers the upper face of the pixel; a small ground contact was placed on a lateral side. With respect to [102], where $Hg_{0.71}Cd_{0.29}Te$ was employed, we changed the HgCdTe mole fraction to obtain a LWIR detector with a cutoff wavelength $\lambda_c = 9.2\,\mu\text{m}$ at $T = 77\,\text{K}$, where approximately $\lambda_c = 1.24/E_g$ and E_g is the material energy gap in eV. We simplified the passivation layer, considering only CdTe, and we improved convergence and realism by adopting a non-abrupt p-n junction, as already done in [3]. The substrate was not included in the computational box.

5.3 Modeling methods

Simulations have been performed with a commercial 3D numerical simulator [1], which solves the drift-diffusion equations using the finite box method, as described in Section 1.3, employing the HgCdTe material library described in Section 2.2. Composition, doping and temperature dependence of the electronic, transport and optical parameters of

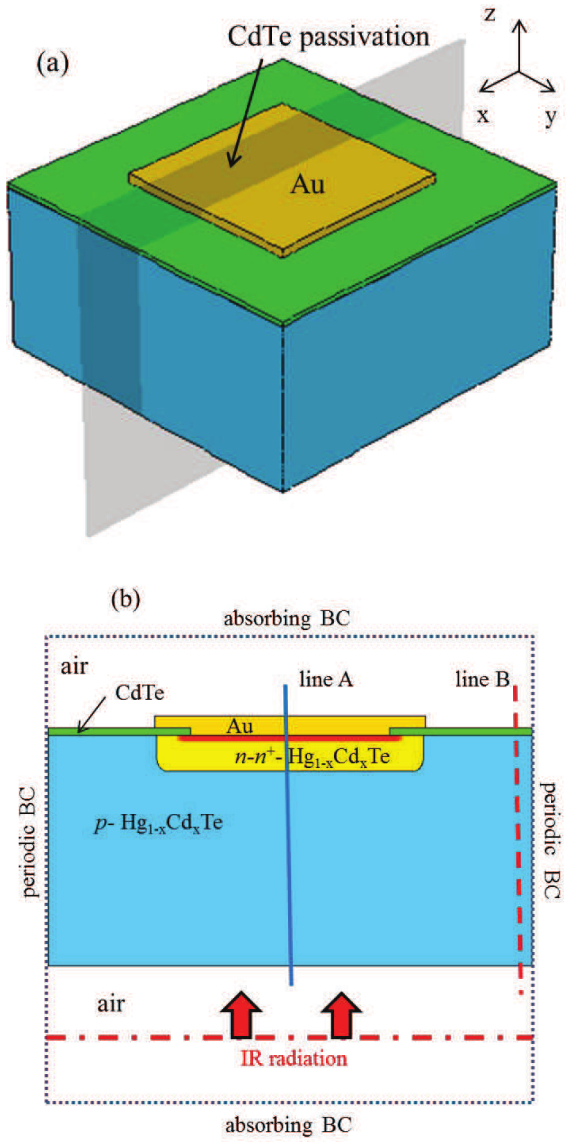

Figure 5.1. (a) 3D single pixel geometry and (b) 2D cut along a vertical symmetry axis. The vertical lines (A, B) mark the 1D cuts along which interesting quantities are plotted in the following. The computational box (dotted line) and optical boundary conditions are also shown.

HgCdTe have been described through the models proposed in [29, p. 1243] for the energy gap E_g, [36, p. 1331] for the electron affinity χ, [34] for electron and hole effective masses, [8, Eq. 14.7] for electron and hole mobilities, and [43, Eqs. 5, 9] for radiative and

Auger lifetimes, where we assumed $F_{12} = 0.2$ for the overlap integral of the Bloch functions and $\gamma = 6$ for the ratio between Auger-7 and Auger-1 intrinsic lifetimes [48, 82]. The SRH recombination processes have been modeled as in [3, 53, 83], considering a lifetime around $2\,\mu$s, and the wavelength- and composition-dependent complex refractive index has been approximated following [55]. Fermi-Dirac statistics and incomplete dopant ionization have been taken into account, with activation energies for HgCdTe alloys estimated according to [8, 57].

The simulated detector, in air and driven in reverse bias (0.1 V), has been illuminated from below with monochromatic plane waves having a propagation vector normal to the illuminated face and an optical power flux $P_0 = 10^{-6}\,\mathrm{W\,cm^{-2}}$. The wavelength of the incident light has been varied in the interval $\lambda = 2 - 12\,\mu$m; hence, the photon flux varies between 10^{13} and $6 \times 10^{13}\,\mathrm{s^{-1}cm^{-2}}$. All simulations have been performed at a temperature $T = 77\,$K. Assuming Auger-limited material (approximate values of Auger, radiative and SRH lifetimes are $\tau_{\text{Auger}} = 0.1\,\mu$s, $\tau_{\text{rad}} = 1\,\mu$s, $\tau_{\text{SRH}} = 2\,\mu$s, respectively), the minority carrier diffusion length is $L_{\text{diff},n} \approx 80\,\mu$m, allowing photogenerated carriers to reach the depleted region and to be efficiently collected.

The pixel structure has been discretized with a meshing tool which generates a denser grid in regions where gradients of current density, electric field, free charge density and material composition are present. All the results reported in the following have been computed with a mesh including about 6×10^5 elements. The j-th mesh element, of volume V_j, absorbs an amount of radiation $A_{\text{opt},j}$, where the spatial distribution $A_{\text{opt}}(x, y, z) = A_{\text{opt},j}/V_j$ is computed according to the selected optical model (either ray tracing or FDTD, see sections 5.3.1 and 5.3.2) and x, y, z are the coordinates of the j-th element barycenter. In its turn, $A_{\text{opt}}(x, y, z)$ enters the continuity equations for the electron and hole current densities $J_{n,p}$ (where $R_{n,p} - G_{n,p}$ is the net recombination rate in absence of carrier photogeneration and q is the elementary charge)

$$\begin{aligned} \nabla \cdot J_n &= q(R_n - G_n - G_{\text{opt}}) + q\frac{\partial n}{\partial t} \\ -\nabla \cdot J_p &= q(R_p - G_p - G_{\text{opt}}) + q\frac{\partial p}{\partial t} \end{aligned} \tag{5.1}$$

through the optical generation rate $G_{\text{opt}} = \eta A_{\text{opt}}$ due to interband optical absorption, which depends also on the model for the quantum yield η, defined as the fraction of absorbed photons which are converted to photogenerated electron-hole pairs. Since the focus of this work is on the comparison between two alternative optical modeling approaches rather than on material- or device-related effects affecting η (e.g., free-carrier absorption or other optical nonlinearities due to population effects), we have set $\eta = 1$ in all our simulations.

5.3.1 Ray tracing

In ray tracing models, rays are used to represent the propagation of electromagnetic waves according to geometrical optics. Fresnel's and Snell's laws of reflection and refraction determine the paths travelled by rays as they interact with different materials. Usually, a large number of rays is traced within an optical system in order to simulate its behavior; in the present simulations, we considered ≈ 2500 rays uniformly spaced in the xy plane, entering the pixel with initial wavevector parallel to the z-axis. Each ray i transmitted into the j-th mesh element corresponds to an optical rate intensity $P_{j,i}$ (number of photons entering the element per second). The ensuing absorbed photon density in V_j is

$$A_{\text{opt},j,i} = \frac{P_{j,i}}{V_j}\left(1 - e^{-\alpha_j L_{j,i}}\right) \tag{5.2}$$

where $L_{j,i}$ is the path length travelled by ray i into the j-th element, α_j is the absorption coefficient α averaged on the j-th element, and α is related to the imaginary part κ of the complex refractive index $\hat{n} = n + i\kappa$ as $\alpha = 4\pi\kappa/\lambda$. Summation over all the ray histories produces the mesh-discretized distribution $A_{\text{opt},j}(x, y, z)$, and hence the optical generation rate $G_{\text{opt}}(x, y, z)$ entering the continuity equations (5.1). This method manages all the internal back-reflections, but without keeping into account interference effects.

5.3.2 FDTD

The FDTD optical method describes the light propagation in a medium according to physical optics: combining divergence and curl Maxwell's equation, one finds for the electric and magnetic fields vectors $\mathbf{E}$ and $\mathbf{H}$ (assuming non-magnetic media)

$$\begin{aligned} \frac{\partial \mathbf{H}}{\partial t} &= -\frac{1}{\mu_0}\nabla \times \mathbf{E} \\ \frac{\partial \mathbf{E}}{\partial t} &= \frac{1}{\epsilon}\nabla \times \mathbf{H} - \frac{\sigma}{\epsilon}\mathbf{E}, \end{aligned} \tag{5.3}$$

where μ_0 is the vacuum magnetic permeability, ϵ and σ are the electric permittivity and conductivity. The FDTD approach discretizes and solves Eqs. (5.3) on a cubic grid known as the Yee grid [90] shown in Fig. 5.2 (algorithmic details can be found e.g. in [92]).

Within the FDTD simulation framework, the material properties are represented by the real part of ϵ and σ, computed from the real part of the material refractive index n and the absorption coefficient α, according to

$$\epsilon = n^2 - \kappa^2 = n^2 - \left(\frac{\alpha\lambda}{4\pi}\right)^2 \tag{5.4}$$

$$\sigma = \frac{n\alpha}{\mu_0 c}, \tag{5.5}$$

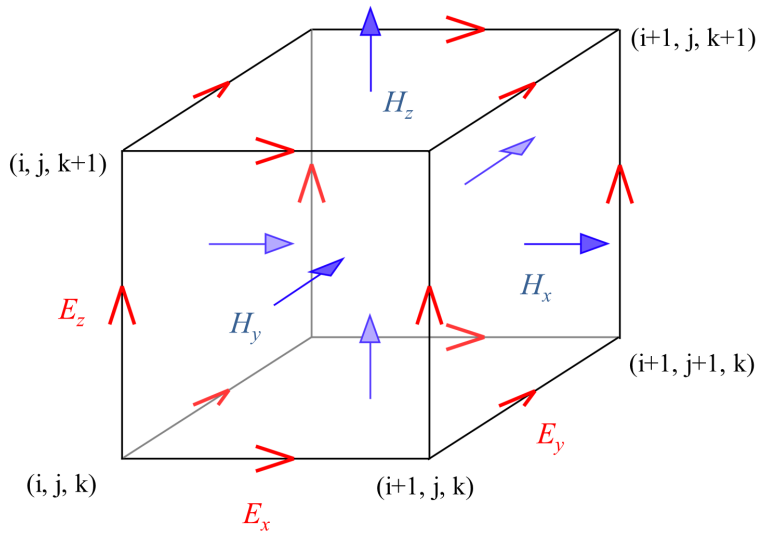

Figure 5.2. Schematic representation of the Yee grid [90]. The position of the electric and magnetic fields in relation to each other implicitly enforces the boundary conditions contained in Maxwell's equations. The electric and magnetic fields are solved at alternating grid points in space.

where c is the speed of light in vacuum. Once the solution of Maxwell's curl equation is available, the absorbed photon density A_{opt} can be evaluated as the divergence of the time-averaged Poynting vector $\langle S \rangle$ [95, 100]

$$A_{\text{opt}} = -\frac{\nabla \cdot \langle S \rangle}{h\nu} = \frac{1}{2h\nu}\sigma|\mathbf{E}|^{\mathbf{2}} \tag{5.6}$$

where $h\nu$ is photon energy. FDTD properly manages interference effects due to back-reflections inside the pixel. When wide contact metallization are present (usually characterized by high infrared reflectance) or the absorber material is semi-transparent (e.g. when $\lambda \approx \lambda_c$), interference is expected to trigger cavity effects, likely to be observed in the QE spectra [100].

According to FDTD, Maxwell's equations are solved at every node of the mesh. A trade-off between accuracy and computational cost must be faced when selecting the number of nodes per wavelength: in the present study, we used 10 and 20 nodes per wavelength in the xy plane and along z, respectively. We imposed Neumann electric boundary conditions (BC) for all not-contacted faces, whereas contacts were considered ideally ohmic. The computational box includes air layers located above and below the pixel, and the optical BC along the upper and lower sides of the box are absorbing (in FDTD, this is obtained with convolutional perfectly matching layers [91]). Periodic optical BC along the lateral sides of the computational box were chosen both for ray tracing and FDTD (see Fig. 5.1(b)). With this choice, a periodic pixel array in the xy plane is simulated, which

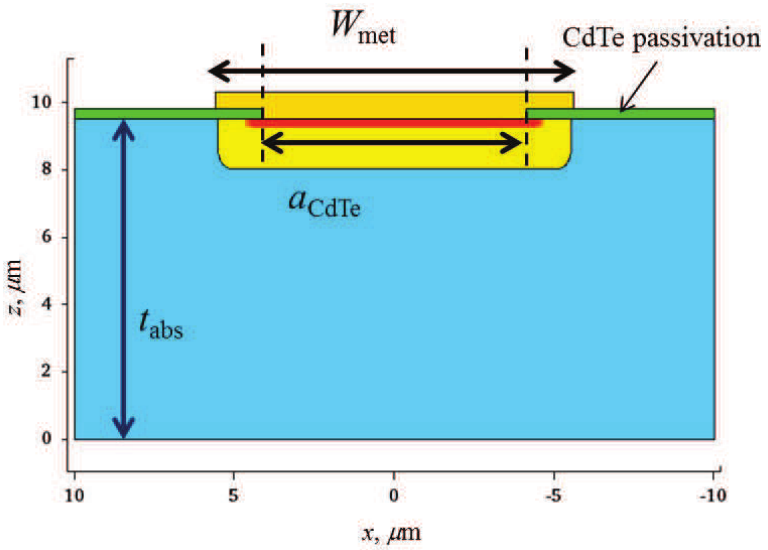

Figure 5.3. Geometrical parameters considered in the present study: absorber thickness t_{abs}, CdTe aperture a_{CdTe}, metallization width W_{met}.

allows FDTD to describe inter-pixel interference effects.

5.4 Simulation results

In the present work we studied the effects of three geometrical parameters (see Fig. 5.3): the absorber thickness t_{abs}, the CdTe aperture a_{CdTe} (which corresponds to the width of the bias contact), and the metallization width above the passivation layer W_{met}. We will not present a complete study of all parameter combinations, but we will focus on a set of significant examples.

In the **first example** we varied the absorber length in the interval $t_{abs} = 9.5 - 10.5\,\mu$m, with constant $a_{CdTe} = 6\,\mu$m and $W_{met} = 11\,\mu$m. The metallization covers 1/3 of the pixel top surface, and back-reflections due to metallization are not intense enough to trigger significant cavity effects. The simulated QE spectra shown in Fig. 5.4 indicate that the ray tracing results are good approximations of the FDTD ones, for all considered wavelengths. Moreover, the absorber thickness t_{abs} does not play a significant role. Weak cavity effects appear in FDTD spectra only when the material is almost transparent ($\lambda \approx 10\,\mu$m); this could be investigated by extending the simulation study to longer wavelengths.

In the **second example** we varied the metallization width in the interval $W_{met} = 11 - 16\,\mu$m, with constant $a_{CdTe} = 6\,\mu$m and $t_{abs} = 9.5\,\mu$m (see Fig. 5.5). At its widest value, the metallization covers 2/3 of the pixel top surface, and back-reflections are quite stronger with respect to the previous example. The simulated QE spectra for the case $W_{met} =$

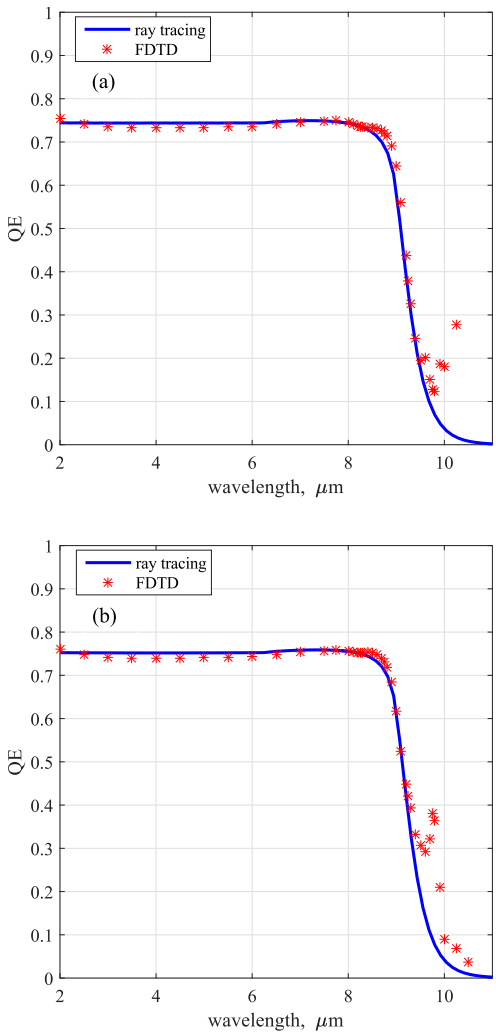

Figure 5.4. First example: QE spectra, simulated with ray tracing and FDTD, for $W_{\text{met}} = 11\,\mu$m, $a_{\text{CdTe}} = 6\,\mu$m. (a) $t_{\text{abs}} = 9.5\,\mu$m, (b) $t_{\text{abs}} = 10.5\,\mu$m.

$16\,\mu$m, reported in Fig. 5.5(b), demonstrate that ray tracing is not a good approximation of FDTD, except for $\lambda \ll \lambda_c$. Cavity effects clearly appear in the FDTD spectra around the cutoff, where the material is transparent enough to trigger them. In addition, if λ_c is defined as the wavelength where the QE is reduced to half its peak value, the presence of a strong resonance also brings a redshift of the λ_c predicted by FDTD with respect to ray tracing; the shift is about $0.2\,\mu$m, not negligible when calibrating a detector material composition.

In the **third example** we verified that the effect of the CdTe aperture is neglegible in the present structure. Setting $t_{\text{abs}} = 9.5\,\mu$m and $W_{\text{met}} = 16\,\mu$m, we varied the CdTe

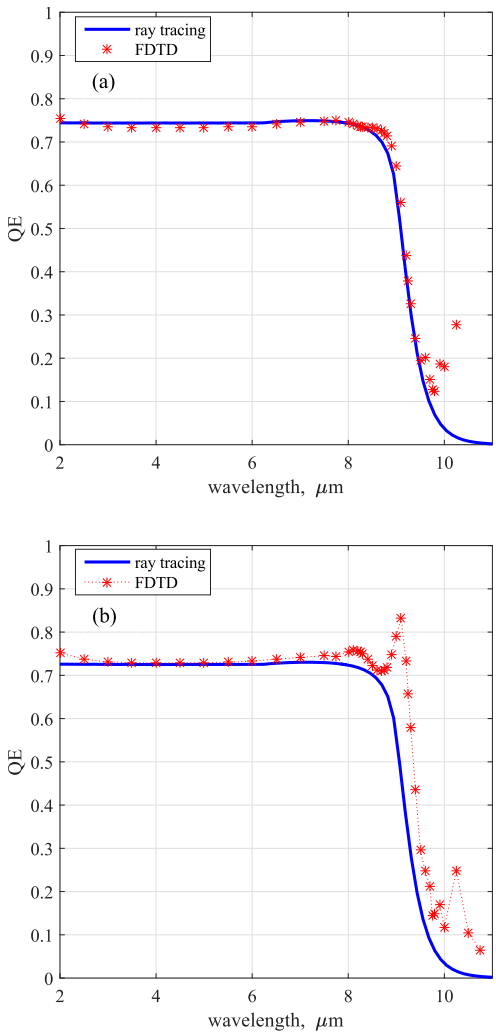

Figure 5.5. Second example: QE spectra, simulated with ray tracing and FDTD, for $t_{\text{abs}} = 9.5\,\mu$m, $a_{\text{CdTe}} = 6\,\mu$m. (a) $W_{\text{met}} = 11\,\mu$m, (b) $W_{\text{met}} = 16\,\mu$m.

aperture considering the cases $a_{\text{CdTe}} = 2,\ 4,\ 6\,\mu$m, always observing significant cavity effects with FDTD (see Fig. 5.6).

Two additional comments about the differences between ray tracing and FDTD simulations are worth making. First, it is interesting to compare the A_{opt} distributions predicted by the two methods at different wavelengths, for $t_{\text{abs}} = 9.5, \mu$m, $a_{\text{CdTe}} = 6\,\mu$m, and $W_{\text{met}} = 11\,\mu$m. In Fig. 5.7 we report the A_{opt} distributions at $\lambda = 9\,\mu$m (i.e., close to the cutoff) along the vertical symmetry plane shown in Fig. 5.1(a). The dissimilarity between the results of the two methods is apparent. If we consider shorter λ values, the differences are gradually attenuated. In Figs. 5.8–5.9, vertical cuts along lines A and B

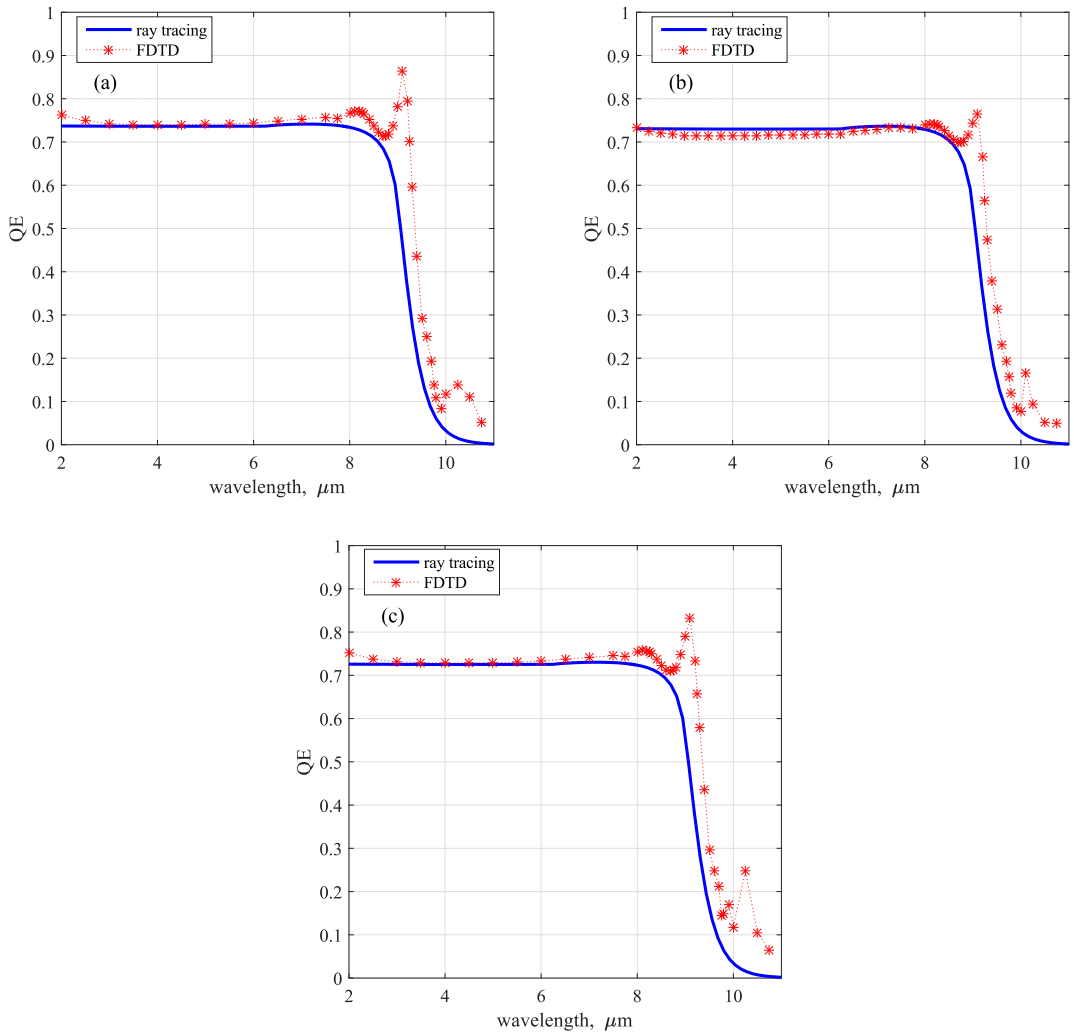

Figure 5.6. Third example: QE spectra, simulated with ray tracing and FDTD, for $t_{abs} = 9.5\,\mu$m, $W_{met} = 16\,\mu$m. (a) $a_{CdTe} = 2\,\mu$m, (b) $a_{CdTe} = 4\,\mu$m, (c) $a_{CdTe} = 6\,\mu$m.

(see definitions in Fig. 5.1(b)) are reported at $\lambda = 8,\ 8.5$ and $9\,\mu$m, and it can be observed that the ray tracing and FDTD distributions become progressively closer as λ is reduced. Along line B, where the metallization back-reflection effects are smaller, differences are less pronounced also near the cutoff.

As a second point, it is interesting to observe how A_{opt} changes if the metallization widens to $W_{met} = 16\,\mu$m, keeping unchanged all other parameters, as in the second example (see Fig. 5.5). The ray tracing and FDTD results at the resonance wavelength $\lambda = 9.1\,\mu$m are compared in Fig. 5.10. Strong peaks in the FDTD A_{opt} distribution can be observed on the pixel lateral sides (see Fig. 5.10(b), at $x \approx \pm 10\,\mu$m and $z \approx 5 - 9\,\mu$m).

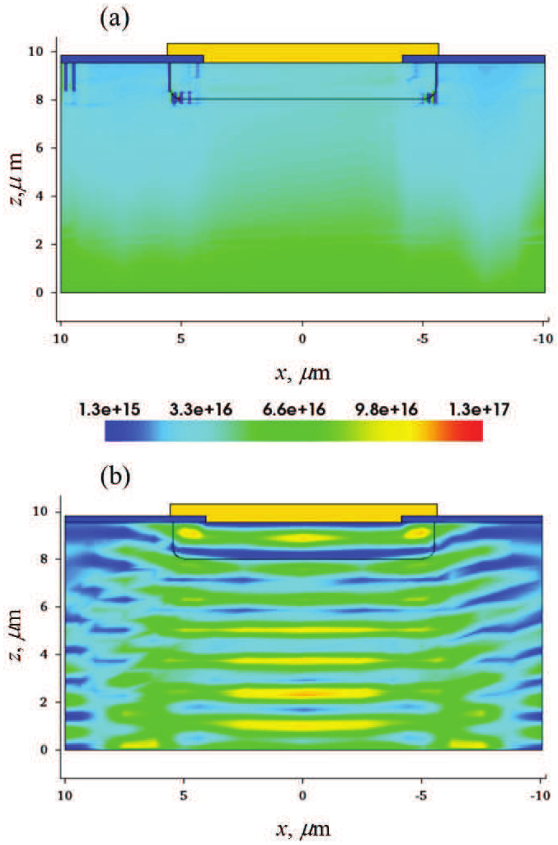

Figure 5.7. A_{opt} distribution ($\text{cm}^{-3}\text{s}^{-1}$) obtained with (a) ray tracing and (b) FDTD, across the vertical symmetry plane shown in Fig. 5.1, for $t_{\text{abs}} = 9.5\,\mu$m, $a_{\text{CdTe}} = 6\,\mu$m, $W_{\text{met}} = 11\,\mu$m at $\lambda = 9\,\mu$m.

In fact, since periodic BC are used along the lateral sides of the computational box, the solution of Maxwell's and transport equations takes into account that the simulated pixel is surrounded on every side in the xy plane by identical pixels. As the metallization becomes wider, its separation from the metallization of its neighbors becomes smaller (it is $9\,\mu$m in the nominal case $W_{\text{met}} = 11\,\mu$m, but just $4\,\mu$m when $W_{\text{met}} = 16\,\mu$m), and we may observe corresponding interference peaks in the absorbed photon density. If the pixel spacing is increased, the resonance strength weakens, and so do the A_{opt} peaks along the lateral sides. All this is not observed with ray tracing, that again underestimates the QE close to λ_c.

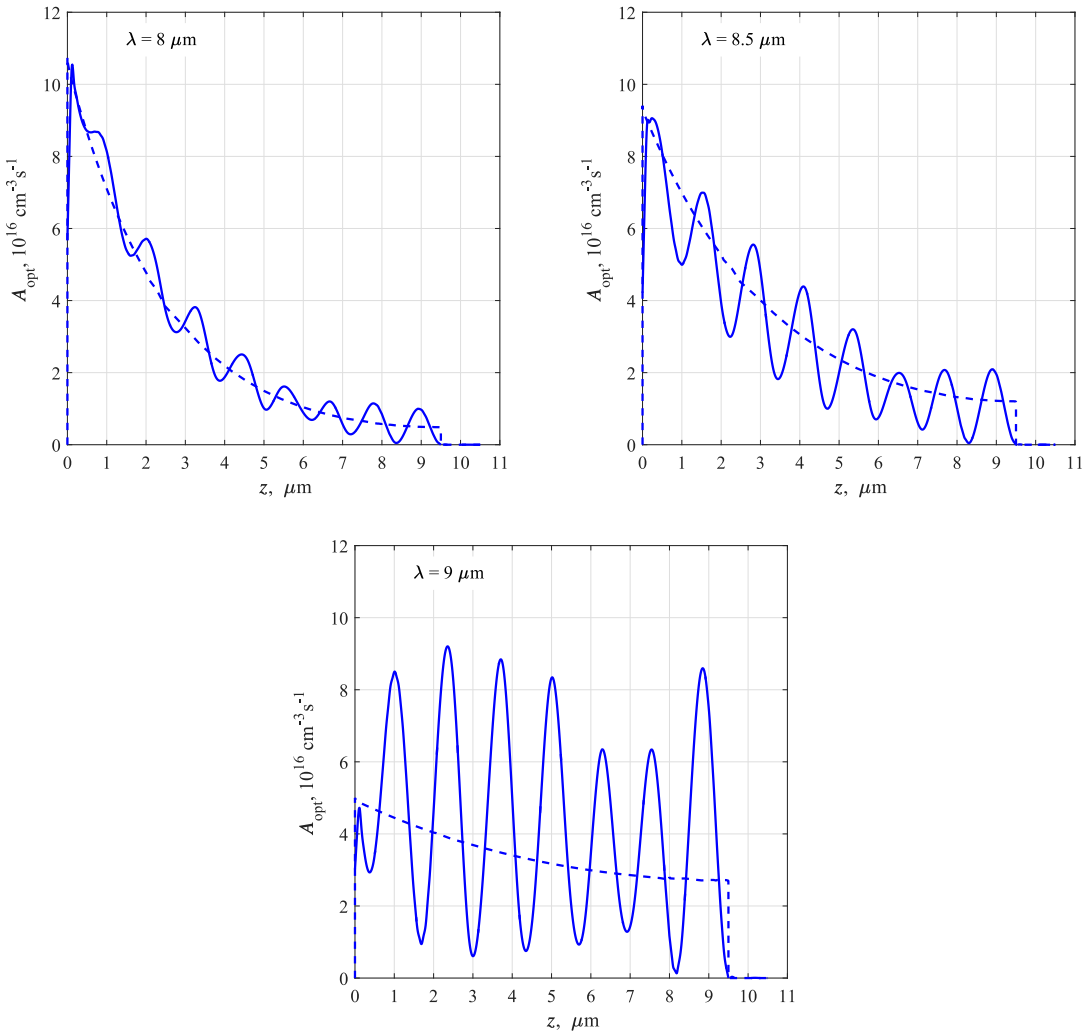

Figure 5.8. A_{opt} obtained with ray tracing (dashed lines) and FDTD (solid lines) along line A (see Fig. 5.1(b)), for $\lambda = 8,\ 8.5,\ 9\,\mu$m.

5.5 Conclusions

In this work we compared A_{opt} distributions and QE spectra obtained by ray tracing and FDTD in planar HgCdTe FPA single pixels, illuminated with monochromatic plane waves. The effect of pixel neighbors was included by imposing periodic electrical and optical BC. Our goal was to assess the possibility to replace FDTD with ray tracing, which is much less computationally intensive. We found that, in general, ray tracing and FDTD yield different results: FDTD predicts A_{opt} interference patterns that, in some conditions, produce QE spectral features which cannot be observed with ray tracing. The customary

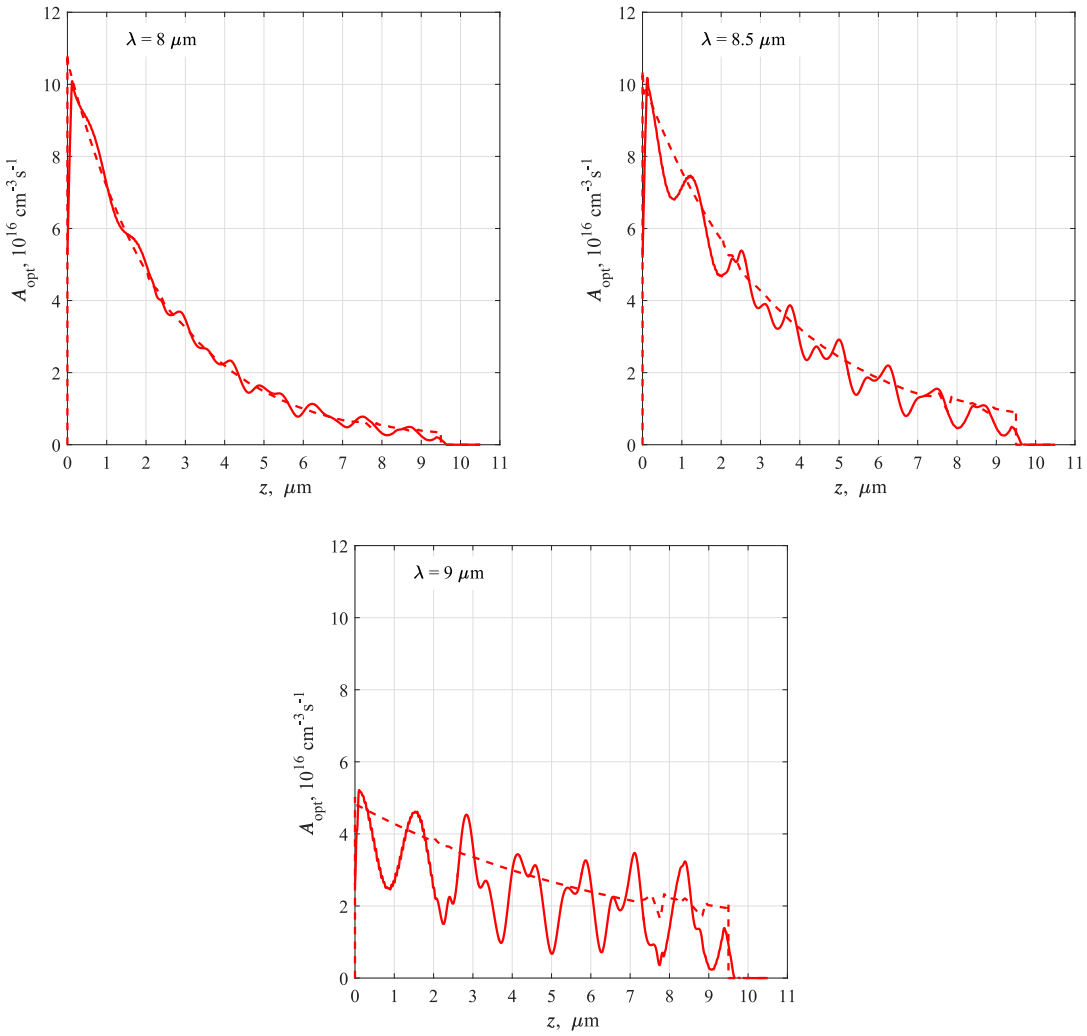

Figure 5.9. A_{opt} obtained with ray tracing (dashed lines) and FDTD (solid lines) along line B (see Fig. 5.1(b)), for $\lambda = 8,\ 8.5,\ 9\,\mu$m.

approach to define λ_c as the wavelength where QE $= 0.5\text{QE}_{\text{max}}$ could also be questioned, since it is significantly affected by the adopted optical model.

However, when internal back-reflections are small, ray tracing can be reliably used to predict QE and photocurrent spectra in our devices. The greatest sources of back-reflections are the interfaces with metallizations, which are assumed to be totally reflecting (reflectivity $R = 1$). By comparison, the reflectivity of a HgCdTe–CdTe interface at $\lambda = 9\,\mu$m is $R \approx 0.022$ for normal incidence (since $\hat{n}_{\text{HgCdTe}} \approx 3.501 + i\,0.062$ and $\hat{n}_{\text{CdTe}} \approx 2.6$), unable to produce appreciable cavity effects. Hence, the most significant QE spectral features are observed in pixels where the metallization covers most of the top

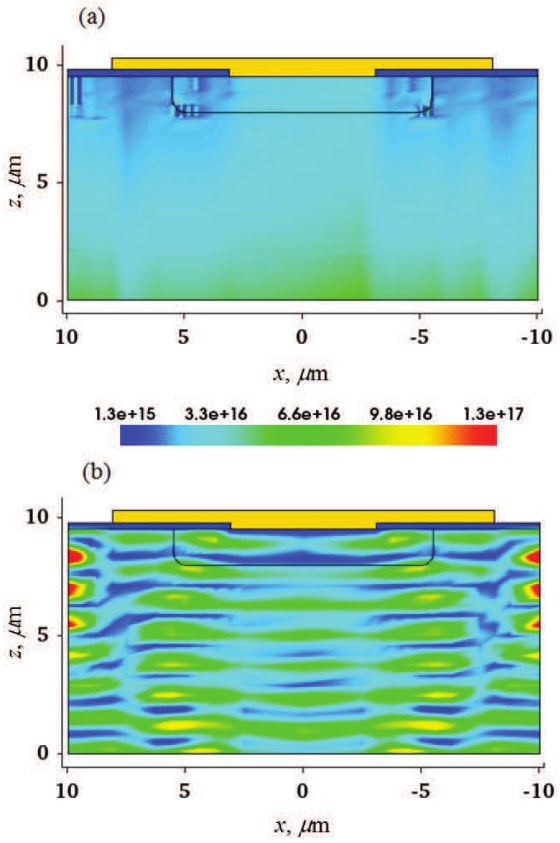

Figure 5.10. A_{opt} distribution ($\text{cm}^{-3}\text{s}^{-1}$) obtained with (a) ray tracing and (b) FDTD, across the vertical symmetry plane shown in Fig. 5.1, for $t_{\text{abs}} = 9.5\,\mu$m, $a_{\text{CdTe}} = 6\,\mu$m, $W_{\text{met}} = 16\,\mu$m at $\lambda = 9.1\,\mu$m.

surface and inter-pixel interference effects are important.

The present conclusions could be revised in a future work considering non-monochromatic illumination and extended sources. As a first step, one could take into account that an extended source produces incoherent light even if monochromatic, since it contributes to the illumination with plane waves propagating along different directions; as a result, one could expect a decrease of cavity effects in the QE spectra. Incoherent extended sources could be simulated by describing the optical field either as the sum of plane waves whose wavevectors are contained in a cone, or as a gaussian beam with appropriate parameters. Both approaches are feasible with FDTD (at significantly higher computational cost with respect to the present study), and will be considered in future investigations.

5.6 Future work

All infrared imaging applications will benefit from reductions in the size, weight, and power of detecting arrays and their supporting subsystems. Recent progress, in small infrared detector fabrication, has raised interest in determining the minimum useful detector size. It has been demonstrated a minimum useful detector size of 2 μm for the mid-wave infrared region (MWIR) and 5 μm for the long-wave infrared region (LWIR) when coupled with an $F/1$ optical system under high signal-to-noise ratio conditions [103–105].

Under this assumption, a preliminary work is in progress at the present time, considering a 3×3 pixel array, whose single pixel architecture is still a $n^+ - n - p$ photodiode, with acceptor density concentration around 5×10^{15} cm^{-3} in the p-region and donor concentration in the $n^+ - n$ region equal to $1 \times 10^{17} \div 5 \times 10^{14}$ cm^{-3} with an error-function graded profile. An example of the absorbed photon density profile is reported in Fig. 5.11, for a HgCdTe LWIR detector with 5 μm pixel pitch and 2 μm thick absorber. The illumination is obtained by monochromatic plane wave with optical power flux $P = 10^{-5}$ W cm^{-2}. This very simple architecture is inspired by Fig. 35(a) of Ref. [105], and it is qualitatively similar to the single-pixel structure we presented at the II-VI U.S. Workshop in Oct. 2015 and described in this chapter [4]. Composition and doping profiles are standard and inspired by Ref. [104].

At the cutoff wavelength, the maximum light absorption occurs in the space between two pixels, confirming what we observed in the single-pixel approximation with Neumann's boundary conditions imposed on its lateral sides. Next steps will be the study of the QE spectra under uniform illumination (monochromatic plane waves) and under illumination obtained by Gaussian beam shining the central pixel.

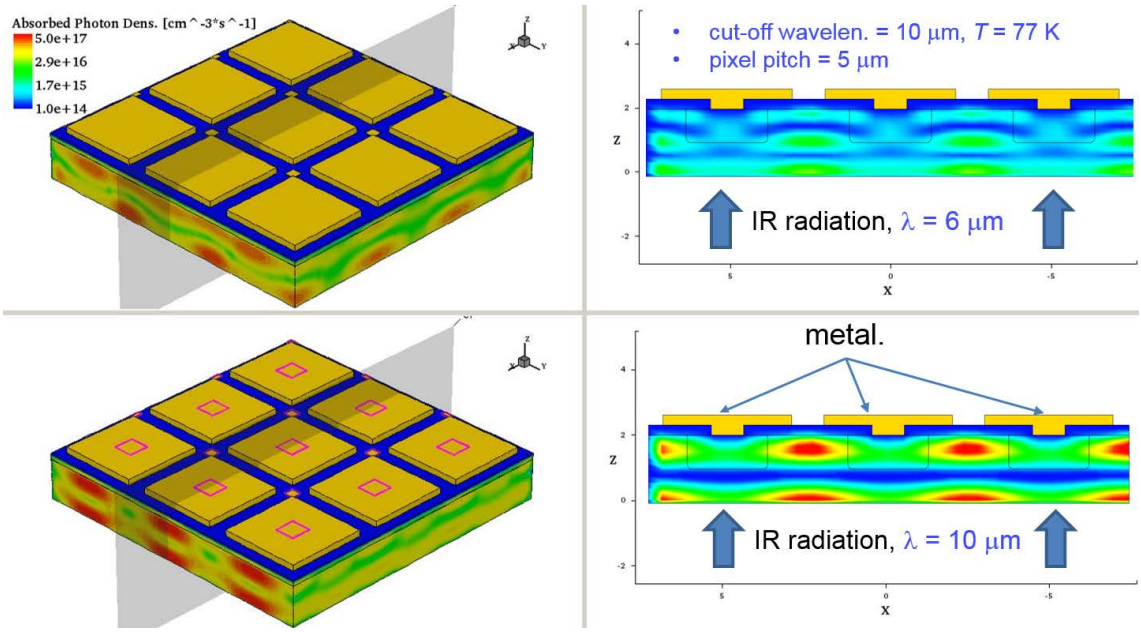

Figure 5.11. A_{opt} obtained with FDTD, for a 3×3 LWIR pixel array with 5 μm pixel pitch (preliminary design).

Part II

Wide band gap devices

Chapter 6

Electron capture time in quantum wells: general theory

Another important topic of industrial research in semiconductor physics deals with nitride-based light-emitting diodes (LEDs). From automotive to streetlights, from lights in our houses to the displays of TVs and smartphones, LED-based technology is making its way in the market. This proliferation would have been impossible without GaN-based LEDs, whose invention by Isamu Akasaki, Hiroshi Amano and Shuji Nakamura has been rewarded with the 2014 Nobel Prize in Physics. Nevertheless, GaN-based LEDs performance is limited by a reduction (droop) of their IQE as the driving current density is increased beyond $\approx 10\,\mathrm{A/cm^2}$, whose physical origin is still under intense debate. Invoked explanations include Auger recombination, electron leakage from the active region, and poor hole injection into the active region (a review can be found in Ref. [106]).

In case of quantum-well (QW) based LEDs, one of the possible issue could be a supposed non-efficient carrier capture into the QW [107], hence in this Chapter, that opens the second part of the present thesis, we present a quantum model, based on the propagator formalism, that allowed to obtain the electron capture time and hot-electron intraband relaxation time in a quantum well for longitudinal optic (LO) phonon emission. The proposed approach produced a closed-form expression for capture time as a function of the carrier densities in QW and barrier states, a topic often neglected in literature, despite its fundamental interest in semiconductor physics. Its application to simple two-population rate equations allowed to reproduce available experimental data with excellent agreement, offering an accurate yet practical alternative to the usual approximation of a constant capture time in modeling light-emitting diodes and lasers.

The formalism is developed and applied to photonic devices in two steps:

1. In this Chapter, after a short recap concerning QWs, some introductory notions will be recalled about the concept of *complex self-energy* and its relation with the

capture of carriers in a QW.

Then, a quantum model for carrier capture in quantum wells is presented, as an extended version of a paper we published in the J. Appl. Phys., vol. 114, p. 053704, Aug. 2013 [5]. It allowed to obtain electron capture and hot electron intraband relaxation times in a quantum well for LO-phonon emission. The model has been applied first to III-V materials, because of the great number of numerical results available in literature for these materials, in view of comparisons. The novelties of the present approach are substantally a) the analytic approach, allowing for a more staighforward understanding of the involved physics, and b) the inclusion of carrier population effects in the model.

2. In Chapter 7 a revision of this model is presented, allowing to obtain a more compact analytic formulation. The model has been applied to InGaN/GaN multi-quantum-well (MQW) LEDs: employing simple two-population rate equations, the model allowed to reproduce literature experimental data with excellent agreement, potentially useful in modeling light-emitting diodes and lasers. The capture process through LO-phonon emission has been found very efficient also for high values of carrier density.

   The concepts here presented substantially follow a paper [6] we published in 2015 in Phys. Status Solidi B, vol. 252, no. 5, pp. 971-976, 2015.

Before going into details, a short introduction about the QW concept, generalities about carrier capture in QW and the concept of complex self-energy will be recalled.

6.1 A brief recall about the QWs

An electron state of mass m_0 in free space is represented in the ordinary space by a plane wave $\psi_{\boldsymbol{k}}(\boldsymbol{r}) = C \exp(i\boldsymbol{k}\cdot\boldsymbol{r})$, solution of the stationary Schrödinger equation

$$-\frac{\hbar^2\nabla^2}{2m_0}\psi_{\boldsymbol{k}}(\boldsymbol{r}) = E_{\boldsymbol{k}}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \tag{6.1}$$

where $\boldsymbol{r}$ is the position vector, $\boldsymbol{k}$ the wavevector and C a normalization constant. The eigenvalue $E_{\boldsymbol{k}}$ is given by the relation parabolic in $|\boldsymbol{k}|$

$$E_{\boldsymbol{k}} = \frac{\hbar^2|\boldsymbol{k}|^2}{2m_0}. \tag{6.2}$$

On the contrary, in a periodic potential of a bulk crystal $U_{\text{ion}}(\boldsymbol{r})$, the periodic lattice potential changes dramatically the electron wavefunction, as the Schrödinger equation becomes

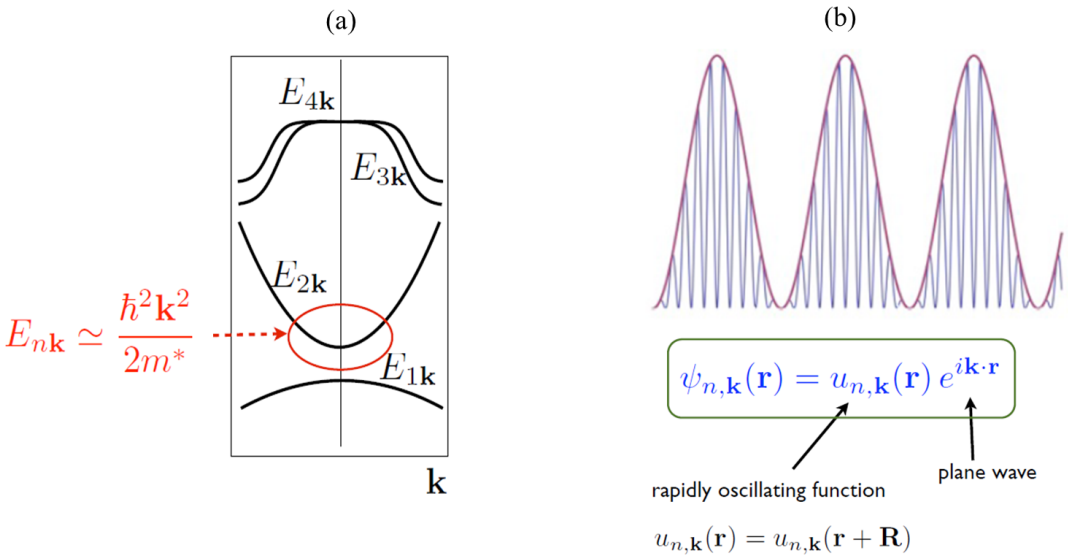

Figure 6.1. (a) Band structure around the Γ point for a typical direct band bap semiconductor. (b) Eigenfunctions of the Hamiltonian operator (electron in the ionic periodic potential).

$$\left(-\frac{\hbar^2\nabla^2}{2m_0} + U_{\text{ion}}(\boldsymbol{r})\right)\psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = E_{n,\boldsymbol{k}}\psi_{n,\boldsymbol{k}}(\boldsymbol{r}) \tag{6.3}$$

whose eigenfunctions read (Block's theorem) as the product of a rapidly oscillating function $u_{n,\boldsymbol{k}}$ reflecting the periodicity of the crystal potential, times a plane wave, the so-called "envelope function":

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = u_{n,\boldsymbol{k}}(\boldsymbol{r})\exp\left(i\boldsymbol{k}\cdot\boldsymbol{r}\right) \tag{6.4}$$

The difference from this "envelope function" and the plane wave of the electron in free space resides in $\boldsymbol{k}$ itself, that now is bounded, since $|\boldsymbol{k}| < \pi/a$, where a is the lattice constant. Also the dispersion relation for the eigenvalues $E_{n,\boldsymbol{k}}$ is quite different from the free space case and represents the so-called band-structure. The index n is a quantum number, representing the band-index, and in a typical direct band gap semiconductor, in proximity of the Γ point of the crystal reciprocal space [15], the band structure can look like in Fig. 6.1.

Around the Γ point the dispersion relation $E_{n,\boldsymbol{k}}$ may often be approximated by a parabola, provided to substitute m_0 with an *effective mass* m in Eq. (6.3) proportional to the inverse of the $E_{n,\boldsymbol{k}}$ curvature.

Now, let us consider a simple planar InGaN/GaN heterostructure, as shown in Fig. 6.2, assuming valid the effective mass approximation, around the Γ point. The InGaN layer

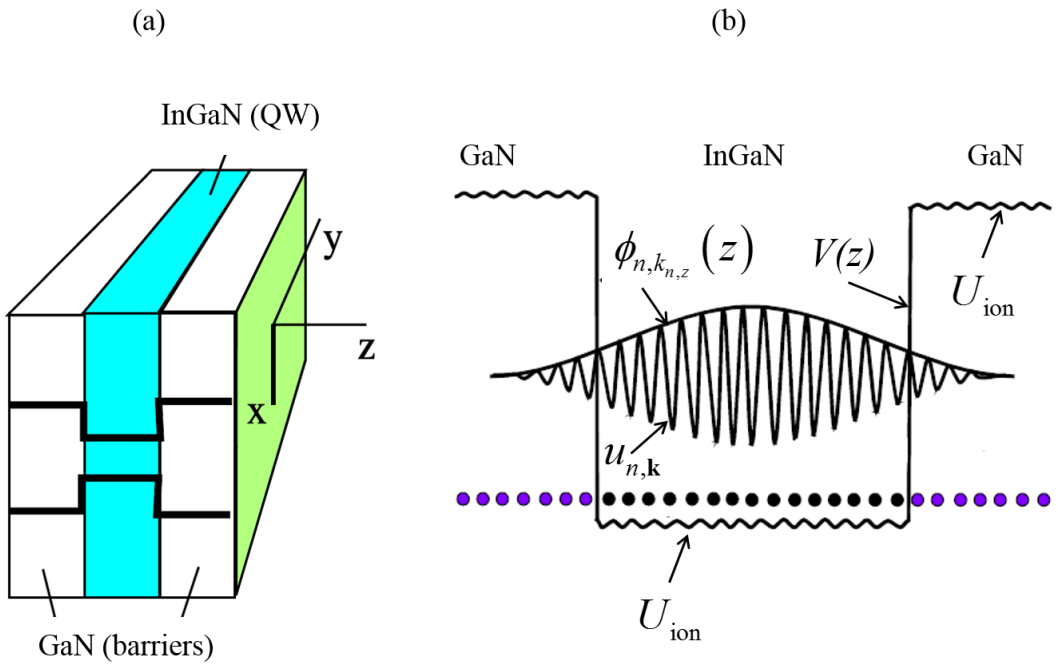

Figure 6.2. (a) InGaN/GaN typical heterostructure scheme. (b) electronic Bloch's state state in the QW. The dots represent the ionic lattice.

– a few nanometers wide – constitutes a 2D reduced dimensional structure, namely a "quantum well", embedded into GaN bulk barriers. The crystal atoms produce a periodic potential $U_{\text{ion}}(\boldsymbol{r})$ and the heterointerfaces produce the confining potential $V(z)$. The 3D character of the structure leads to a 2D equation for the x, y directions and a 1D equation for the z direction, with some major differences among them and with respect to the bulk case:

- electronic states in the barrier are still Bloch's wavefunctions $u_{n,\boldsymbol{k}} \exp(i\boldsymbol{k}\cdot\boldsymbol{r})$, and the envelope wavefunction is a plane wave;
- electronic states in the QW are still Bloch's states, but the envelope wavefunction can be factorized as $\exp(ik_x x)\exp(ik_y y)\phi_{n,k_{n,z}}(z)$, where k_x, k_y and $k_{n,z}$ are the components of $\boldsymbol{k}$ along the reference axes;
- in x, y directions the QW envelope wavefunctions are still plane waves (in the QW plane, electrons are free-moving particles along x, y directions);
- eigenvalues and eigenfunctions $\phi_{n,k_{n,z}}(z)$ can be found as solutions of the Schrödinger equation:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + V(z)\right)\phi_{n,k_{n,z}}(z) = E_n\phi_{n,k_{n,z}}(z) \tag{6.5}$$

where the index n is a quantum number distinguishing the allowed discrete QW states (subbands) and $k_{n,z}$ is fixed by the quantization condition; the eigenvalue E_n is given by $E_n = \hbar^2 k_{n,z}^2/(2m)$;

- in the z direction, the envelope wavefunction $\phi_{n,k_{n,z}}(z)$ describes a z-confined state, with a probability density $\phi^*_{n,k_{n,z}}(z)\phi_{n,k_{n,z}}(z)$ characterized by a more or less pronounced peak in correspondence of the QW, according to the profile of $V(z)$, the energy eigenvalue E_n and the effective mass value m. Fig. 6.2(b) shows a possible Bloch wavefunction for the z-direction;

- the total energy of the state, measured from the QW bottom, in the effective mass approximation is given by

$$E_{n,\boldsymbol{k}} = E_n + \frac{\hbar^2\left(k_x^2 + k_y^2\right)}{2m}. \tag{6.6}$$

Carrier capture in QW is often visualized in a naive fashion as illustrated in Fig. 6.3: this way can be considered a pictorial and self-explaining way to consider a capture process, but it is very far from what happens in a semiconductor, since electrons must be described as quantum states, and the motion of an electron in QW systems can never be considered as classical.

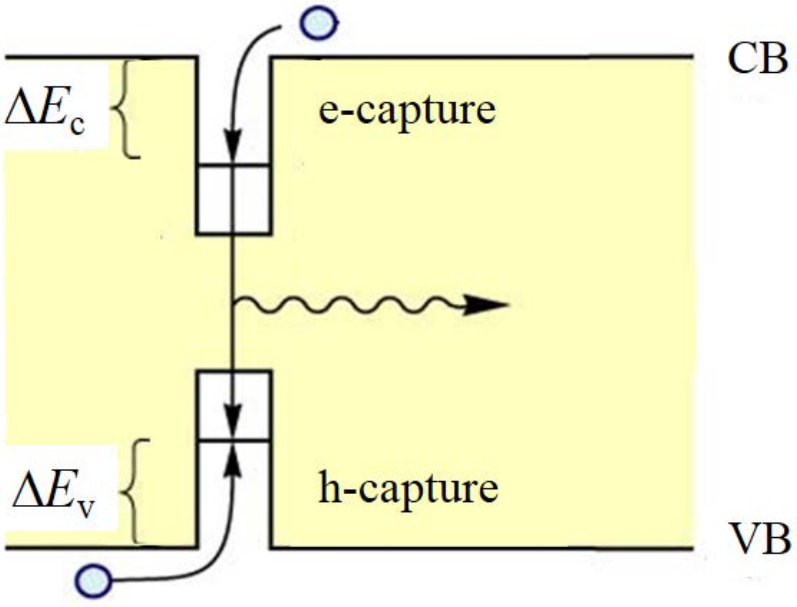

Figure 6.3. Carrier capture process in QWs: a pictorial view.

6.2 The quantum capture

Since in the barrier the electron envelope function is a 3D quantum eigenstate represented by a plane wave, in the effective mass approximation the particle must be thought of as fully delocalized with an energy $E = \hbar^2|\boldsymbol{k}|^2/2m$ above the energy of the barrier conduction band (CB), measured from an arbitrary reference.

The barrier state has a finite probability to decay to the localized, confined quasi-2D eigenstate belonging to the QW, with a potential energy lower than the barrier CB energy by an amount ΔE, often called *confinement* or *binding energy* (in addition, the electron in the QW can have a kinetic energy $\hbar^2(k_x^2 + k_y^2)/(2m)$ quadratic in the in-plane wavevector components): thus the QW state is energetically favored with respect to the barrier state. In order to decay into the QW, the barrier state must lose part of its energy, and the most important mechanisms are a) the emission of a LO-phonon and b) the carrier-carrier scattering [108–112], illustrated in Fig. 6.4.

The two mechanisms are very different: according to the first one, an electron in the barrier state with energy E and momentum $\boldsymbol{k}$ emits al LO-phonon of momentum $\boldsymbol{q}$ and energy $\hbar\omega_m$, decaying in a QW state of energy E' and momentum $\boldsymbol{k'}$.

Instead, by the electron-electron scattering mechanism, an electron in the barrier state with energy E_1 and momentum $\boldsymbol{k_1}$ interacts with another electron whose energy and momentum are E_2 and $\boldsymbol{k_2}$ by the exchange of a *virtual photon*, decaying in a QW state of

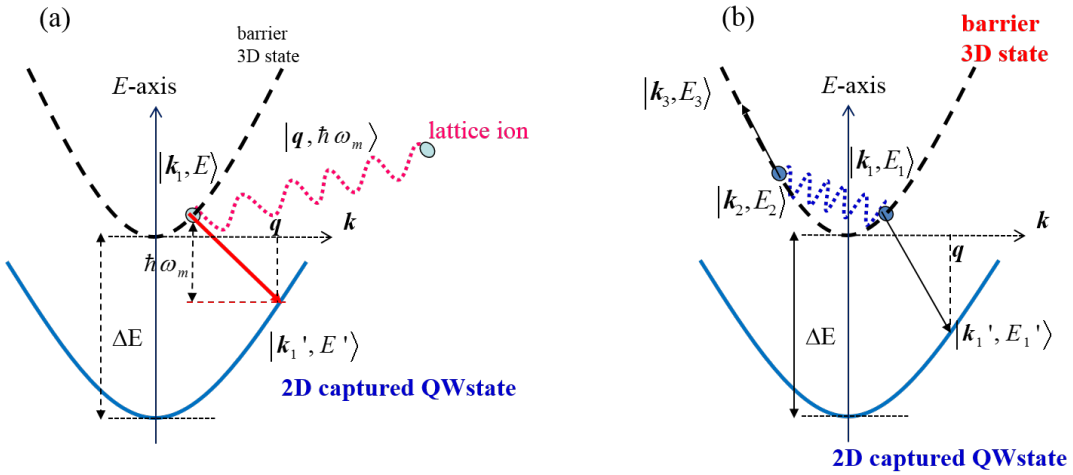

Figure 6.4. Carrier capture process in QWs. (a) An electron in the barrier state with energy E and momentum $\boldsymbol{k_1}$ emits al LO-phonon of momentum $\boldsymbol{q}$ and energy $\hbar\omega_m$, decaying in a QW state of energy E' and momentum $\boldsymbol{k_1'}$. (b) Electron-electron scattering: an electron in the barrier state with energy E_1 and momentum $\boldsymbol{k_1}$ interact with another electron whose energy and momentum are E_2 and $\boldsymbol{k_2}$ by the exchange of a *virtual photon*, decaying in a QW state of energy E_1' and momentum $\boldsymbol{k_1'}$.

energy E_1' and momentum $\boldsymbol{k_1'}$.

In the present work, only the capture by LO-phonon emission will be addressed. According to our preliminary calculations (a paper is now in preparation), the capture by electron-electron scattering is much less efficient, also for very high carrier density.

In order to describe the mechanism, it is important to consider the time-evolution of a quantum state in a medium. The condensed matter and many-body physics are the tools needed to investigate the underlying physics, but this is not the right place to discuss them in detail (see e.g textbooks as those by A. L. Fetter and J. D. Walecka [113] and G. D. Mahan [114]). Nevertheless, something of qualitative, but sufficient to situate the problem, can be said.

6.3 The concept of complex self-energy

Let us consider the temporal evolution of an electronic state represented by the wavefunction $\phi_{\boldsymbol{k}}(\boldsymbol{r}, t)$, eigenstate of the *free* Hamiltonian H (i.e. in absence of any interaction), whose eigenvalue is the energy $E_{\boldsymbol{k}}$ (eigenvalue of the stationary Schrödinger equation). Its general expression is

$$\phi_{\boldsymbol{k}}(\boldsymbol{r}, t) = \phi_0(\boldsymbol{r}) \exp\left(\frac{i}{\hbar} E_{\boldsymbol{k}} t\right). \tag{6.7}$$

It is well known that if there are no interaction at all, the modulus of the eigenstate amplitude $\phi_0(\boldsymbol{r})$ remains constant in time, and the eigenstate evolution is described simply by the unitary transformation given by the exponential factor. This applies also for two (and even more) *non interacting* electrons in a QW: the Hamiltonian ground eigenstate is the antisymmetric state $1/\sqrt{2}\,(\phi_1(z_1)\phi_2(z_2) - \phi_2(z_1)\phi_1(z_2))$ (they are undistinguishable fermions), but the ground eigenstate is stable: its amplitude probability never changes in time.

What if H includes an interaction term? Also in this case (for example, we consider the Coulomb interaction between the two electrons in the QW), it is possible to build exact eigenstates of the whole Hamiltonian, normally organized in multiplets, whose amplitude is stable in time.

Let us now consider a more general scenario, with an electron merged in a semiconductor plasma (electrons, holes, positive ions background), interacting each other by the Coulomb interaction, and for a moment let us forget the required antisymmetry of the total wavefunction of such huge N-particle system. Let us also suppose the local charge neutrality. The considered electron repels the other electrons (Coulomb interaction, represented by the potential $V_i(r)$), exposing the positive background due to crystal ions:

as a consequence, the electron is surrounded by a *positive cloud* and moves as a quasi-particle (the bare electron plus the positive cloud). All this means that another electron (surrounded itself by a similar positive cloud), when approaching it, experiences *weaker* Coulomb repusion with respect to the repulsion experienced by two bare electrons in vacuum. The reason is the presence of the positive cloud: this is the phenomenon of the *potential screening*. The potential energy $E'_{\boldsymbol{k}}$ gained by the second electron when approaching the first one from infinity – as a consequence – is a little less with respect to the energy $E_{\boldsymbol{k}}$ gained in the case of two electrons approaching each other in vacuum. The difference of energy is called **self-energy**. As a remark, a hole behaves as a quasi-particle as well: in addition it is attracted by (and attracts as well) electrons, contributing to their positive cloud. It is because of this fact that the *screening* of the Coulomb potential in an electron-hole (e-h) plasma in a semiconductor strongly depends on the carrier density.

If we followed just the simple argument indicated above, we obtained the so-called **Hartree** self-energy, that we indicate with Σ_H (and normally is negative), and it would be $E'_{\boldsymbol{k}} = E_{\boldsymbol{k}} + \Sigma_H$. This self-energy is just a number depending on carrier density, and is $\boldsymbol{k}$-independent [115].

At a microscopic level, this process can be viewed as an electron with momentum $\boldsymbol{k}$ that emits a virtual photon [113, Ch. 3], representing the Coulomb interaction: it produces an electron-hole (*e-h*) couple (that is, it knocks a Fermi sea's electron out of the Fermi surface), that immediately destroys (i.e. it immediately recombines with the hole left into the Fermi sea): the virtual photon is re-absorbed by the first electron, that continues propagating with the same momentum $\boldsymbol{k}$.

Anyway, there is also another possible source of self-energy [113, Ch. 3]: the first electron and the knocked-out one can exchange each-other, since they are undistinguishable fermions: the obtained self-energy term is the **Fock**'s term Σ_F, that turns out to be $\boldsymbol{k}$-dependent.

In condensed matter physics the self-energy is represented by means of Feynman's diagrams, as shown in Fig. 6.5 for the Hartree and Fock terms. In these diagrams, the arrowed straight lines represent particles, or particle *propagators* indicated by G_0, and the wavy line a particle-particle potential interaction, indicated by V_i. Using a small number of simple rules, each Feynman's diagram can be readily expressed in its corresponding algebraic form and employed to obtain the Hamiltonian eigenvalues, corrected by its contribution to the total self-energy. In order to be a little more precise, the one-particle propagator is the probability amplitude that a particle put in $\boldsymbol{r}_1$ at time t_1 is observed in $\boldsymbol{r}_2$ at time t_2. Expressing it in the conjugate space, it is the probability amplitude that a particle propagates with wavevector $\boldsymbol{k}$ and energy E. Its algebraic form will be presented in a moment; first, I prefer to remind two concepts more, because thay are important.

First, it is possible to demonstrate [113, 115] that for an electron gas in a metal (or semiconductor) *for which charge neutrality is fulfilled*, the Hartree term is identically cancelled out by a similar term due to the positive background of ions.

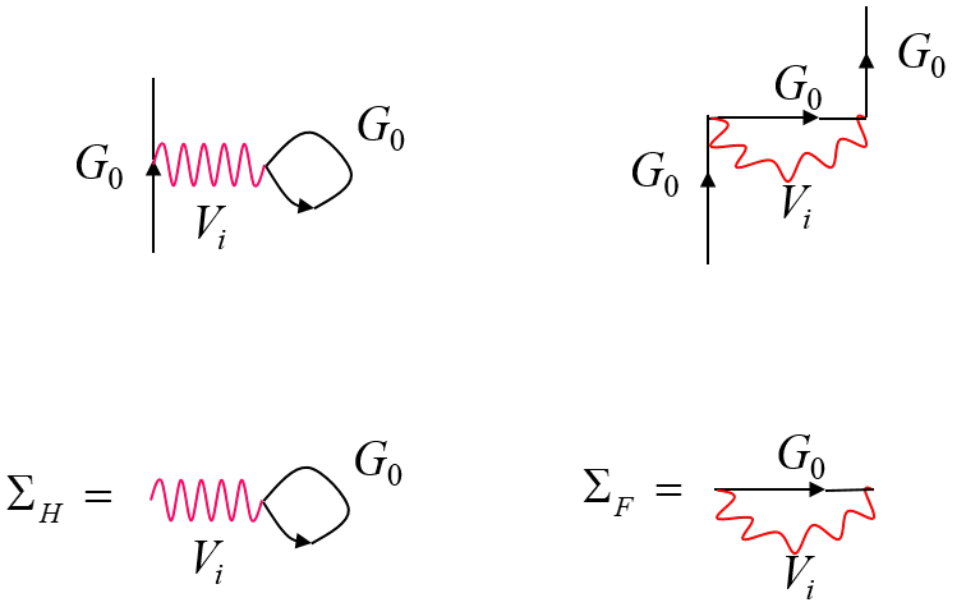

Figure 6.5. Hartree-Fock corrections to the free propagator: the sum of the two diagrams gives the correction to the free propagator G_0. The selfenergy contributions Σ_H and Σ_F are given by a free propagator times the considered interaction, integrating on all the possible frequency and momentum values of the exchanged virtual photon, imposing the energy-momentum conservation in the vertices of the interaction.

Second, things are more complicated: translating the Fock's diagram into its corresponding algebraic form, the result is infinite (it diverges) at the Fermi surface, giving the self-energy an infinite contribution. This problem can be solved considering the fact that in the real world the "bare" Coulomb potential never exists: injecting current in a semiconductor, the plasma density (electrons and holes) increases of orders of magnitude, contributing to the screening in a very effective way. Holes surround electrons and vice-versa, reducing the intensity of the effective Coulomb interaction. At an approximation acceptable studying the band gap renormalization (but *not* for the capture mechanisms we are addressing) the effect of the screening on the Coulomb interaction due to the plasma may be expressed by the Yukawa potential [113, 115]:

$$V_{i,s}(r) = V_i(r)\exp\left(-\kappa r\right) = \frac{q_e}{4\pi\epsilon_0}\frac{\exp\left(-\kappa r\right)}{r}. \tag{6.8}$$

Here ϵ_0 and q_e are the vacuum dielectric permittivity and the elementary charge. The factor $\exp\left(-\kappa r\right)$ multiplies the bare Coulomb potential: κ is known as *screening wavevector*, a characteristic length rendering *finite* the interaction range. At this point it is possible to obtain a more realistic estimate of the Hartree-Fock self-energy replacing V_i with the Yukawa potential $V_{i,s}$, obtaining the so-called *screened Hartree-Fock approximation* of

the self-energy, that remains *finite (and real)*.

In the most general case in which also Σ_H contributes, the eigenvalue is corrected by the presence of carrier plasma as $E'_{\boldsymbol{k}} = E_{\boldsymbol{k}} + \Sigma_H + \Sigma_F(\boldsymbol{k})$, and corrections like this hold both for electrons in the conduction band and for holes in the valence band.

In addition, another and generally more important term, known as **Coulomb-Hole** [116, Ch. 7,9], must be included in the total self-energy: the total energy of the system reduces because electrons avoid each other, due to their mutual Coulomb repulsion (correlation energy). Nevertheless *all these contributions to self-energy turn out to be real numbers*, and the importance of this remark in the study of the capture problem will be clear in a moment.

What does it happen if the total self-energy Σ includes – for some reason, at the end of its calculation – a small imaginary part $\Im\Sigma$, with $\Im\Sigma \ll E_k$? The Hartree-Fock self-energy is real, anyway let us consider this possibility. In this case the state evolves in time like an approximate eigenstate, with a small exponential damping [113, Ch. 3]:

$$\begin{aligned}\phi_k(\boldsymbol{r},t) &= \phi_0(\boldsymbol{r})\exp\left(\frac{i}{\hbar}E'_{\boldsymbol{k}}t\right)\\ &= \phi_0(\boldsymbol{r})\exp\left(\frac{i}{\hbar}\left(E_{\boldsymbol{k}}+\Re\Sigma+i\Im\Sigma\right)t\right)\\ &= \phi_0(\boldsymbol{r})\exp\left[\frac{i}{\hbar}(E_{\boldsymbol{k}}+\Re\Sigma)t\right]\exp\left(\frac{-\Im\Sigma}{\hbar}t\right).\end{aligned} \tag{6.9}$$

This means that the probability amplitude of the approximate eigenstate

$$|\phi_k(\boldsymbol{r},t)|^2 = |\phi_0(\boldsymbol{r})|^2\exp\left(\frac{-2\Im\Sigma}{\hbar}t\right) \tag{6.10}$$

will decay in a characteristic time τ given by

$$\tau = \left(\frac{2}{\hbar}\Im\Sigma\right)^{-1}. \tag{6.11}$$

We still have to understand how to calculate Σ, but the most interesting fact is that, if $\Im\Sigma \neq 0$, the electron will decay to some other state in a time τ. Since the decay is spontaneous, the energy of the final state must be lower than the initial state's one. *This is exactly the scenario of an electron in a barrier state that, after a characteristic time τ, spontaneously decays into a QW's state – it is captured –, because QW's eigenstates are energetically favored with respect to barrier's.*

To be more precise, let us create a barrier's state electron in $\boldsymbol{r'}$ with energy $E_{\boldsymbol{k}}$ and effective mass m. This scenario is well described by the stationary, non-homogeneous Schrödinger equation for a unknown function G_0, with a source term in $\boldsymbol{r'}$:

$$\left(-\frac{\hbar^2\nabla^2}{2m} - E_{\boldsymbol{k}}\right)G_0(\boldsymbol{r},\boldsymbol{r'};E_{\boldsymbol{k}}) = \delta(\boldsymbol{r}-\boldsymbol{r'}). \tag{6.12}$$

The solution G_0, called the **free propagator** (as anticipated) or **Green's function**, can be easily found taking the Fourier transform in the conjugate variable $\boldsymbol{k}$, obtaining

$$\left(\frac{\hbar^2|\boldsymbol{k}|^2}{2m} - E_{\boldsymbol{k}}\right) G_0(\boldsymbol{k}; E_{\boldsymbol{k}}) = 1. \tag{6.13}$$

from which we obtain

$$G_0(\boldsymbol{k}; E_{\boldsymbol{k}}) = \frac{1}{\frac{\hbar^2|\boldsymbol{k}|^2}{2m} - E_{\boldsymbol{k}}}. \tag{6.14}$$

In the momentum $\boldsymbol{k}$ space, the two momentum eigenvalues $\boldsymbol{k} = \pm\hat{k}\sqrt{2mE_{\boldsymbol{k}}}/\hbar$ coincide with the pole of the propagator ($\hat{k}$ is the versor of $\boldsymbol{k}$). This is true also if the motion takes place in a medium: the energy $E'_{\boldsymbol{k}}$ is given by the energy $E_{\boldsymbol{k}}$ corresponding to a free electron in vacuum, plus the self-energy correction Σ, that $E'_{\boldsymbol{k}} = E_{\boldsymbol{k}} + \Sigma$, and the momentum eigenvalues $\boldsymbol{k} = \pm\hat{k}\sqrt{2m(E_{\boldsymbol{k}} + \Sigma)}/\hbar$ still coincide with the poles propagator. In the end, the "perturbed" propagator is:

$$G(\boldsymbol{k}; E_{\boldsymbol{k}}) = \frac{1}{\frac{\hbar^2|\boldsymbol{k}|^2}{2m} - E_{\boldsymbol{k}} - \Sigma} \tag{6.15}$$

and describes the motion of a quasi-particle in a medium. Eventually Σ may be complex and depending on $\boldsymbol{k}$.

In the framework of the perturbation theory, the motion of a particle may be much more complicated with respect to what considered so far. In short, we can say:

at the 0^{th} perturbative order the electron proceeds freely, and the probability amplitude is represented simply by a free propagator (Fig. 6.6, 1^{st} line).

At the 1^{st} perturbative order, we find the already examined Hartree and Fock diagrams: an electron can interact by means of a virtual photon of momentum $\boldsymbol{q}$ and frequency ω_m (or also a phonon: in this case the interaction may be the Frölich one) with an electron emerging from the Fermi sea and immediately annihilating (first diagram in the 2^{nd} line of Fig. 6.6, the Hartree term). As anticipated, when charge neutrality is fulfilled this diagram can be discarded, since its contribution can be demonstrated to be exactly cancelled out [113]. In the other cases the result is a real number. Another possibility is the emission and re-absorption of a photon or phonon (second diagram in the 2^{nd} line of Fig. 6.6). This is the already treated Fock's diagram, yielding an infinite.

At the 2^{nd} perturbative order (two vertices of interactions, 3^{rd} line in Fig. 6.6) there are 10 possible diagrams, many of them, if translated in algebraic form and integrated over exchanged momentum, yield an infinite result. Classifying the diagrams with respect to the order of divergence, it is expected that the greatest contribution to the amplitude of propagation probability is brought by *the most diverging one*: we keep that diagram only, despite the result is infinite.

Following this rule, we keep the most diverging diagram at each perturbative order (4^{th} line in Fig. 6.6), discarding the others: this is the Random Phase Approximation (RPA) [113]. Removing (or amputating) the lower-most and the upper-most straight lines (external propagators) in the diagrams shown in Fig. 6.6 (these so-called external lines correspond to prescribed values for momentum and energy), one retains a contribution to the self-energy, represented at each perturbative order by:

- at the 1^{st} order, a propagator times an interaction (a wavy line), second diagram in the second line of Fig. 6.6;

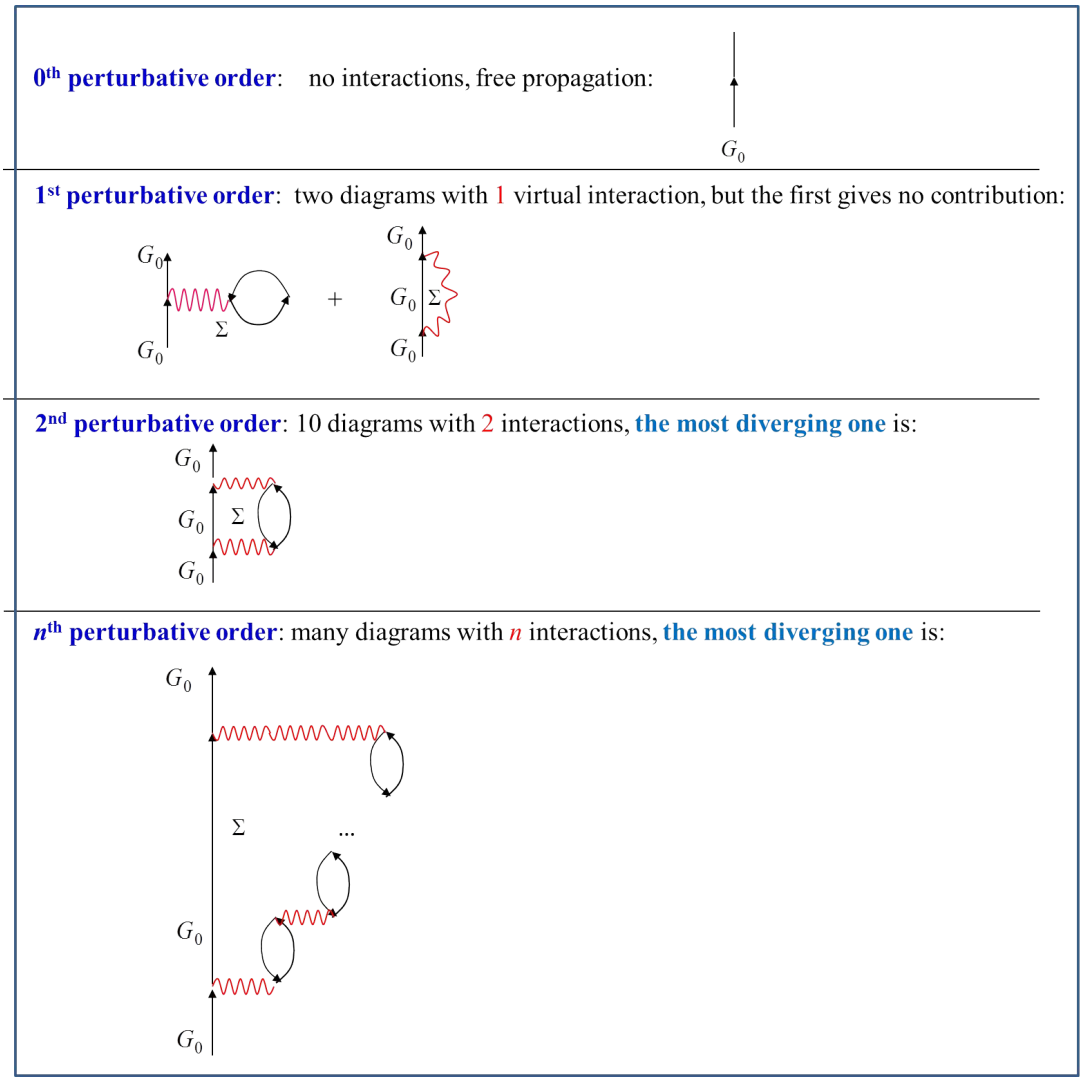

Figure 6.6. Dressed propagator G, as sum of self-energy contributions.

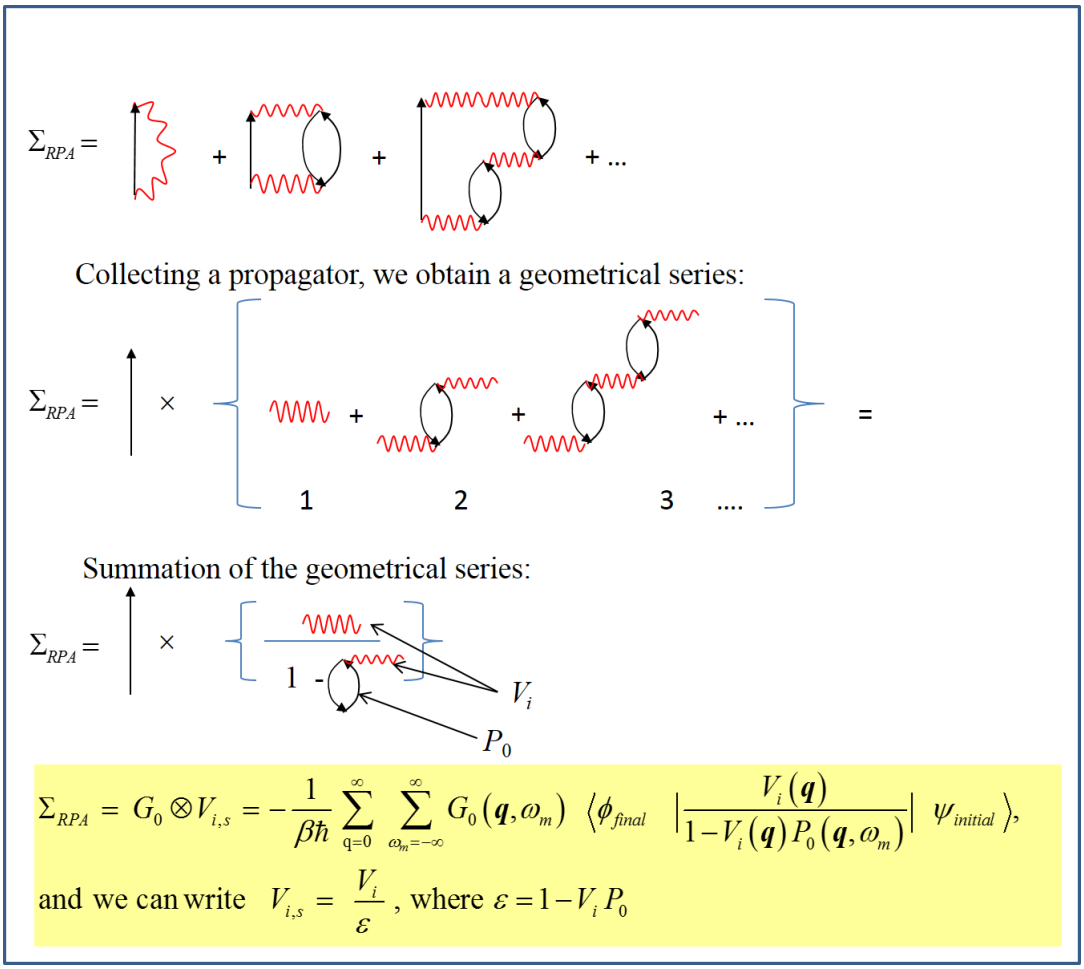

Figure 6.7. Construction of the RPA self-energy: in the final formula, $\boldsymbol{q}$ and ω_m are the momentum and frequency of the exchanged virtual photon or phonon, ψ and ϕ are the initial and final interacting states, P is the polarization and ϵ is the RPA dielectric function.

- at the 2nd order, a propagator times a sort of more complicated diagram, third line of Fig. 6.6;
- at the n^{th} order, a propagator times a still more complicated diagram, as represented in the fourth line of Fig. 6.6.

Referring now to Fig. 6.7, we "collect" a propagator, that multiplies a series, that is a geometrical series. Although each term diverges, *the result is finite*: it is the RPA self-energy, given by a free propagator G_0, times a "dressed" or "screened" interaction $V_{i,s}$, in which contributions up to infinite order are present: this is a great difference, with respect to stop the calculation at, say, the second perturbative order or so. As indicated in Fig. 6.7, the calculation includes a summation over the frequencies ω_m and momentum $\boldsymbol{q}$

of the exchanged quantum of energy related to the considered interaction (a photon, if we are dealing with Coulomb electron-electron interaction, or a phonon if we are considering the electron-lattice interaction via the Frölich interaction). A remark: the frequency ω_m has a subscript m because it is a bosonic frequency, discrete and multiple of an elementary frequency, as it will be clear in the next Section.

Still in Fig. 6.7 there are the definitions of the polarization bubble P and the dynamical RPA dielectric function ϵ. The challenge of the calculation we present in the next Section is to avoid any $\boldsymbol{q}$-truncation during the $\boldsymbol{q}$-summation, from zero to infinite, unavoidable when numerical integration is considered: because of this fact, the approach to this point will be analytic. In addition, we will not perform neither the *static* approximation (setting $\omega_m = 0$ in the polarization P), nor the *long-wavelength* approximation (setting $\boldsymbol{q} = 0$ in the dielectric function ϵ): all this would lead to incorrect results.

In summary, we will describe the capture time of an electron from a barrier to a QW state as

$$\tau = \left(\frac{2}{\hbar} \Im \Sigma_{RPA} \right)^{-1}, \tag{6.16}$$

where Σ_{RPA} is evaluated as in Fig. 6.7, through a summation from zero to infinity over both ω_m and $\boldsymbol{q}$. Referring to this figure, the central problem is to write the interaction potential V_i corresponding to the considered capture mechanism, together with the RPA dielectric function ϵ expressing the screening of the potential, for which we will choose the Single Plasmon Pole form [114].

6.4 Quantum well electron capture time: An analytic approach

The Section that follows is based on a paper we published in the J. Appl. Phys., vol. 114, p. 053704, Aug. 2013 [5], and constitutes an extended version of it.

Coming back to the capture process into a QW, we can recall the Fig. 6.4: two electrons (for holes the question is similar) can interact exchanging a *virtual* photon or LO-phonon, as shown by the two Feynman's diagram in Fig. 6.8(a) (i.e the first line of the figure). The simplest approach to evaluate the capture time by LO-phonon emission following the path outlined in the previous Section would apparently consist in writing the Σ_{RPA} considering just the Frölich interaction. In the perturbative approach, the path would be to consider Feynman's diagrams containing only dashed lines, introducing in some way the effect of the plasma screening.

This cannot be done easily, because beyond the first perturbative order the two scattering processes are deeply connected and both interactions (the wavy and the dashed line)

are present in the same diagram: at the second perturbative order there are four possible "most diverging" diagrams (Fig. 6.8(b)) and at the n-th order they are 2^n. The *bare* (unscreened) electron-electron (*ee*) Coulomb and the LO-phonon (*ph*) interaction energies V_{ee}^{∞} and V_{ph} (the wavy and dashed line in the diagrams, respectively) are given in the $\boldsymbol{q}$-space by [114]

$$\begin{aligned} V_{ee}^{\infty} &= \frac{4\pi q_e^2}{\epsilon_0 \epsilon_\infty |\boldsymbol{q}|^2} \\ V_{ph} &= -M^2(|\boldsymbol{q}|) D_0(\omega_m) \end{aligned} \tag{6.17}$$

where

$$M^2(|\boldsymbol{q}|) = (1/2) K_\epsilon \omega_{LO} \frac{4\pi q_e^2}{\epsilon_0 |\boldsymbol{q}|^2} \tag{6.18}$$

is the square of the unscreened electron-phonon matrix element, $K_\epsilon = \epsilon_\infty^{-1} - \epsilon_s^{-1}$, whereas ω_{LO} is the longitudinal optical phonon frequency, and

$$D_0(\omega_m) = \frac{2\omega_{LO}}{\omega_m^2 - \omega_{\mathrm{LO}}^2} \tag{6.19}$$

is the unscreened phonon propagator.

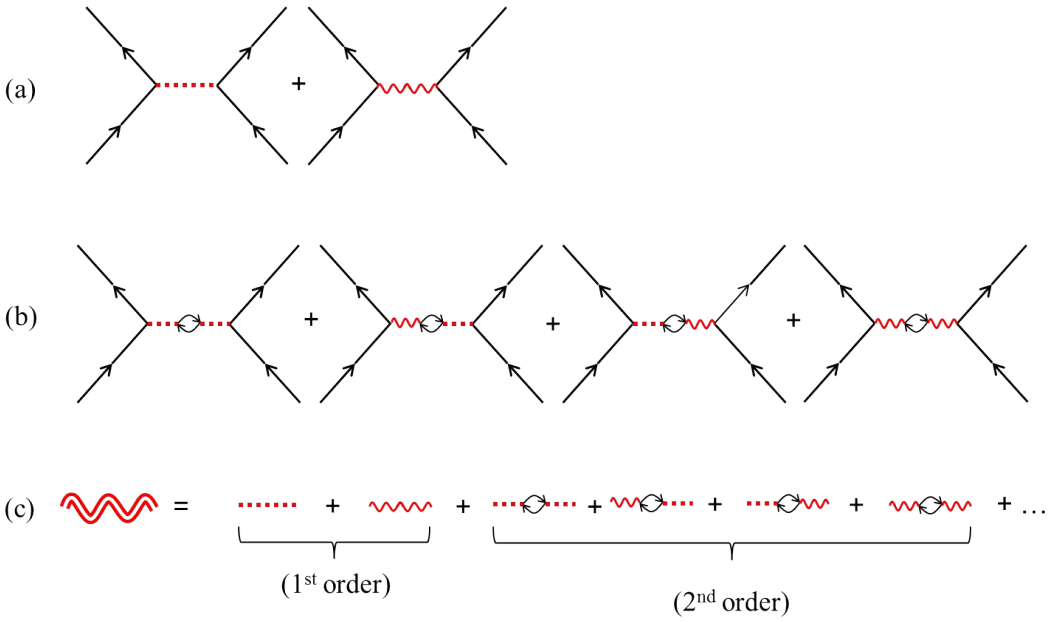

Figure 6.8. (a) First order Feynman's diagrams for electron scattering, exchanging a phonon (dashed line) or a photon (waved line). (b) The second order Feynman's diagram with one polarization bubble. RPA consists in summing such n-bubbles contributions to infinite-order, obtaining (c) the effective *dressed* RPA interaction (double waved line).

The effective *ee* plus *ph* interaction V_{eff} in the RPA [114, 117] can be found summing up to infinite perturbative order all possible n-bubble diagrams (Fig. 6.8(c)) according to the Dyson's equation, obtaining

$$V_{\text{eff}} = \frac{V_{ee}^{\infty} + V_{ph}}{1 - (V_{ee}^{\infty} + V_{ph})\, P}\,, \tag{6.20}$$

where $P = (1-\epsilon)/V_{ee}^{\infty}$ is the polarization of the electron gas and ϵ is the electron-electron part of the dielectric function. An important and non-trivial step is the separation of the effective *dressed* (dynamically screened) RPA interaction into the *ee* and *e-ph* contributions $V_{\text{ee,s}}$ and $V_{\text{ph,s}}$, obtained in the end, following Ref. [114],

$$V_{ee,s} = \frac{V_{ee}^{\infty}}{\epsilon}, \qquad V_{ph,s} = \frac{V_{ph}}{\epsilon^2\ (1 - M^2(|\boldsymbol{q}|) D_0 P/\epsilon)}\,, \tag{6.21}$$

where $\epsilon = \epsilon\,(|\boldsymbol{q}|, \omega)$ is the RPA dynamic dielectric function. The capture time by LO-phonon emission *in presence of carrier population* is proportional to $\Im\Sigma_{RPA,ph}$, where $\Sigma_{RPA,ph}$ is the RPA-phonon self-energy given by a free-propagator times $V_{ph,s}$, summed over phonon frequencies ω_m and phonon momentum $\boldsymbol{q}$, but before going into details of the developed model, let us examine what has been done in literature.

Several authors have investigated capture process via electron-electron scattering and via emission of LO-phonon, especially for III-V heterostructures [108–112], calculating the corrensponding scattering rates at different levels of approximation. A first group of works calculate the capture time from a bulk–state into a QW, describing its oscillating behavior with the QW width [118–121]. These works, despite their unquestionable value, do not deal with the dependence of capture time on the carrier density, having considered the Frölich LO–phonon hamiltonian with the unscreened Coulomb potential.
Other works [110, 122, 123] offer numeric and very complete investigations about these two population relaxation mechanisms in QWs and also in quantum–dots [124], making use of the full frequency and wave vector dependent dielectric function and the renormalized phonon frequencies due to phonon–plasmon coupling. However, the full understanding of the physical mechanisms at play may be in this case quite difficult and a brief comparison between their and our approaches are given is Sec. 6.4.5. Nevertheless its comprehension is crucial not only in optoelectronics, but also for the simulation of hot carrier relaxation in solar cells [125] and LEDs, for which in several cases efficiency droop may be observed [107].
The intent of this work is to make use of a known formalism to develop an original method for *a*) calculating the hot electrons relaxation time in QW confined states and *b*) for the evaluation of electrons capture/escape times into and from QW, via emission/absorption of coupled LO–phonon–plasmon [126, 127] for arbitrary value of carrier density.
The electron–LO-phonon interaction is addressed here only just as an example, but with

the appropriate changes all other types of phonon interactions can be considered as well. The carrier–carrier scattering and its interplay with phonon scattering will be presented in a future separate contribution. In fact, the same formalism may be employed to describe many other relaxation processes and the scope of this work is not to include them all, but to take this process as example of application of a powerful method of investigation, offering a new formulation and allowing to shed some light into certain peculiarities of the underlying physics that, in the author's opinion, have not been highlighted yet in a clear way.
It is convenient to emphasize that the development of simple model (ready to be implemented in device modeling tools) that calculates scattering time as a function of carrier density can have many applications, among which its inclusion in rate equation systems describing laser and semiconductor optical amplifier dynamics, but also the evaluation of phonon–assisted Auger recombination coefficient [128, 129]; furthermore, its inclusion in light–emitting diodes modeling tools may be helpful in understanding their efficiency droop mechanisms [107, 130]. In Sec. 6.4.1 the general formalism is developed, whereas in Sec. 6.4.3 and Sec. 6.4.4 the intraband relaxation and capture time in QWs are respectively evaluated with the application of the developed concepts, then in Sec. 6.4.5 some conclusions and final remarks will be given.

6.4.1 General formalism

The formalism that is developing in this Section offers a general formulation of the scattering processes in quantum wells, mediated by the emission or absorption of bulk LO-phonon in arbitrarily dense carrier plasma. The LO-phonon screened interaction is described by the Frölich hamiltonian, in which the dynamically screened electrostatic potential is considered, according to the Single Plasmon Pole (SPP) description of the RPA formalism [116]. The usual long–wavelength and static limits of SPP (the Thomas–Fermi limit) are not made, retaining instead the full dynamic expression of the dielectric function. Although the dynamic SPP is a well known formalism, it is not usually employed in this context because of its complexity. Nevertheless its advantage is consistent and leads in a natural way to the coupling between plasmon and LO–phonon, describing screening and Fermi band–filling effects in a self–consistent fashion.

All interfaces in heterostructure are assumed to be planar, so that momenta and wavefunctions can be decomposed in *in–plane* and *orthogonal* components. Therefore, if $r = r(x, y)$ is the position vector in the QW plane and z is the growth direction, the *in–plane* motion of electrons with *in–plane* wavevector k is described by a plane wave $\exp(ik \cdot r)$, whereas in the z direction the motion is described by a Bloch state, whose envelope wavefunction is $\psi(z)$. In addition and for further simplification, the effective mass approximation and parabolic bands have been assumed valid, given that the focus of this work is on the role of the dynamical screening of the interaction, more than on the band–structure.

It is considered a carrier in an initial state ψ_1 of energy E and momentum $\boldsymbol{k}$ that emits a phonon of momentum $\boldsymbol{q}$ and frequency ω_m, making a transition to the state ψ_2 of energy E' and momentum $\boldsymbol{k'} = \boldsymbol{k} - \boldsymbol{q}$. Decomposing momenta in the in–plane components k, k', $q_\parallel$ and orthogonal components k_z, k'_z, q_z, it is possible to write the unscreened in–plane interaction Coulomb potential V_0 as [116]:

$$V_0(q_\parallel) = \frac{4\pi q_e^2}{\epsilon_0} \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} d\zeta \, \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_z \times \psi_1^*(z)\psi_2^*(\zeta) \frac{\exp\left(iq_z|z-\zeta|\right)}{q_\parallel^2 + q_z^2} \psi_2(\zeta)\psi_1(z)\,, \tag{6.22}$$

Extending the integration to the complex plane and adding to the path a semicircle at infinity in the positive half–plane giving no contribution, the integration in q_z is straighforward, exploiting the contour integration around the integrand pole in $q_z = iq_\parallel$, yielding

$$V_0(q_\parallel) = \frac{2\pi q_e^2}{\epsilon_0} \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} d\zeta \; \psi_1^*(z)\psi_2^*(\zeta) \frac{\exp\left(-q_\parallel|z-\zeta|\right)}{q_\parallel} \psi_2(\zeta)\psi_1(z)\,. \tag{6.23}$$

It is worth of note the fact that Eq. (6.22) always contains the 3D Coulomb potential form

$$V_{3D}(q_\parallel, q_z) = \frac{4\pi q_e^2}{\epsilon_0 \left(q_\parallel^2 + q_z^2\right)} \tag{6.24}$$

regardless of the interacting involved states. For realistic QW of width L_z it is obtained obtain

$$V_0(q_\parallel) = \frac{2\pi q_e^2}{\epsilon_0} \frac{I(q_\parallel)}{q_\parallel}\,, \tag{6.25}$$

in which the form factor $I(q_\parallel)$ only depends on the details of envelope eigenfunctions [131]. Instead, if we were interested in obtaining the interaction potential between two perfectly 2D quantum well states (that is, considering $\psi_1(z)$ and $\psi_2(z)$ as δ–Dirac functions), we obtained the well known 2D Coulomb potential $V_0(q_\parallel) = 2\pi q_e^2 / \left(\epsilon_0 q_\parallel\right)$, that is $I(q_\parallel) = 1$.

The present discussion is not limited to the simple 2D case, but it is addressed the general situation of a finite–width QW, whose orthogonal envelope wavefunction $\psi(z)$ can be evaluated at the desired accuracy, keeping into account strain with the most appropriate treatment. Following the path here outlined, the carrier–phonon interaction matrix

element $V_{\text{eff}}(q_{\|};\omega_m)$ can be written as [132]

$$V_{\text{eff}}(q_{\|};\omega_m) = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} d\zeta \, \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_z \times \psi_1^*(z)\psi_2^*(\zeta) \exp\left(iq_z|z-\zeta|\right) \times \frac{-M^2(\boldsymbol{q})\, D(\boldsymbol{q},\omega_m)}{\epsilon^2(\boldsymbol{q},\omega_m)} \psi_2(\zeta)\psi_1(z)\,, \tag{6.26}$$

in which the renormalized propagator for a phonon of frequency ω_m is [132]

$$D(\boldsymbol{q},\omega_m) = \frac{2\,\omega_{LO}}{\omega_m^2 - \omega_{LO}^2 - \omega_{LO}\frac{M^2(\boldsymbol{q})P(\boldsymbol{q},\omega_m)}{\epsilon(\boldsymbol{q},\omega_m)}}. \tag{6.27}$$

Here $\epsilon(\boldsymbol{q},\omega_m)$ is the screened dielectric function, given by [132]

$$\frac{1}{\epsilon(\boldsymbol{q},\omega_m)} = 1 + \frac{\Omega_{pl}^2}{\omega_m^2 - \omega_q^2}. \tag{6.28}$$

where Ω_{pl} and ω_q are respectively the plasma and the effective plasmon frequencies [116] function of carrier density N and wavevector $\boldsymbol{q}$, and $P(\boldsymbol{q},\omega_m)$ is the polarization, related to the dielectric function through $\epsilon(\boldsymbol{q},\omega_m) = 1 - V_0(q)P(\boldsymbol{q},\omega_m)$. The phonon propagator can be re–written, making use of Eqs. (6.22,6.18,6.28) as

$$D(\boldsymbol{q},\omega_m) = \frac{2\,\omega_{LO}\left(\omega_m^2 - \omega_q^2\right)}{\left(\omega_m^2 - \omega_+^2\right)\left(\omega_m^2 - \omega_-^2\right)} \tag{6.29}$$

where

$$\omega_\pm^2 = \frac{\omega_q^2 + \omega_{LO}^2}{2} \pm \frac{1}{2}\sqrt{\left(\omega_q^2 - \omega_{LO}^2\right)^2 + 4K_\epsilon \omega_{LO}^2 \Omega_{pl}^2} \tag{6.30}$$

in which the N– and $\boldsymbol{q}$– dependencies are implicitly understood in the definition of ω_q and Ω_{pl}.

With all these concepts in mind, it is considered the free, not renormalized propagator of a particle of effective mass m, energy $E = \hbar\omega$ and momentum $\boldsymbol{k}$ that just emitted a phonon of frequency ω_m and momentum $\boldsymbol{q}$, decaying in a state of momentum $\boldsymbol{k} - \boldsymbol{q}$. Following the Matsubara's formalism [132,133], the free propagator reads (μ is the Fermi energy, calculated considering a two–dimensional QW with a single conduction band, parabolic dispersion energy $E(k)$ and effective mass approximation [116]):

$$G_0(k, q_{\|}; i\omega, i\omega_m) = \frac{1}{i\omega_m + i\omega - \omega_{kq} - \frac{\mu}{\hbar}} \tag{6.31}$$

where $\hbar\omega_{kq} = \hbar^2|\boldsymbol{k} - \boldsymbol{q}|^2/(2m)$. The propagator allows to calculate the retarded self–energy Σ_r gained by the particle at various levels of approximation. At the lower order in

electron–phonon coupling, it can be visualized as a Feynman's diagram made of a particle line (the propagator G_0) that emits and reabsorbs a screened phonon (represented by the interaction V_{eff}). Its real part $\Re\Sigma_r$ gives rise to a shift in the position of the electron energy levels and appears as an extra term in the denominator of G_0. However it is customary in modeling devoted applications to set the real part of the self–energy $\Re\Sigma_r$ to zero, as the phonon contribution to the electronic energy can be thought of already included in the band structure.

The imaginary part of self–energy $\Im\Sigma_r$, instead, gives rise to a level broadening and yields the scattering rate, from which the relaxation time follows: this constitutes the object of the present study. Although it is customary in literature to make use of the static limit of RPA-SPP dielectric function (the Thomas–Fermi approximation [116], much simpler to manage), this path will not be followed, because of the remarkable amount of physics that would be lost in that way. Instead $\Im\Sigma_r$ can be evaluated to the lowest order in the electron phonon coupling [132], but making use of the *dynamic* RPA-SPP dielectric function, as:

$$\begin{aligned}\Sigma_r(k,\omega) = \quad & -\frac{1}{\beta\hbar}\sum_{q_\parallel}\sum_{\omega_m} V_{\text{eff}}(q_\parallel;\omega_m) \\ & \times G_0\left(k, q_\parallel; i\omega, i\omega_m\right). \end{aligned} \tag{6.32}$$

The motivation for this choice consists in the fact that phonon–plasmon coupling becomes increasingly important as the carrier density increases and its effects cannot be discarded. Now, firstly the integration in q_z can be made, following the same contour technique employed in obtaining Eq. (6.23). Then, after the integration in z and ζ, the phonon–mediated interaction matrix element $V_{\text{eff}}(q_\parallel;\omega_n)$ to be employed in Eq. (6.32) reads:

$$\begin{aligned} V_{\text{eff}}(q_\parallel;\omega_m) \quad & = -\frac{2\pi q_e^2 K_\epsilon \omega_{LO}^2 I\left(q_\parallel\right)}{\epsilon_0\epsilon_s\, q_\parallel} \\ & \times \frac{(\omega_m^2 - \omega_q^2 + \Omega_{pl}^2)^2}{(\omega_m^2-\omega_q^2)\,(\omega_m^2-\omega_+^2)\,(\omega_m^2-\omega_-^2)} \end{aligned} \tag{6.33}$$

Now, the frequency ω_m summation in Eq. (6.32) can be made following the Matsubara's method [132, 133], converting it to an integral in ω_m and extending it to the complex plane.

6.4.2 The frequency and momentum summations: detailed calculation

In the published paper the content of this Section was only summarized: here all the calculations have been reported explicitly.

Inserting Eq. (6.33) and Eq. (6.31) in Eq. (6.32) we obtain

$$\Sigma_r(k,\omega) = \frac{2\pi q_e^2 K_\epsilon \omega_{LO}^2}{\beta\hbar\epsilon_0\epsilon_s} \sum_{q_\parallel} \frac{I\left(q_\parallel\right)}{q_\parallel} \times \sum_{\omega_m} \frac{(\omega_m^2 - \omega_q^2 + \Omega_{pl}^2)^2}{\left(\omega_m^2 - \omega_q^2\right)\left(\omega_m^2 - \omega_+^2\right)\left(\omega_m^2 - \omega_-^2\right)\left(i\omega_m + i\omega - \omega_{kq} - \frac{\mu}{\hbar}\right)} . \tag{6.34}$$

Bosonic frequencies ω_m are discrete. Any function, analytic except for simple poles, can be expressed as a sum over poles and residues. In particular, extending the the Bose-Einstein distribution to the complex plane by an analytic continuation, the distribution can be expressed as a series of infinite poles, even multiples of $i\pi/(\beta\hbar)$ [114, Chap. 3] as

$$n_B(z) = \frac{1}{e^{\beta\hbar z} - 1} = -\frac{1}{2} + \frac{1}{\beta\hbar} \sum_{m=-\infty}^{\infty} \frac{1}{z - 2m\frac{i\pi}{\beta\hbar}} \tag{6.35}$$

where the poles $z_m = 2mi\pi/(\beta\hbar)$ are on the imaginary axis in the complex frequency z-plane. It is easy to check in Matlab that for real values of z, the value of $n_B(z)$ obtained making use of the summation in Eq. (6.35) converges rapidly to the (real) value obtained making use of the Bose-Einstein standard expression.

Let us define an integral I_B as

$$I_B = \oint_C f(z)\, n_B(z)\, dz \tag{6.36}$$

where the integration path is a circle at infinity in the complex plane. The function $f(z)$ is supposed analytic except for a number of its poles z_n. The function n_B, expressed as the infinite sum in Eq. (6.35) is analytic as well, except for its infinite poles. For the Jordan theorem, the result of the integral must be zero, hence the sum of the residues in the poles of $f(z)$ (if any) must be equal and opposite to the sum of residues in the poles of n_B in z_m:

$$\sum_{m=-\infty}^{\infty} \text{Res}\left[f(z)n_B(z)\right]_{z=z_m} = -\sum_n \text{Res}\left[f(z)n_B(z)\right]_{z=z_n} \tag{6.37}$$

The residues of n_B are equal to $1/(\beta\hbar)$, hence

$$\sum_{m=-\infty}^{\infty} \text{Res}\left[f(z)n_B(z)\right]_{z=z_m} = \frac{1}{\beta\hbar} \sum_{m=-\infty}^{\infty} f\left(2m\frac{i\pi}{\beta\hbar}\right) . \tag{6.38}$$

As a consequence:

$$\sum_{m=-\infty}^{\infty} f\left(2m\frac{i\pi}{\beta\hbar}\right) = -\beta\hbar \sum_n \text{Res}\left[f(z_n)n_B(z_n)\right] \tag{6.39}$$

that can also be written, remembering our notations:

$$\sum_{m=-\infty}^{\infty} f\left(\omega_m\right) = -\beta\hbar \sum_n \text{Res}\left[f(z_n) n_B(z_n)\right] \tag{6.40}$$

valid for all functions $f(\omega_m)$ with only poles as non-analyticity points, provided ω_m is a bosonic frequency (as in the case of phonons). In our discussion the function $f(z)$ is

$$f(z) = \frac{(z^2 - \omega_q^2 + \Omega_{pl}^2)^2}{\left(z^2 - \omega_q^2\right)\left(z^2 - \omega_+^2\right)\left(z^2 - \omega_-^2\right)\left(z - \omega_G\right)}, \tag{6.41}$$

having definined

$$\omega_G = -i\omega + \omega_{kq} + \frac{\mu}{\hbar}. \tag{6.42}$$

The function $f(z)n_B(z)$ has poles in

$$z_n = \left(\pm\omega_+,\, \pm\omega_-,\, \pm\omega_q,\, \omega_G\right). \tag{6.43}$$

Looking at Eq. (6.41), it is easy to verify that

$$\frac{1}{z^2 - \omega_+^2} = \frac{1}{2\omega_+}\left(\frac{1}{z - \omega_+} - \frac{1}{z + \omega_+}\right) \tag{6.44}$$

and similarly for the other two factors in ω_- and ω_q in Eq. (6.41). Adding an infinitesimal imaginary frequency $i\eta$ to each frequency, we have three poles in the upper half plane (positive frequencies) and 3 in the lower one (negative frequencies). Hence, in right-hand-ride of Eq. 6.40 we have 6 residues from such poles and 1 residue from the pole in ω_G.

It is easy to show that $n_B(-z) = -\left(1 + n_b(z)\right)$, hence we are able to calculate the

residues in the poles:

$$\begin{aligned}
\text{Res}(f\,n_B)_{\omega_+} =& \frac{(\omega_+^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_+\left(\omega_+^2-\omega_q^2\right)(\omega_+^2-\omega_-^2)}\frac{n_B(\omega_+)}{(\omega_+-\omega_G)}\\
\text{Res}(f\,n_B)_{-\omega_+} =& -\frac{(\omega_+^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_+\left(\omega_+^2-\omega_q^2\right)(\omega_+^2-\omega_-^2)}\frac{-(1+n_b(\omega_+))}{(-\omega_+-\omega_G)}\\
\text{Res}(f\,n_B)_{\omega_-} =& \frac{(\omega_-^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_-\left(\omega_-^2-\omega_q^2\right)(\omega_-^2-\omega_+^2)}\frac{n_B(\omega_-)}{(\omega_--\omega_G)}\\
\text{Res}(f\,n_B)_{-\omega_-} =& -\frac{(\omega_-^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_-\left(\omega_-^2-\omega_q^2\right)(\omega_-^2-\omega_+^2)}\frac{-(1+n_b(\omega_-))}{(-\omega_--\omega_G)}\\
\text{Res}(f\,n_B)_{\omega_q} =& \frac{\Omega_{pl}^4}{2\omega_q\left(\omega_q^2-\omega_+^2\right)\left(\omega_q^2-\omega_-^2\right)}\frac{n_B(\omega_q)}{(\omega_q-\omega_G)}\\
\text{Res}(f\,n_B)_{-\omega_q} =& -\frac{\Omega_{pl}^4}{2\omega_q\left(\omega_q^2-\omega_+^2\right)\left(\omega_q^2-\omega_-^2\right)}\frac{-(1+n_b(\omega_q))}{(-\omega_q-\omega_G)}
\end{aligned}$$

Now we must sum the residues (but we must not forget the pole in ω_G, that we will treat in a moment):

$$\begin{aligned}
\text{Res}(f\,n_B)_{\omega_+}+\text{Res}(f\,n_B)_{-\omega_+} =& -\frac{(\omega_+^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_+\left(\omega_+^2-\omega_q^2\right)(\omega_+^2-\omega_-^2)}\left(\frac{n_B(\omega_+)}{(\omega_G-\omega_+)}+\frac{1+n_B(\omega_+)}{\omega_G+\omega_+}\right)\\
=& -B_+\left(\frac{n_B(\omega_+)}{(\omega_G-\omega_+)}+\frac{1+n_B(\omega_+)}{\omega_G+\omega_+}\right)\\
\text{Res}(f\,n_B)_{\omega_-}+\text{Res}(f\,n_B)_{-\omega_-} =& -\frac{(\omega_-^2-\omega_q^2+\Omega_{pl}^2)^2}{2\omega_-\left(\omega_-^2-\omega_q^2\right)(\omega_-^2-\omega_+^2)}\left(\frac{n_B(\omega_-)}{(\omega_G-\omega_-)}+\frac{1+n_B(\omega_-)}{\omega_G+\omega_-}\right)\\
=& -B_-\left(\frac{n_B(\omega_-)}{(\omega_G-\omega_-)}+\frac{1+n_B(\omega_-)}{\omega_G+\omega_-}\right)\\
\text{Res}(f\,n_B)_{\omega_q}+\text{Res}(f\,n_B)_{-\omega_q} =& -\frac{\Omega_{pl}^4}{2\omega_q\left(\omega_q^2-\omega_+^2\right)\left(\omega_q^2-\omega_-^2\right)}\left(\frac{n_B(\omega_q)}{(\omega_G-\omega_q)}+\frac{1+n_B(\omega_q)}{\omega_G+\omega_q}\right)\\
=& -B_q\left(\frac{n_B(\omega_q)}{(\omega_G-\omega_q)}+\frac{1+n_B(\omega_q)}{\omega_G+\omega_q}\right).
\end{aligned}$$

We also define for brevity of notation the triplet of poles $\omega_j=(\omega_q,\omega_-,\omega_+)$ as $\omega_j=(\omega_1,\omega_2,\omega_3)$; in this way, the coefficient B_j is written as

$$B_j=\frac{\left(\omega_j^2+\Omega_{pl}^2-\omega_q^2\right)^2}{2\omega_j\prod\limits_{n\neq j,\,n=1}^{3}\left(\omega_j^2-\omega_n^2\right)}. \tag{6.45}$$

Now we can evaluate the residue in the pole $z = \omega_G$. From Eq. (6.41) we have

$$\text{Res}(f\, n_B)_{\omega_G} = \frac{(\omega_G^2 - \omega_q^2 + \Omega_{pl}^2)^2}{\left(\omega_G^2 - \omega_q^2\right)\left(\omega_G^2 - \omega_+^2\right)\left(\omega_G^2 - \omega_-^2\right)} n_B(\omega_G) \tag{6.46}$$

We have to evaluate $n_B(\omega_G)$, keeping in mind that this is a *fermionic* frequency, hence an odd multiple of $i\pi/(\beta\hbar)$ [114, Chap. 3]. We obtain the Fermi occupation factor, as expected:

$$\begin{aligned}
n_B(\omega_G) =& \frac{1}{\exp\left(\beta\hbar\omega_G\right) - 1} \\
=& \frac{1}{\exp\left[\beta\hbar\left(-i\omega + \omega_{kq} + \frac{\mu}{\hbar}\right)\right] - 1} \\
=& \frac{1}{\exp\left(-i\beta(\hbar\omega - \frac{\mu}{\hbar})\right)\exp\left(\beta\hbar\omega_{kq}\right) - 1} \\
=& \frac{1}{\exp\left(i(2n+1)\pi\right)\exp\left(\beta\hbar\omega_{kq}\right) - 1} \\
=& \frac{1}{-\exp\left(\beta\hbar\omega_{kq}\right) - 1} \\
=& -\, n_F\left(\omega_{kq}\right)
\end{aligned}$$

where

$$n_F(x) = \frac{1}{\exp\left(\beta\hbar x\right) + 1} \tag{6.47}$$

is the definition of the Fermi-Dirac distribution for the frequency x. With the same procedure illustrated for the other poles, we can now rewrite the residue in a different way:

$$\begin{aligned}
\text{Res}(f\, n_B)_{\omega_G} =& -\frac{(\omega_G^2 - \omega_q^2 + \Omega_{pl}^2)^2}{\left(\omega_G^2 - \omega_q^2\right)\left(\omega_G^2 - \omega_+^2\right)\left(\omega_G^2 - \omega_-^2\right)} n_F\left(\omega_{kq}\right) \\
=& -\left(\frac{F_+}{\omega_G - \omega_+} + \frac{F_-}{\omega_G - \omega_-} + \frac{F_q}{\omega_G - \omega_q}\right) n_F\left(\omega_{kq}\right) + \\
&+ \left(\frac{F_+}{\omega_G + \omega_+} + \frac{F_-}{\omega_G + \omega_-} + \frac{F_q}{\omega_G + \omega_q}\right) n_F\left(\omega_{kq}\right)\ ,
\end{aligned}$$

having defined the coefficients F_j as we did for B_j,

$$F_j = \frac{\left(\omega_G^2 + \Omega_{pl}^2 - \omega_q^2\right)^2}{2\omega_j \prod\limits_{n\neq j,\, n=1}^{3}\left(\omega_G^2 - \omega_n^2\right)}\ . \tag{6.48}$$

Now we can write the sum of all residues, obtaining (we remind that the frequencies ω_j are the three frequencies $\omega_+, \omega_-, \omega_q$):

$$\begin{aligned}\sum_{m=-\infty}^{\infty} f\left(\omega_m\right) =& -\beta\hbar \sum_{\omega_m} \frac{\left(\omega_m^2 - \omega_q^2 + \Omega_{pl}^2\right)^2}{\left(\omega_m^2 - \omega_q^2\right)\left(\omega_m^2 - \omega_+^2\right)\left(\omega_m^2 - \omega_-^2\right)\left(i\omega_m - i\omega + \omega_{kq} + \frac{\mu}{\hbar}\right)} = \\ =& -\beta\hbar \sum_{j=1}^{3} \left\{ -B_j \left(\frac{n_B(\omega_j)}{(\omega_G - \omega_j)} + \frac{1 + n_B(\omega_j)}{\omega_G + \omega_j} \right) - \frac{F_j n_F\left(\omega_{kq}\right)}{\omega_G - \omega_j} + \frac{F_j n_F\left(\omega_{kq}\right)}{\omega_G + \omega_j} \right\} \\ =& \beta\hbar \sum_{j=1}^{3} \frac{B_j[1 + n_B(\omega_j)] - F_j\, n_F\left(\omega_{kq}\right)}{\omega_G + \omega_j} + \beta\hbar \sum_{j=1}^{3} \frac{B_j n_B(\omega_j) + F_j\, n_F\left(\omega_{kq}\right)}{\omega_G - \omega_j}\,.\end{aligned}$$

The sum over residues on these poles forms a very general self-energy expression in which Bose and Fermi occupation factors n_B and n_F arise in a natural way [132, 133]. Each residue gives an additive contribution to the self-energy. In this way the self-energy has three additive contributions for emission and three for absorption of a phonon-plasmon. As it has been shown, the pole in ω_G contributes to each phonon–plasmon mode $\pm\omega_j$ both in emission and in absorption, in a way that depends on the considered scattering process.

The first sum in the last line describes the *emission* of a LO-phonon (the capture process), the second sum describes the escape. We can write together the two contributions as

$$\sum_{m=-\infty}^{\infty} f_{\pm}\left(\omega_m\right) = \beta\hbar \sum_{j=1}^{3} \frac{B_j[\frac{1}{2} \pm \frac{1}{2} + n_B(\omega_j)] - F_j\, n_F\left(\omega_{kq}\right)}{\omega_G \pm \omega_j} \tag{6.49}$$

Substituting now this expression in Eq. (6.34), the self–energy reads:

$$\begin{aligned}\Sigma_r(k,\omega) =& \frac{2\pi q_e^2 K_\epsilon \omega_{LO}^2}{\epsilon_0 \epsilon_s} \sum_{q_\parallel} \frac{I\left(q_\parallel\right)}{q_\parallel} \\ & \times \sum_{j=1}^{3} \frac{B_j[\frac{1}{2} \pm \frac{1}{2} + n_B(\omega_j)] - F_j\, n_F\left(\omega_{kq}\right)}{\omega_G \pm \omega_j}.\end{aligned}$$

Converting the 2D $q_\parallel$–summation into an integral, exploiting the translational invariance, we obtain

$$\begin{aligned}\Sigma_{r\pm}(k,E) =& \frac{K_\epsilon \omega_{LO}^2 \alpha_0 \hbar\, c}{\pi \epsilon_s} \int_0^{2\pi} d\theta \int_0^{\infty} q_\parallel\, dq_\parallel\, \frac{I\left(q_\parallel\right)}{q_\parallel} \\ & \times \sum_{j=1}^{3} \frac{\left(\frac{1}{2} \pm \frac{1}{2} + n_{Bj}\right) B_j \mp n_F\left(\omega_{kq}\right) F_j}{\omega_G \pm \omega_j}\end{aligned} \tag{6.50}$$

where α_0 is the fine structure constant, and θ is the angle between $q_\parallel$ and k. In all what follows, for brevity n_F and n_{Bj} will be written instead of $n_F(\omega_{kq})$ and $n_B(\hbar\omega_j)$; furthermore, the *upper* (*lower*) sign in $\Sigma_{r\pm}$ refers to emission (absorption) of a phonon–plasmon quantum of energy.

The quantities B_j and F_j come respectively from the residues in $\pm\omega_j$ and ω_G and they are not present in literature about scattering times: this because they do not appear in calculations making use of unscreened or statically–screened Coulomb interaction, whereas in totally numerical calculations their role is obscured in the general formulas. They should be intended as correcting terms that multiply the usual Bose and Fermi distributions, representing the interplay between dynamical screening and statistical occupancy factors.

In order to obtain semi–analytic results, it is more convenient to change the integration variable to ω_G, exploiting its definition given in Eq. (6.42), obtaining $q_\parallel$ as a function of ω_G

$$q_\parallel(\omega_G) = k\,\cos(\theta) + \sqrt{\frac{2m\,(E-\mu+\hbar\omega_G)}{\hbar^2} - k^2\sin^2(\theta)} \tag{6.51}$$

Then the integration can be extended to the negative axis, dividing the integrand by 2:

$$\begin{aligned}\Sigma_{r\pm}(k,E) = &\frac{K_\epsilon\omega_{LO}^2\alpha_0\hbar\,c}{2\pi\epsilon_s}\int_0^{2\pi}d\theta\int_{-\infty}^{\infty}d\omega_G\left(-\frac{2m}{\hbar}\right)\frac{I\,(\omega_G)}{2\sqrt{\frac{2m(E-\mu+\hbar\omega_G)}{\hbar^2}-k^2\sin^2(\theta)}}\\ &\times\sum_{j=1}^{3}\frac{\left(\frac{1}{2}\pm\frac{1}{2}+n_{Bj}\right)B_j\mp n_F\,(\omega_{kq})\,F_j}{\omega_G\pm\omega_j}\end{aligned} \tag{6.52}$$

where the square root come from the Jacobian $d_\parallel/d\omega_G$ arising from the change of integration variable. Simplifying we obtain

$$\begin{aligned}\Sigma_{r\pm}(k,E) = &-\frac{K_\epsilon\omega_{LO}^2\alpha_0 m\hbar\,c}{2\pi\epsilon_s}\int_0^{2\pi}d\theta\int_{-\infty}^{\infty}d\omega_G\frac{I\,(\omega_G)}{2\sqrt{\frac{2m(E-\mu-\hbar\omega_G)}{\hbar^2}-k^2\sin^2(\theta)}}\\ &\times\sum_{j=1}^{3}\frac{\left(\frac{1}{2}\pm\frac{1}{2}+n_{Bj}\right)B_j\mp n_F\,(\omega_{kq})\,F_j}{\omega_G\pm\omega_j}.\end{aligned} \tag{6.53}$$

Now the integration can be extended to the complex plane, adding a counterclockwise semicircle at infinity in the upper plane giving no contribution. The integration in ω_G can now be done considering the formal Dirac's identity

$$\lim_{\eta\to 0+}\frac{1}{x+i\eta} = P\left(\frac{1}{x}\right) - i\pi\delta(x) \tag{6.54}$$

and observing that the argument of Dirac–δ is a function with simple zeroes only, given by roots of equation

$$\omega_G \pm \omega_j\,(\omega_G) = 0\,, \tag{6.55}$$

whose solutions (let us name them $\hat{\omega}_j$) can be easily found numerically, resulting functions of N, E and θ, with values that depend on the sign considered in Eq. (6.50). If the minus

sign is chosen, Eq. (6.55) gives two positive roots, ω_-^{rel} and ω_+^{rel}, representing the lower and upper *relaxation* frequencies. Instead, if the plus sign is chosen, the two positive roots represent the two *excitation* frequencies ω_-^{exc} and ω_+^{exc}. The contribution to the imaginary part of the self-energy due to the mode $\hat{\omega}_j$ results

$$\Im\Sigma_{r\pm}^{j}(k,E,\hat{\omega}_j) = \frac{K_\epsilon \omega_{LO}^2 \alpha_0 m\hbar\, c}{2\epsilon_s}\int_0^{2\pi} d\theta \times \frac{I\left(\hat{\omega}_j\right)\left[\left(\frac{1}{2}\pm\frac{1}{2}+n_{Bj}\right)B_j \mp n_F F_j\right]\Theta\left(E-\mu\Theta(\mu)\mp\hbar\hat{\omega}_j\right)}{\sqrt{2m\left(E-\mu\mp\hbar\hat{\omega}_j\right)-\hbar^2k^2\sin^2(\theta)}}, \tag{6.56}$$

in which the θ integration is only a matter of software implementation (here Θ is the Heaviside step-function). The Fermi function must be evaluated in the pole as:

$$n_F\left(\omega_{kq}\right) = \frac{1}{\exp\left[\beta\left(E-\mu-\hbar\hat{\omega}_j\right)\right]+1}. \tag{6.57}$$

The *relaxation* and *excitation* times (respectively $\tau_\pm^{\text{rel}}$ and $\tau_\pm^{\text{exc}}$) corresponding to $\omega_\pm^{\text{rel}}$ and $\omega_\pm^{\text{exc}}$ are given by

$$\begin{aligned} 1/\tau_\pm^{\text{rel}} &= \frac{2}{\hbar}\,\Im\Sigma_{r+}^{\pm}(k,E) \\ 1/\tau_\pm^{\text{exc}} &= \frac{2}{\hbar}\,\Im\Sigma_{r-}^{\pm}(k,E)\,. \end{aligned} \tag{6.58}$$

In the remainder of this Section, it is examined the limiting case which is not complicated by this integration, considering $\theta = 0$; the development of the calculation depends now on the particular scattering process under consideration. It is convenient to stress that B_j, F_j, n_{Bj}, n_F and $\hat{\omega}_j$ are all functions of N, E, k and θ through Eq. (6.51) evaluated at the integrand pole in Eq. (6.50).

6.4.3 Intraband relaxation in quantum wells

A quasi–two–dimensional system is considered, made of an unstrained and undoped InGaAs single quantum well at room temperature having a width of 8 nm, whose states of energy $E(k)$ have a parabolic dispersion kinetic energy $E(k) = E_n + \hbar^2k^2/(2m)$; E_n are the stationary Schrödinger equation eigenvalues. Effective perpendicular carrier masses m and quantized level positions E_n in the well have been computed according to Ref. [134], finding $m = 0.047\, m_0$ and $E_0 = 51$ meV above the bottom of QW conduction band (m_0 is the free electron mass). The confinement is obtained with InP barriers and the conduction band offset between well and barrier results to be 249 meV (therefore, the binding energy of the electron level is 198 meV). The form factor $I(q_\parallel)$ has been calculated according to Ref. [131] just as an example: different and more rigorous forms

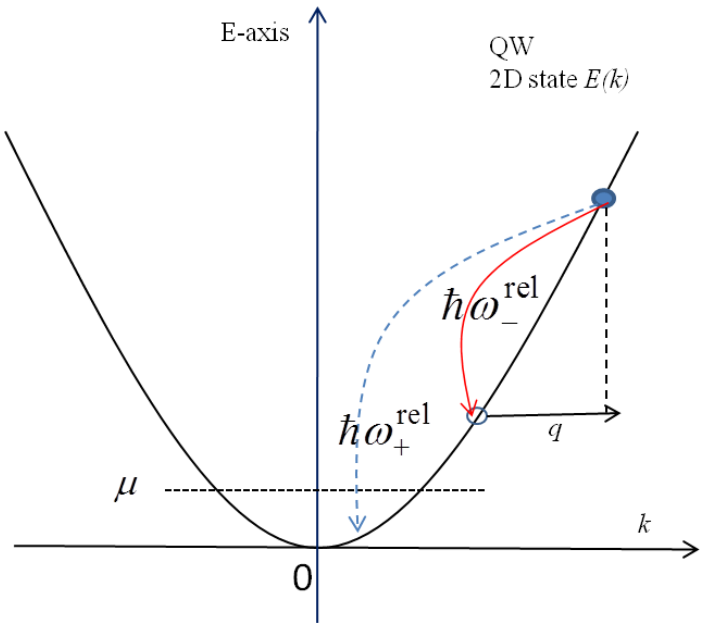

Figure 6.9. Relaxation of an electron with energy $E(k)$ by the emission of a phonon–plasmon. In the example, the mode ω_-^{rel} (blue solid arrow) is allowed, whereas ω_+^{rel} (red dashed) is forbidden.

can be employed, depending on what material and band structure is considered, nothing changing of the present formalism. Without losing generality, it has been set $E_n = 0$, moving the energy reference on the confined level.

As already stressed, an electron can lose energy emitting phonon–plasmons of frequency ω_-^{rel} (lower plasmon mode) and ω_+^{rel} (higher plasmon mode). The intraband relaxation time for the emission of a phonon–plasmon of frequency ω_j^{rel} (here $j = \pm$) is then given by Eq. 6.58 and it is important to point up the fact that for zero carrier density it is found $B_+ = 0$, $B_- = F_- = 1/(2\omega_{LO})$, hence the self–energy recovers an usual form [135]. Furthermore, it must be stressed that each phonon–plasmon mode ω_j^{rel} separately contributes to the self–energy and the restriction imposed by the energy conservation separately applies to each of them.

Considering for simplicity the zero temperature limit where the Bose term $n_{Bj} = 0$, the factor $1 - n_F$ in the numerator assures that a real phonon–plasmon can be emitted only if the final electron state is allowed. This means that the electron initial energy must be at least $\hbar\omega_\pm^{rel}$ above the lowest available state, as shown in Fig. (6.9), that is:

$$E - \mu\,\Theta(\mu) > \hbar\omega_\pm^{rel}\,. \tag{6.59}$$

Frequencies ω_-^{rel} and ω_+^{rel} have been calculated versus N and shown in Fig. 6.10 for the given structure, for $E(k) = 150\,\text{meV}$ (for low carrier density, $\omega_-^{rel} \approx \omega_{LO}$, around

35 meV in this material at $T = 300\,\mathrm{K}$). The energy conservation imposes a maximum allowed frequency transition, reported in the figure as well (dash–dotted line), defined by Eq. (6.59). It can be observed that, if N is below $\approx 1.5 \times 10^{11}\,\mathrm{cm}^{-2}$, the relaxation via LO-phonon emission is the only allowed: the higher plasma mode frequency ω_+^{rel} lies above the $(E - \mu\,\Theta(\mu))/\hbar$ line, so it is forbidden, whereas for density $1.5 \times 10^{11} \lessapprox N \lessapprox 1.5 \times 10^{12}\,\mathrm{cm}^{-2}$ both modes are allowed. All this reflects in the corresponding electron relaxation times τ_-^{rel} and τ_+^{rel} calculated by Eq. (6.58) and shown in Fig. 6.11 as a function of two–dimensional carrier density, for $\theta = 0$ (only for simplicity) for $E(k) = 50\,\mathrm{meV}$ and $E(k) = 150\,\mathrm{meV}$. Commenting the case for $E(k) = 150\,\mathrm{meV}$, the lower plasma mode is the only one allowed if $N \lessapprox 1.5 \times 10^{11}\,\mathrm{cm}^{-2}$ (the dash-dotted

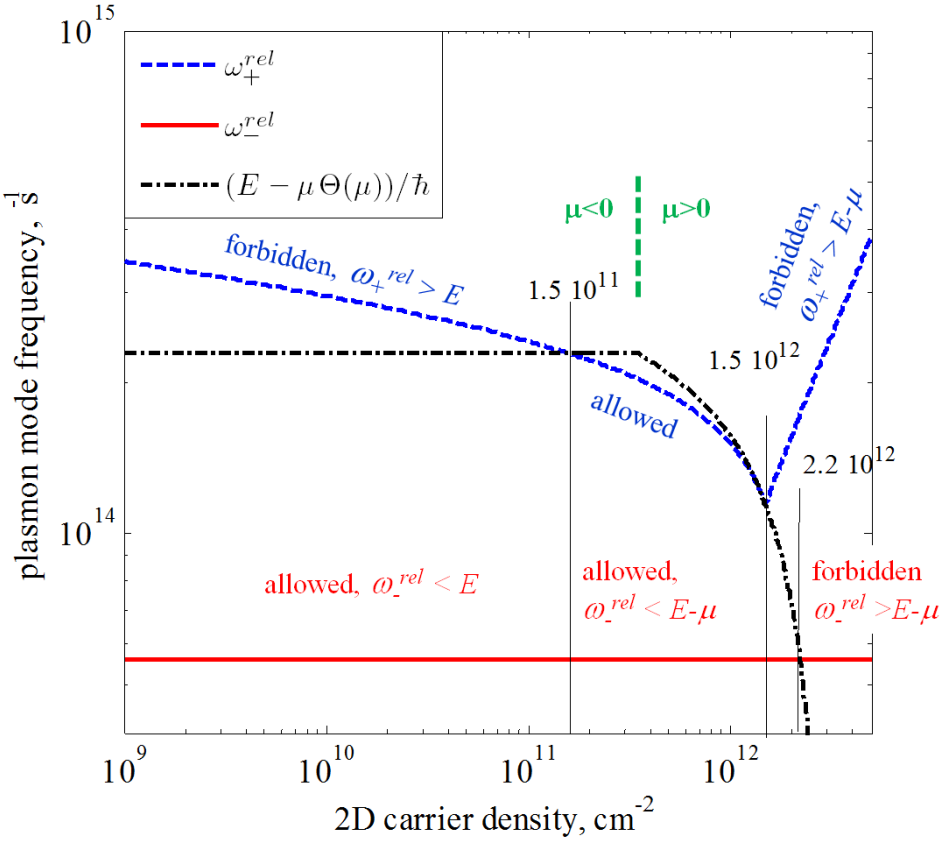

Figure 6.10. Lower (red solid) and higher (blue dashed) plasma modes ω_-^{rel} and ω_+^{rel}, with the upper allowed frequency transition $(E(k) - \mu\,\Theta(\mu))/\hbar$, imposed by energy conservation (dash–dotted line). The allowed and forbidden regions for each transitions are marked in the graph, where it is also indicated the N value above which Fermi energy enters in conduction band (CB), $\mu > 0$. The values of carrier density $1.5 \times 10^{11}\,\mathrm{cm}^{-2}$, $1.5 \times 10^{12}\,\mathrm{cm}^{-2}$ and $2.2 \times 10^{12}\,\mathrm{cm}^{-2}$ discussed in the text are indicated as well with vertical black lines.

line), whereas for $1.5 \times 10^{11} \lessapprox N \lessapprox 1.5 \times 10^{12}\,\mathrm{cm}^{-2}$ both modes are allowed (although $\omega_{-}^{\mathrm{rel}}$ remains the dominant mode, being the fastest one). Infact, looking at Fig. 6.10, it can be observed that in this interval of N both modes lie *below* the upper allowed frequency transition (the dash–dotted line, $E(k) - \mu\Theta(\mu)$). When carrier density increases, the relaxation via $\omega_{+}^{\mathrm{rel}}$ plasma mode becomes increasingly competitive, reducing its scattering time by two orders of magnitude. The higher plasma mode $\omega_{+}^{\mathrm{rel}}$ becomes again forbidden when the electron Fermi energy μ rises above $E(k) - \hbar\omega_{+}^{\mathrm{rel}}$ (this happens for carrier density above $\approx 1.5 \times 10^{12}\,\mathrm{cm}^{-2}$), and relaxation via the lower plasma mode $\omega_{-}^{\mathrm{rel}}$ becomes again the only allowed process. In the end, when carrier density approaches $2.2 \times 10^{12}\,\mathrm{cm}^{-2}$, no relaxation is possible via any LO-phonon-plasmon mode. Of course other and more efficient decaying channels are dominant in this regime (e.g. electron-electron scattering) but they are not addressed in this work.

The curve for $E(k) = 50\,\mathrm{meV}$ shows a strong resonance at $N \approx 4.7 \times 10^{11}\,\mathrm{cm}^{-2}$ (red

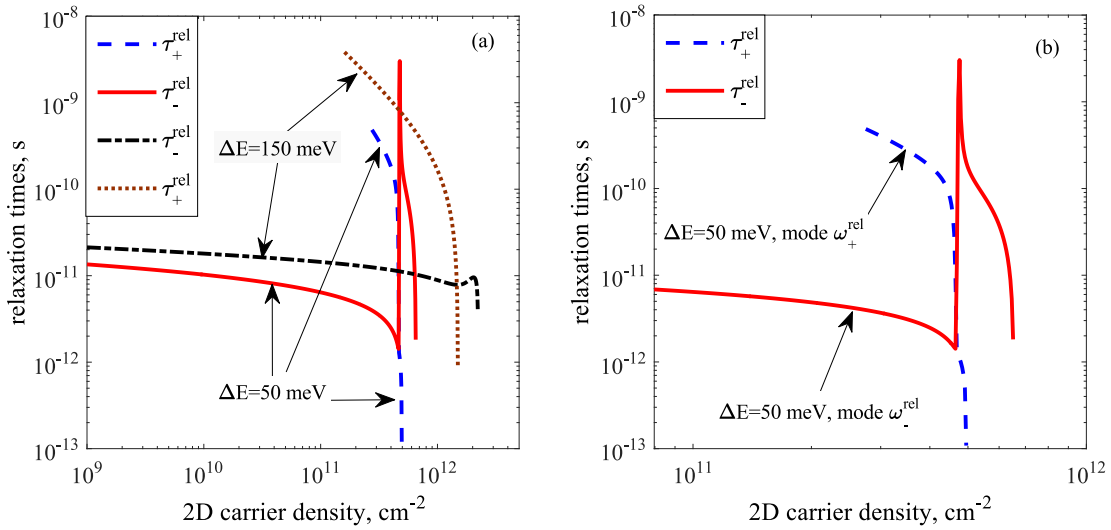

Figure 6.11. (a) Relaxation times τ_{-}^{rel} (solid and dash-dotted lines) and τ_{+}^{rel} (dashed and dotted lines) for an electron in a confined state of a 8 nm wide quantum well, with a kinetic energy respectively of $E(k) = 50$ and $150\,\mathrm{meV}$. (b) Detail of the case for $E(k) = 50\,\mathrm{meV}$, showing the exchange of the two modes.

solid line in Fig. 6.11); this feature takes place whenever, for some value of carrier density N, the two modes $\omega_{-}^{\mathrm{rel}}(N)$ and $\omega_{+}^{\mathrm{rel}}(N)$ take the same value and are both allowed. For this value of N the two plasmon-modes exchange their roles and the self-energy related to $\omega_{-}^{\mathrm{rel}}$ has a zero, corresponding to a (virtually) infinite lifetime. The other mode (dashed line in Fig. 6.11(b)) takes its role and the total relaxation time τ^{tot} follows the fastest of the two modes with continuity. The reduction of relaxation times $\tau_{\pm}^{\mathrm{rel}}$ for increasing carrier density, a quite counter–intuitive feature, is worth of an explanation: until

the Fermi filling effect is not dominant (below $N = 10^{12}\,\mathrm{cm}^{-2}$), the dependence of $\tau_{\pm}^{\mathrm{rel}}$ on the factor $\sqrt{2m\left(E(k) - \mu(N) - \hbar\hat{\omega}_{\pm}\right) - \hbar^2 k^2 \sin^2(\theta)}$ determines their reduction for increasing values of N through the rise of the Fermi energy μ, favoring the phonon–plasmon emission probability. Instead, when carrier density increases above $10^{12}\,\mathrm{cm}^{-2}$ (in the given example), Fermi band filling reduces the density of available final states, making the relaxation time to increase up to a value of N above which the relaxation becomes forbidden.

6.4.4 Capture time into quantum wells

It is consider here the quantum capture of an electron from a 3D barrier state with momentum $k = 0$, to a bound quantum well electron state, whose energy is ΔE below the bottom of barrier conduction band (where there is the reference of energy), through the emission of a phonon–plasmon. Following the shown path, the relaxation via the modes $\omega_{\pm}^{\mathrm{rel}}$ must now be intended as giving origin to capture processes, so for coherence it is convenient to call them $\omega_{\pm}^{\mathrm{cap}}$. The imaginary part of the self-energy is given by, after the integration in the scattering angle θ:

$$\begin{aligned} \Im\Sigma_{r+}^{\pm} = \quad & \frac{K_\epsilon \omega_{LO}^2 \sqrt{m}\alpha_0 \hbar\, c\pi}{\sqrt{2}\epsilon_s} \\ \times \quad & I\left(\omega_{\pm}^{\mathrm{cap}}\right) \frac{\Theta\left(\Delta E - \mu - \hbar\omega_{\pm}^{\mathrm{cap}}\right)}{\sqrt{\left(\Delta E - \mu\Theta(\mu) - \hbar\omega_{\pm}^{\mathrm{cap}}\right)}} \\ \times \quad & \left[\left(1 + n_{B\pm}\right) B_{\pm} - n_F F_{\pm}\right] . \end{aligned} \tag{6.60}$$

The corresponding capture times of the two modes are given by $1/\tau_{\pm}^{\mathrm{cap}} = 2/\hbar\,\Im\Sigma_{r+}^{\pm}$ similarly to $\tau_{\pm}^{\mathrm{rel}}$, whereas the total capture time is given by $\tau_{\mathrm{tot}}^{\mathrm{cap}} = 1/[1/\tau_{+}^{\mathrm{cap}} + 1/\tau_{-}^{\mathrm{cap}})]$. The position $k = 0$ does not result in a limitation: an initial thermalized state it is considered, but only to better show the mechanism: nevertheless Eq. (6.56) (which the calculation is based on) can be employed for all possible values of k. In Fig. 6.12 two different situations are considered: the first one (Fig. 6.12(a)) describes the case in which the lower plasmon mode is allowed, because $\Delta E > \hbar\omega_{-}^{\mathrm{cap}}$. The higher plasmon mode is forbidden, given that no available states exist at $E = -\hbar\omega_{+}^{\mathrm{cap}}$ (for simplicity in the drawing the Fermi energy is considered to lie below the QW state).
The second situation, shown in Fig. 6.12(b), considers the case in which ΔE is greather than both $\hbar\omega_{\pm}^{\mathrm{cap}}$: in this case, capture can occur via both $\omega_{\pm}^{\mathrm{cap}}$ throught the emission of phonon–plasmons of momentum $q_{\parallel}$. The minimum capture time is expected for $\Delta E \approx \hbar\omega_{-}^{\mathrm{cap}}$, given that in this case it can occur also for $q \approx 0$.

In Fig. 6.13 the capture time is reported as function of ΔE, for some values of carrier

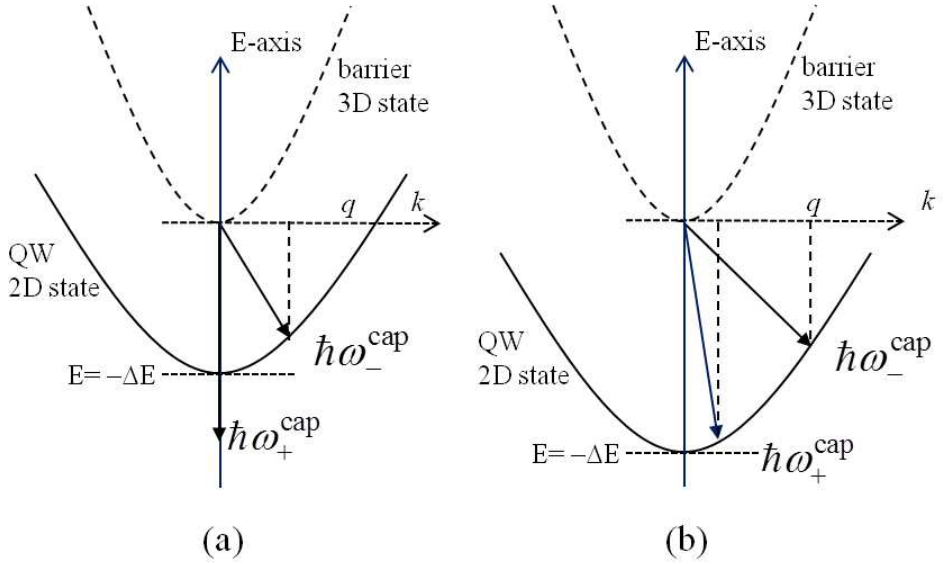

Figure 6.12. Capture scattering scheme from a parabolic E_k bulk barrier state (dashed) to a quantum well state (solid) with a binding energy ΔE. (a) No capture is possible via phonon-plasmon emission with frequency $\omega_+^{\rm cap}$, whereas capture via the $\omega_-^{\rm cap}$ mode is allowed with the emission of a phonon-plasmon of momentum $q_\parallel$. (b) In this case the capture is possible through both plasmon modes.

density N. The increase of ΔE can be thought of as an increase of quantum well width or barrier height, progressively confining more the bound level. For $N = 10^9$ and 10^{10}cm^{-2} the capture through $\omega_-^{\rm cap}$ can happen only for ΔE above $\approx$ 39 meV, because of energy conservation prescription. The capture through $\omega_+^{\rm cap}$ can happen only if ΔE is above $\approx 80 - 90$ meV. However, in this case the capture time through $\omega_+^{\rm cap}$ is in the order of 10^{-8} s and it does not contribute in practice. Instead, considering the cases $N = 10^{11}$ and 10^{12}cm^{-2}, at low energy $\omega_+^{\rm cap}$ is the only allowed mode: nevertheless it corresponds to a quite long capture time. The large reduction of capture time occurring respectively around ≈ 40 and ≈ 90 meV is due to $\omega_-^{\rm cap}$, becoming allowed as soon as ΔE overcomes it, constituting the feature known as resonant capture [110, 121].

Concerning the momentum $q_\parallel$ of the emitted phonon–plasmon, it must be stressed that this is the value at the *pole* of Eq. (6.50) integrand: from a quantum point of view, there are infinitely virtual modes of emission from which, integrating over all possible values of $q_\parallel$, the value at the *pole* results as a consequence of energy–momentum conservation laws.

In Fig. 6.14 the total capture time versus carrier density N is reported. In the upper graph (a), capture time is shown as function of N for several values of ΔE. The lines for $\Delta E = 50,\ 100$ and 250 meV have got contributions from both $\omega_-^{\rm cap}$ and $\omega_+^{\rm cap}$, whereas the line for $\Delta E = 30$ meV gets contribution only from $\omega_+^{\rm cap}$, in the N–interval in which the

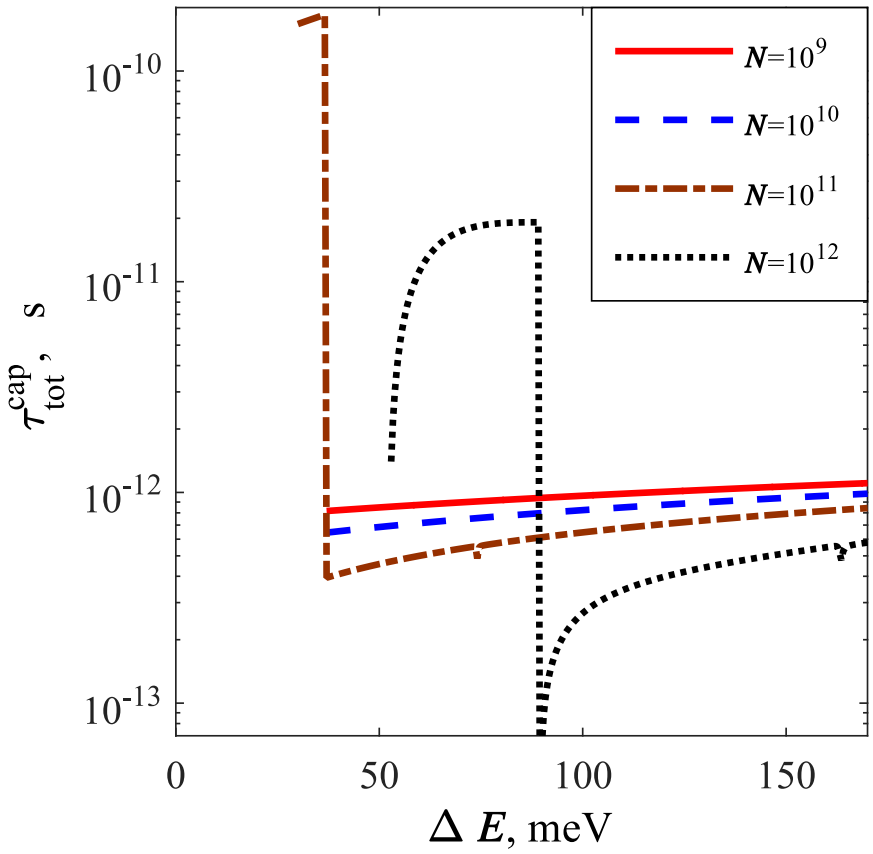

Figure 6.13. Total capture time versus quantum well binding energy, for several values of 2D carrier density N (in the legend, values are in cm^{-2}).

transition is allowed.
It is interesting to observe again that capture time is proportional to the factor $\sqrt{\Delta E - \mu - \hbar\omega_{\pm}^{\mathrm{cap}}}$, so it is an increasing function of ΔE and a *decreasing* function of N, through the Fermi energy μ (in agreement with Ref. [110], for example). The explanation of the general behavior is easily attained: considering as an example the case for $\Delta E = 50$ meV, described by the dashed line in Fig. 6.14(a) (blue line), in Fig. 6.14(b) τ_{-}^{cap} and τ_{+}^{cap} are separately reported: for $N = 4 \times 10^{10}$ cm^{-2} the two plasmon modes exchange their roles, but the total capture time practically follows the lowest curve (that is, the fastest capture mode) until it exists. When the dashed line ends (that is, for N higher than 4.5×10^{11} cm^{-2}), the slowest mode becomes the only allowed one, giving origin to the sudden increase of capture time (well visible in Fig. 6.14(a) as well).
Still considering Fig. 6.14(a), the most visible effect of raising the confinement energy

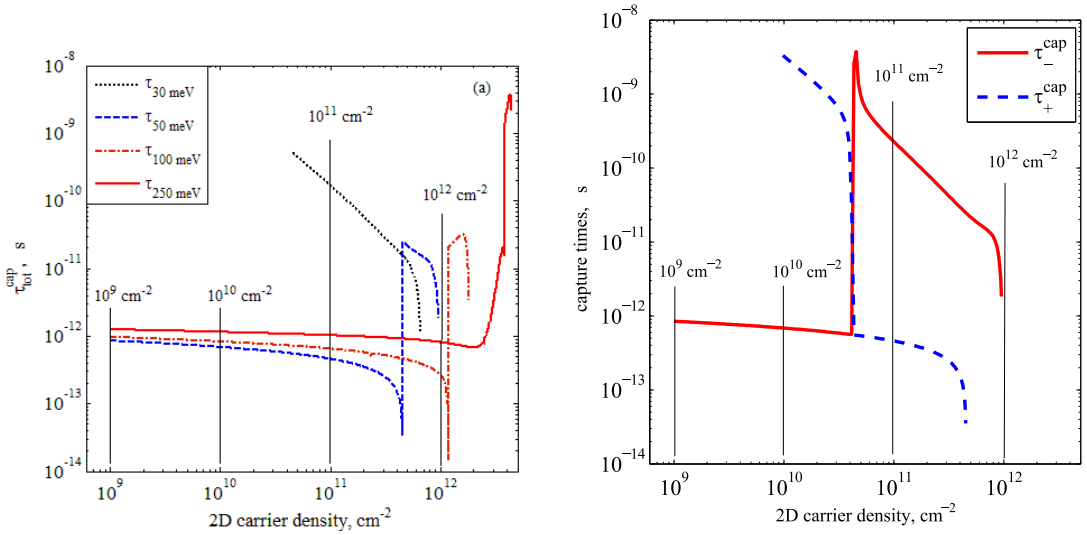

Figure 6.14. (a) Total capture time versus 2D carrier density, for several values of quantum well state binding energy (30, 50, 100 and 250 meV). (b) Capture times $\tau_-^{\rm cap}$ and $\tau_+^{\rm cap}$ versus carrier density (respectively via $\omega_-^{\rm cap}$ and $\omega_+^{\rm cap}$) for $\Delta E = 50\,$meV. The vertical black lines correspond to density values as in Fig. 6.13.

ΔE consists in an increase of the N-point value at which the $\omega_-^{\rm cap}$ mode becomes forbidden, leaving the $\omega_+^{\rm cap}$ mode as the only possible one (this points is clearly visible in the figure for the curves $\Delta E = 50\,$meV and $\Delta E = 100\,$meV as a sharp "cusp"). Another effect, visible at low-medium values of carrier density, is a small increase of the total capture time for a given value of carrier density.
However there is another more interesting effect better visible when higher values of ΔE are considered: as already pointed out in Sec. 6.4.3, the reduction of relaxation/capture times for increasing carrier density (anti-screening effect) is a counter-intuitive effect that comes from the rise of the Fermi energy μ, favoring the phonon–plasmon emission probability. This effect can be better appreciated in the curve for $\Delta E = 250\,$meV that is also shown on a linear time-scale in Fig. 6.15: the decrease continues until $N \lesssim 2\times10^{12}\,$cm^{-2}, then the Fermi band filling prevails and the capture time fast increases.

6.4.5 Final remarks

In the present work the calculation of intraband relaxation and carrier capture into quantum wells has been addresed, obtaining semi–analytic expressions as function of carrier density with dynamic SPP dielectric function, a topic often neglected in literature, despite its fundamental interest in semiconductor physics.

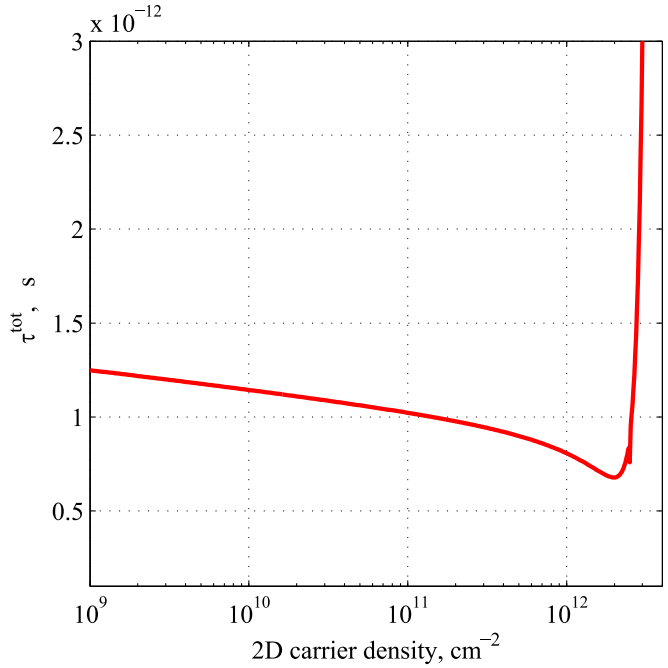

Figure 6.15. Total capture time versus 2D carrier density, for $\Delta E = 250$ meV, where the anti-screening effect is better visible.

This approach has allowed to point out the role of coupling between LO–phonon and plasmon, describing in a comprensible way how the carrier density affects the most effective phonon–plasmon scattering mode. The increase of carrier density makes the intraband relaxation and interband capture time firstly to reduce, because of a resonance between $E - \mu$ and the energy of the phonon–plasmon $\hbar\omega_j^{\text{rel}}$ (or $\hbar\omega_j^{\text{cap}}$ for the capture process), and successively to increase, because of Fermi band filling increase.
Furthermore, the Bose and Fermi occupation factors appear to be corrected by factors expressing the interplay with the dynamical screening, a feature not present in literature due to the scarce attention to these issues. All these facts and the figures presented in this Section demonstrate how inadequate can be to keep scattering times as constant respect to N and $E(k)$ in modeling calculations.

Population effects (dynamical screening of interacting potential and Fermi occupation factors) must be kept into account especially in laser and semiconductor optical amplifier (SOA) dynamics, photodetector efficiency calculations, etc. For many high-speed applications, SOAs must have a fast gain recovery to avoid system penalties arising from bit pattern dependencies. The gain recovery of SOAs is limited by the carrier lifetime, which itself depends on the QW carrier density, determined by a non-trivial interplay of applied current and optical intensity in the active layer. Furthermore SOAs and LEDs based on

multiple-QW active layer normally require as uniform as possible QW carrier population, that dynamically changes during input-signal and strongly depends on capture and escape times. Therefore their nonlinear behavior with carrier density must be considered in the formulation of a rate-equation set that realistically pretend to describe the system performances.

As stressed in the introduction, in literature there are many works addressing capture times in quantum wells. Most of them focus on rigorous multi-quantum-well subband calculations, evaluating the relaxation rate among them, but without addressing population effects. Good examples of them are Ref. [118–121]. It is not possible to compare directly our work with those, because of their different scope: in fact their main purpose consists in discussing the oscillating behavior of capture time with quantum well width. This is a very important point for laser and optical amplifier efficiency and it has received rightly much attention. Instead, there are few works addressing carrier density effects offering good examples of full numerical approach. Among them, good examples are represented by Ref. [110, 122]. They report calculations in the RPA using full multiple-subband and frequency-dependent screening, implementing very complete algorithms, finding an increasing capture rate for increasing carrier density, in qualitative agreement with our finding (they report total capture times $\tau_{\mathrm{tot}}^{\mathrm{cap}}$ versus QW width, so a comparison is not immediate; nevertheless the order of magnitude and the trend of $\tau_{\mathrm{tot}}^{\mathrm{cap}}$ with density N are consistent). In another good example of numerical calculation [136], intrasubband scattering rates are semi-analytically calculated in the static screening approximation of RPA and numerically with a full RPA, pointing up the necessity of using the dynamic dielectric function to describe the enhancement of the intrasubband scattering rates. Nevertheless they do not show neither scattering rates as explicit functions of carrier density, nor the relative contributions of the plasmonic modes separately, but only the total scattering rates, although in good qualitative agreement with our rates calculated as a function of confining energy (compare for example Fig. 3 in Ref. [136] and our Fig. 6.13).
In few words, the focus of the present work is on the exploitation of semi-analytic methods to address population effects, the interplay between occupancy factors and dynamical screening, and the coupling of plasmon modes, and it is possible to say that this analytic and other numerical approaches complement each other. In addition, the evaluation of phonon assisted Auger recombination [128, 129], in which the electron–phonon scattering self–energy is a crucial ingredient, may receive further understanding considering this approach, for its fast implementation in any numeric algorithm.

As a final remark, although the present method automatically includes one–phonon contributions from all orders in the carrier–phonon coupling constant (implicit in RPA formalism), it is convenient to point out that there are still many terms that have been omitted from the perturbation expansion. Besides the interactions with all other kind of phonons, it is worthwhile to mention the neglected contributions of multiple–phonon processes and the carrier–carrier scattering and its coupling with carrier-phonon scattering, that could be addressed in a future work.

Chapter 7

Electron capture time model: revision, validation and application to InGaN/GaN quantum wells

A revision of the quantum capture model discussed in Chapter 6 is presented, allowing to obtain an equivalent but more compact analytic formulation. The model has been applied to InGaN/GaN multi-quantum-well LEDs: employing simple two-population rate equations, the model allowed to reproduce available experimental data with excellent agreement, offering an accurate yet practical alternative to the usual approximation of a constant capture time in modeling light-emitting diodes and lasers.

The concepts here presented substantially follow a paper[1] we published in Phys. Status Solidi B, vol. 252, no. 5, pp. 971-976, 2015 [6].

7.1 Introduction

Nitride-based LEDs have emerged as an important technology for lighting applications, but their performance is limited by a reduction (droop) of their IQE as the driving current density is increased beyond $\approx 10\,\mathrm{A/cm^2}$, whose physical origin is still under intense debate [106, 130, 139–143]. IQE depends critically on carrier transport across the active region, and in particular is determined by a number of time scales, including all the radiative/nonradiative recombination lifetimes and the carrier capture time from the 3D unconfined states at or above the barrier energy to the confined 2D quantum well states.

[1] Experimental data we employed to test our model were published by W. H. Fan *et al.* in Ref. [137]. Authors published an erratum [138] which we were not aware at time of our publication (they exchanged the reported QW/barrier lengths). For this thesis, we repeated the fitting employing the correct QW/barrier length, however in order to obtain a satisfactory agreement we needed to consider carrier density values *in the barrier* well below the very high experimental values, about which we still have some doubts.

Capture in QWs may take place through carrier-carrier scattering, relaxation on defects, phonon emission, etc. Among these possibilities, in InGaN/GaN the capture time via polar LO-phonon emission $\tau_{c,ph}$, besides concurring to the determination of the total QW capture time τ_c to a considerable extent [137, 144], is interesting from a theoretical point of view because it is determined by the LO-phonon contribution to the electron self-energy Σ_{ph}, a quantity that also enters in the phonon-assisted Auger recombination lifetime, object of intensive investigations [129, 145–148]. As a consequence there have been considerable efforts to determine experimentally and theoretically the value of τ_c or $\tau_{c,ph}$. Among others, Mansour *et al.* [149] calculated $\tau_{c,ph}$ at zero carrier density in InGaN/GaN QWs with a Monte Carlo transport model, obtaining values oscillating with the QW thickness, a well known result already calculated and observed in GaAs-based materials [110, 119–121, 150–152]. J. Wang *et al.* [153] calculated $\tau_{c,ph}$ as a function of carrier density, including piezoelectric effects, Fermi occupation factors and many-body corrections to the band structure, but without any plasma screening of the electron-phonon interaction potential. X. Zhang *et al.* [154] exploited time- and wavelength-resolved cathodoluminescence to investigate the impact of $In_xGa_{1-x}N$ compositional fluctuations on capture efficiency. Ü. Özgür *et al.* [155] used differential transmission optical spectroscopy to measure τ_c in $In_{0.15}Ga_{0.85}N/In_{0.05}Ga_{0.95}N$ multiple QWs (MQWs) for several values of pump energy density. H.-C. Wang *et al.* [156] exploited femtosecond pump-and-probe technique to measure τ_c in InGaN thin films with In-rich nanoclusters. In all cited works, τ_c decreases for increasing pump energy density, but no estimates are given about the corresponding carrier density values. As an exception, W. H. Fan *et al.* [137] exploited time-resolved differential transmission spectroscopy to measure the carrier density dependence of τ_c for InGaN/GaN MQWs.

7.2 Model and method

In the present work the role of plasma screening on the Frölich interaction in InGaN/GaN QWs is investigated, considering an energy-, wavevector- and density- dependent dielectric function, and obtaining $\tau_{c,ph}$ in closed form *in presence of carrier population*, a topic not discussed yet in the way presented here. The spectral density function [114]

$$A(\boldsymbol{k},\omega) = \frac{-\frac{1}{\pi}\Im\Sigma_{ph}(\boldsymbol{k},\omega)}{\left[\hbar\omega - \hbar\omega_{\boldsymbol{k}} - \Re\Sigma_{ph}(\boldsymbol{k},\omega)\right]^2 + \left[\Im\Sigma_{ph}(\boldsymbol{k},\omega)\right]^2} \tag{7.1}$$

is related to $\Sigma_{ph}(\boldsymbol{k},\omega)$ and enters in the expression of $\tau_{c,ph}$ through $\tau_{c,ph}^{-1} = 2\Im\Sigma_{ph}/\hbar$. This function has been used also in phonon-assisted Auger recombination [129], allowing for the formal inclusion of phonon effects to the infinite order ($\boldsymbol{k}$ and $\omega_{\boldsymbol{k}}$ are electron wavevector and frequency dispersion expression). The presence of electron-phonon scattering gives rise to a shift in the position of the electron energy levels as well as to a level broadening, described respectively by $\Re\Sigma_{ph}$ (that can be safely set to zero as the

phonon contribution to the electronic structure is already included in the experimental band structure) and $\Im\Sigma_{ph}$.

The method presented here considers only the self-energy contribution due to LO-phonon emission/absorption, but can be extended to include the contribution of carrier-carrier scattering [157], possibly important for high values of carrier density. This point will be investigated in a future work. Nevertheless, despite its incompleteness, this formalism confirmes the remarkable observed dependence of τ_c on carrier density [137], and demonstrates the need to consider it not just as an external fitting parameter, but as a function of the device working point and characteristics. For all these reasons the development of simple but rigorous expressions of $\Im\Sigma_{ph}$ due to polar LO-phonon is of interest for next generation modeling tools suitable for the simulation of LED, lasers and other optoelectronic devices.

The QW capture process consists in a 3D barrier state electron of initial wavevector $\boldsymbol{k}$ that emits a quantum of energy $\hbar\omega_{\mathrm{m}}$ and wavevector $\boldsymbol{q} = \boldsymbol{k'} - \boldsymbol{k}$ ending in a 2D QW state of final wavevector $\boldsymbol{k'}$. Here $\boldsymbol{q} = \left(\overrightarrow{q_{\parallel}}, \overrightarrow{q_z}\right)$, where $\overrightarrow{q_{\parallel}}$ is the component of vector $\boldsymbol{q}$ in the QW plane, $\overrightarrow{q_z}$ is the component along the growth direction; the same notation applies to all other vectors. The self-energy associated to the process was evaluated through the expression

$$\Sigma_{ph}\left(\boldsymbol{k}, \omega\right) = -\frac{1}{\beta\hbar}\sum_{\boldsymbol{q},\omega_{\mathrm{m}}} V_{\mathrm{eff}}\left(\boldsymbol{q}, \omega_{\mathrm{m}}\right) G\left(\boldsymbol{k}, \boldsymbol{q}, \omega, \omega_{\mathrm{m}}\right) \ . \tag{7.2}$$

Here β is the inverse temperature in energy units, G is the dressed propagator [114, 117] of a carrier of momentum $\hbar\boldsymbol{k}$ and energy $\hbar\omega$, and V_{eff} is the density- and temperature-dependent plasma-screened potential interaction,

$$V_{\mathrm{eff}}\left(\boldsymbol{q}, \omega_m\right) = \int_{-\infty}^{\infty} dz\, \Phi^{*}_{\boldsymbol{k'}}\left(z\right) \frac{e^{iq_z z}}{2\pi} V_{\mathrm{ph}}(q_{\parallel}, q_z, \omega_m)\Psi_{\boldsymbol{k}}\left(z\right) \ , \tag{7.3}$$

where $q_{\parallel} = |\overrightarrow{q_{\parallel}}|$, $q_z = |\overrightarrow{q_z}|$, $\Psi_{\boldsymbol{k}}$ and $\Phi_{\boldsymbol{k'}}$ are the initial (barrier) and final (QW) electron wavefunctions, and $V_{\mathrm{ph}}(q_{\parallel}, q_z, \omega_m)$ is the plasma-screened electron-phonon potential interaction. $V_{\mathrm{ph}}(q_{\parallel}, q_z, \omega_m)$ was treated according to [114],

$$V_{\mathrm{ph}}(q_{\parallel}, q_z, \omega_m) = -\frac{4\pi q_e^2 K_{\epsilon}}{\epsilon_0 \epsilon_s \left(q_{\parallel}^2 + q_z^2\right)} \frac{D\left(\boldsymbol{q}, \omega_m\right)}{\epsilon^2\left(\boldsymbol{q}, \omega_m\right)}, \tag{7.4}$$

where $D\left(\boldsymbol{q}, \omega_m\right)$ is the screened phonon propagator [114]. The dielectric function $\epsilon(\boldsymbol{q}, \omega_m)$ was evaluated in the Single Plasmon Pole (SPP) description of the Lindhard's formula known as the RPA [114, 116], $\epsilon(\boldsymbol{q}, \omega_m)^{-1} = 1 + \Omega_{pl}^2\left(\omega_m^2 - \omega_q^2\right)^{-1}$, where carrier density of barrier states N_{barr} enters through $\omega_q\left(\boldsymbol{q}, N_{\mathrm{barr}}\right)$ and $\Omega_{pl}\left(\boldsymbol{q}, N_{\mathrm{barr}}\right)$, respectively the effective plasmon and the plasma frequency (their analytic forms are extensively discussed in [114, 116]). The customary static limit of $\epsilon(\boldsymbol{q}, \omega_m)$ will not be taken: this would be an easier task, but a large part of physical details would be lost. The limiting form at

zero carrier density of Eq. (7.4) is of course the standard Frölich interaction. It is important to stress that the phonon propagator $D(\boldsymbol{q}, \omega_m)$ was derived managing an original expression [114, Ch. 6] that contains self-energy terms arising from the electron gas polarization, $D = D_0/(1 - |M_q|^2 D_0 P/\epsilon)$, where $D_0 = \omega_{LO}^2/(\omega^2 - \omega_{LO}^2)$ is the unscreened phonon propagator, ω_{LO} is the polar LO-phonon frequency, $|M_q|^2$ and P are respectively the electron-phonon coupling matrix element and the polarization of the electron gas. This expression was derived from a Dyson's equation that already accounts for the coupling of phonon with the electron-electron interaction, and the result consists in the presence of the extra term $1/\epsilon$ in the phonon self-energy term $|M_q|^2 D_0 P/\epsilon$ (see [114, Ch. 6] for detailed derivation) that considerably weakens the self-energy effects.

Σ_{ph} was calculated by Eq. (7.2) exploiting the standard frequency ω_{m} summation in Matsubara formalism, Fermi n_F and Bose n_B occupation factors stemming during the summation in a natural way [114].

The $\boldsymbol{q}$-summation in Eq. (7.2) was converted to an integral exploiting the translational invariance, and the integration could be done analytically by residue theorems, after having extended the integration to the complex plane, without making any truncation in the $\boldsymbol{q}$-integration as customary done in numerical approaches (see details in Ch. 6, regarding integration in q_z and $q_{\|}$). In the end the following self-energy expression have been obtained, where *plus* or *minus* signs refer to the emission of a phonon-plasmon of frequency $\omega_{\pm}$ (they are distinct possibilities, with different resulting emission probabilities):

$$\Im\Sigma_{\pm ph}(k, E) = \frac{K_\epsilon \omega_{LO}^2 \alpha_0 m \hbar\, c}{\epsilon_s} I_0 \int_0^{2\pi} d\theta \times F_{\pm} \frac{[1 + n_B - n_F]\ \Theta(E_{\mathrm{res}})}{\sqrt{2mE_{\mathrm{res}} - \hbar^2 k^2 \sin^2(\theta)}}, \tag{7.5}$$

where for brevity it has been set $E_{\mathrm{res}} = E - \mu\,\Theta(\mu) - \hbar\omega_{\pm}$ and

$$\omega_{\pm}^2 = \frac{\omega_q^2 + \omega_{LO}^2}{2} \pm \frac{1}{2}\sqrt{\left(\omega_q^2 - \omega_{LO}^2\right)^2 + 4K_\epsilon \omega_{LO}^2 \Omega_{pl}^2}. \tag{7.6}$$

Here $k = |\overrightarrow{k_{\|}}|$, θ is the angle between $\overrightarrow{q_{\|}}$, μ is the Fermi energy, and m is the electron effective mass. Θ is the Heaviside step function, and I_0 is a form factor ensuing from the wavefunctions overlap integral. $F_{\pm}$ originates from the analytic $\boldsymbol{q}$-integration and its expression for arbitrary values of carrier density is

$$F_{\pm} = \frac{\left(\omega_{\pm}^2 + \Omega_{pl}^2 - \omega_q^2\right)^2}{2\omega_{\pm}\left(\omega_{\pm}^2 - \omega_q^2\right)\left(\omega_{\pm}^2 - \omega_{\mp}^2\right)}. \tag{7.7}$$

It must be noticed that, after the q-integration yielding Eq. 7.5, all q-dependent quantities in $F_{\pm}$ and E_{res} (that is $\omega_{\pm}$, ω_q and Ω_{pl}) were evaluated in $q_{\mathrm{pole}} = k\cos\theta + \sqrt{2mE_{\mathrm{res}}/\hbar^2 - k^2\sin^2\theta}$, the q-pole of the propagator G, that explicitly depends on k and θ.

For zero carrier density the limiting forms of $F_{\pm}$ are $F_{-} = 1/(2\omega_{LO})$ and $F_{+} = 0$, and recover known expressions, as in P. Sotirelis and K. Hess [110]. An extended description of their quite complicated analytic derivation for arbitrary values of carrier density can be found in [5].

Regarding k, a thermal statistical average over initial $\overrightarrow{k_{\|}}$ states was considered, calculating $\Im\Sigma_{\pm ph}(k, E)$ for $k = \sqrt{2mk_BT}/\hbar$. The Fermi energy μ was calculated making use of its analytic 2D expression [116], in which the 2D equivalent carrier density in the QW enters (here estimated as $N_{\mathrm{QW}}L_{\mathrm{w}}$, where N_{QW} is the 3D density for the QW states and L_{w} is the QW width). Considering that capture can take place through emission of any of the two possible coupled phonon-plasmon modes of frequency ω_{+}, ω_{-}, with corresponding times τ_{+}, τ_{-}, the total capture time $\tau_{c,ph}$ by phonon emission follows as:

$$\tau_{\pm} = \frac{2}{\hbar}\Im\Sigma_{\pm ph}\,, \qquad \tau_{c,ph} = \left(\tau_{+}^{-1} + \tau_{-}^{-1}\right)^{-1}. \tag{7.8}$$

The barrier density N_{barr} enters in all density-dependent terms ($F_{\pm}$, $\omega_{\pm}$), except for the expression of Fermi energy μ and consequently n_F, that depend on 2D QW carrier density. The QW Fermi occupation factor $1 + n_B - n_F$, normally inserted as an external term in semiclassical LEDs or lasers rate equations, is already included in the expression of $1/\tau_{c,ph}$.

7.3 Model validation

In order to validate the application of the present model to InGaN/GaN QWs LEDs, we considered the experiment described by W. H. Fan *et al.* [137]. The complexities of carrier transport in electrically injected InGaN-based devices were avoided in [137] by using an optical pump-probe experiment on a nominally undoped 2.5 nm/7.5 nm $In_{0.08}Ga_{0.92}N$/GaN MQW, grown on sapphire substrate and capped with a 20 nm-thick GaN layer. Capture times were deduced from the rise time of the QW differential transmission (DT) signal following photoexcitation of carriers in the barriers by femtosecond pulses. Starting from the measured DT signal, the pulse width and the hot-carrier cooling time were deconvolved through a rate equation analysis, eventually obtaining the elementary QW capture time τ_c, separating this process from all other relaxation and recombination mechanisms. The authors of [137] remark that the DT rise time was expected to be determined primarily by the electron dynamics, since the hole capture time in nitride QWs is three orders of magnitude shorter than that of the electrons [153].

In the present calculations, direct capture processes from barrier bulk states to the ground state (GS) of a single QW with flat bands have been considered, neglecting possible transitions mediated by intermediate bound or quasi-bound states. Electron eigenfunction profiles have been evaluated in the effective mass approximation, obtaining strain-dependent energy gap and band offset according to [158]. Band bending and piezoelectric charge effects would surely led to more realistic modeling and could be considered

Table 7.1. Radiative and Auger recombination coefficients estimated in the literature for InGaN/GaN LEDs. When known, also the corresponding peak wavelength λ_{peak} is reported.

	B, cm$^3/s$	C, cm$^6/s$	λ_{peak}, nm
Ref. [159]	10^{-11}	10^{-30}	450
Ref. [160]	$2 \cdot 10^{-11}$	$1.5 \cdot 10^{-30}$	407
Ref. [161]	10^{-10}	$8 \cdot 10^{-29}$	460
Ref. [162]	$1.2 \cdot 10^{-12}$	$3.5 \cdot 10^{-31}$	green
Ref. [163]	$7 \cdot 10^{-11}$	10^{-29}	430
Ref. [164]	$3 \cdot 10^{-12}$	$4.5 \cdot 10^{-31}$	violet
Ref. [165]	$3 \div 6 \cdot 10^{-12}$	$0.8 \div 1 \cdot 10^{-31}$	$440 - 470$
Ref. [142]	–	$2.7 \div 5.7 \cdot 10^{-30}$	470

without substantial modifications to the model. Capture time, when polarization charges warp eigenfunctions, could be numerically different, but the overall behavior with density is maintained and the present model offers a satisfactory level of accuracy, as differences would be limited to the form factor I_0 in Eq. 7.5. The evolution of barrier and QW populations were described by a system of rate equations in which capture, escape and recombination processes were included by means of characteristic time scales τ_c, τ_{esc}, τ_{rec}:

$$\frac{dN_{\text{barr}}}{dt} = -\frac{N_{\text{barr}}}{\tau_c} - \frac{N_{\text{barr}}}{\tau_{\text{rec}}} + \frac{N_{\text{QW}}}{\tau_{\text{esc}}} \tag{7.9a}$$

$$\frac{dN_{\text{QW}}}{dt} = -\frac{N_{\text{QW}}}{\tau_{\text{rec}}} - \frac{N_{\text{QW}}}{\tau_{\text{esc}}} + \frac{N_{\text{barr}}}{\tau_c}. \tag{7.9b}$$

In the following, it has been assumed $\tau_c \approx \tau_{c,ph}$ [137, 144]. The recombination time τ_{rec} was evaluated considering only radiative and Auger recombination processes, whose coefficients B and C are still affected by significant uncertainties (see Table 7.1). A discussion about the limitations of standard ABC recombination models in nitrides (where A is the Shockley-Read-Hall recombination coefficient, not considered in the present study as well as leakage effects) can be found in [142] and references therein. In calculations it has been chosen $B = 2 \times 10^{-11}\,\text{cm}^3\text{s}^{-1}$ and $C = 3 \times 10^{-30}\,\text{cm}^6\text{s}^{-1}$ in QW and barrier, leaving as fitting parameter only the escape time τ_{esc}. A different choice for B and C would lead to a different τ_{esc}, but would not change the gist of the present investigation.

Possible intermediate scattering events involving the QW excited states ES_n could be formally included by writing, instead of Eqs. 7.9, a multi-level rate equations system considering a barrier state, the GS and all the ES_n states. However, employing Eq. 7.2

in a multi-level version of Eqs. 7.9 would result in a complicated combination of the unscreened potentials and polarization functions: in order to model the screening, it would not be possible to define a dielectric function such as the presently employed $\epsilon(\boldsymbol{q}, \omega_m)$ (this difficulty is described, but not addressed, in [136] for GaAs-based materials). An inclusion of the excited states would be feasible in the static long-wavelength limit of SPP screening [136, 166], but such a description would be inadequate to the goals of the present investigation, focused on the inclusion of dynamical plasma screening effects on capture, although accepting some approximations in view to obtain fast computer codes suitable for LED simulation.

The solution of Eqs. (7.9) in stationary conditions provides self-consistently $\tau_{c,ph}$, τ_+, τ_- and N_{QW} as a function of N_{barr}. Following an optimization process, the best agreement with experimental capture times from [137] was obtained for $\tau_{\text{esc}} \approx 0.3\,\text{ps}$. With this estimate, Fig. 7.1 reports separately the calculated τ_+ and τ_- over a wide range of N_{barr}, along with the experimental data.

It would not be possible to reproduce the measured capture times considering τ_+ or τ_- alone: in fact, τ_- reproduces well the experimental data only for the lowest value of N_{barr}. Conversely, in Fig. 7.2 the calculated total $\tau_{c,ph}$ is shown to agree well with the experimental data, confirming the need to consider both phonon-plasmon modes in the capture description.

The same figure reports $\tau_{c,ph}$ for the same QW composition, but for $L_{\text{w}} = 1\,\text{nm}$ and $L_{\text{w}} = 5\,\text{nm}$, corresponding – only at the lowest values of carrier density – to shorter/longer capture times due to the decreased/increased GS confining potential, respectively. Increasing the carrier density, the trend appears inverted, due to interplay of τ_+ and τ_-. The increase of τ_- for the highest values of carrier density in Fig. 7.1 is due to Fermi band-filling in the QW.

The impact of a density-dependent $\tau_{c,ph}$ becomes evident observing how the QW GS populates for increasing N_{barr}: Fig. 7.3 reports N_{QW} as a function of N_{barr}, calculated with the present model (Eqs. 7.5, 7.8 and 7.9) and with a fixed, "typical" capture time $\tau_{c,0} = 0.3\,\text{ps}$ for the QW considered in [137].

It is evident that the density-dependent model for $\tau_{c,ph}$ tends to favor the progressive filling of the QW GS. In order to explain this behavior, we must consider that for the highest N_{barr} values, the capture time is mainly determined by the shortest one, τ_+, that progressively reduces much for increasing N_{barr}. This is due to the fact that also ω_+ decreases for increasing N_{barr}, and when it approaches the resonance condition $\omega_+ \approx E - \mu$, the capture time tends virtually to zero (see also Fig. 6.10 in Chapter 6). Therefore, when carrier density increases, we correctly observe an increasingly larger capture efficiency, well visible in Fig. 7.3. The effect of QW GS Fermi band filling reflects in a progressive roll-off of N_{QW} for the highest values of N_{barr}. Nevertheless, in this case other processes like carrier-carrier scattering intervene, and a description based just on relaxation through phonon emission could be not realistic.

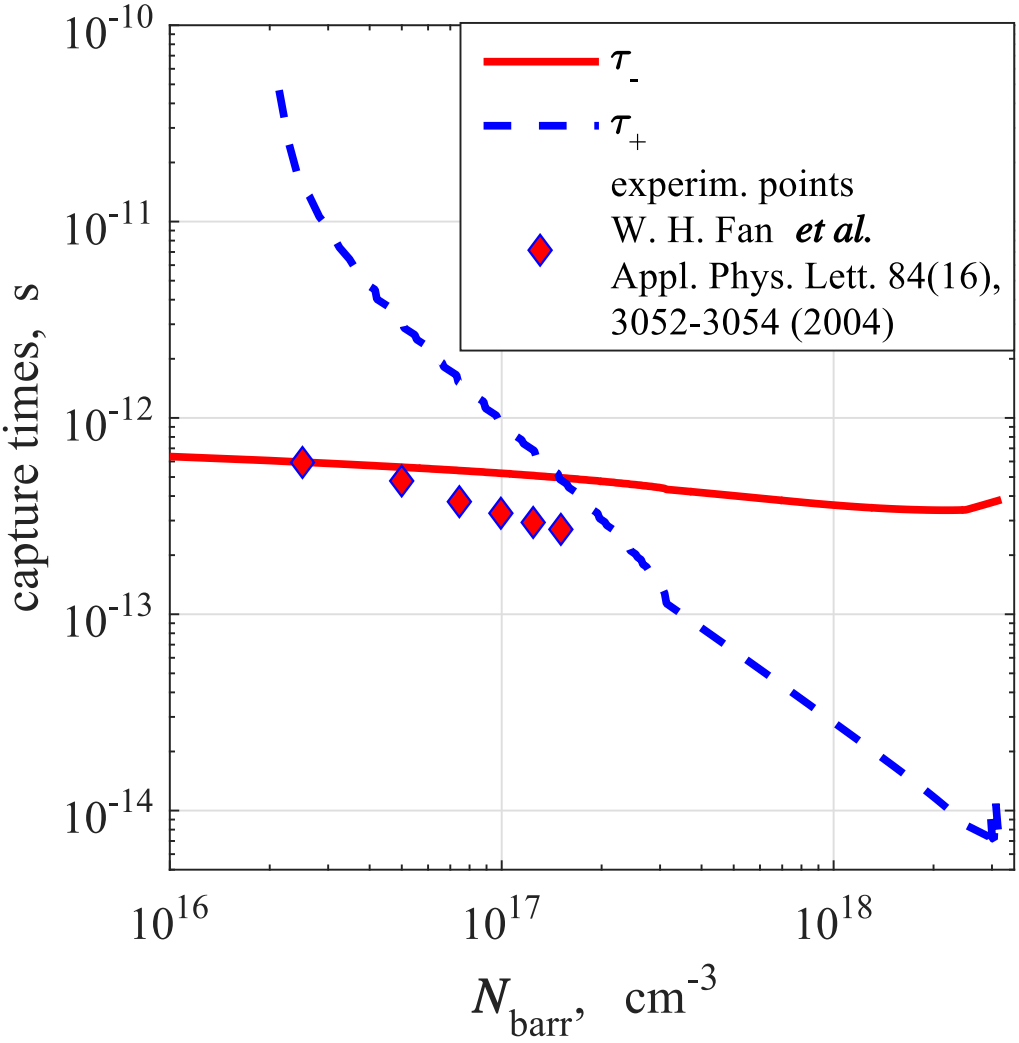

Figure 7.1. Electron capture times $\tau_{\pm}$ calculated for the QW considered in [137]. Experimental points are shown as symbols.

7.4 Final remarks

A quantum model allowing to semi-analytically calculate $\tau_{c,ph}$ as a function of N_{barr} has been successfully validated, abandoning the fixed-capture-time approximation customarily adopted in device simulations. Novel features of our approach are the use of an energy-, $\boldsymbol{q}$- and density-dependent dielectric function in the Frölich Hamiltonian (dynamical screened potential) for InGaN/GaN materials, and the $\boldsymbol{q}$-integration performed *exactly*

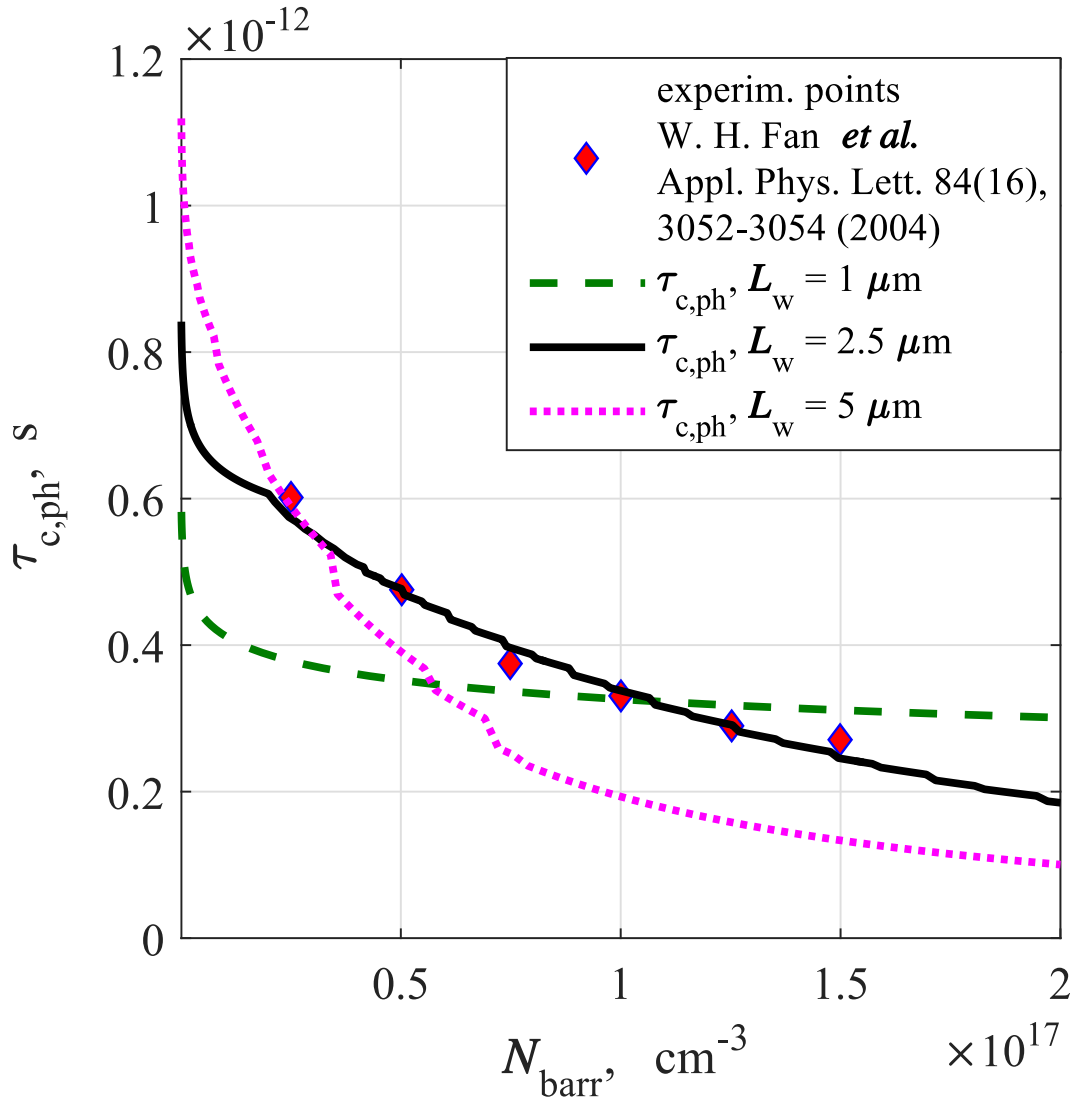

Figure 7.2. Electron capture time $\tau_{c,ph}$ calculated for the QW considered in [137] (black solide line) and for two different values of L_{w}. Experimental points are shown as symbols.

(that is, without $\boldsymbol{q}$-truncation), exploiting residue theorems. This has produced a closed-form expression for $\tau_{c,ph}$, suitable for device modeling tools. The present work is only an example of a methodology applicable to other phonon modes and carrier-carrier scattering mechanisms. Furthermore, it can be easily extended to more realistic QW descriptions without substantial modifications.

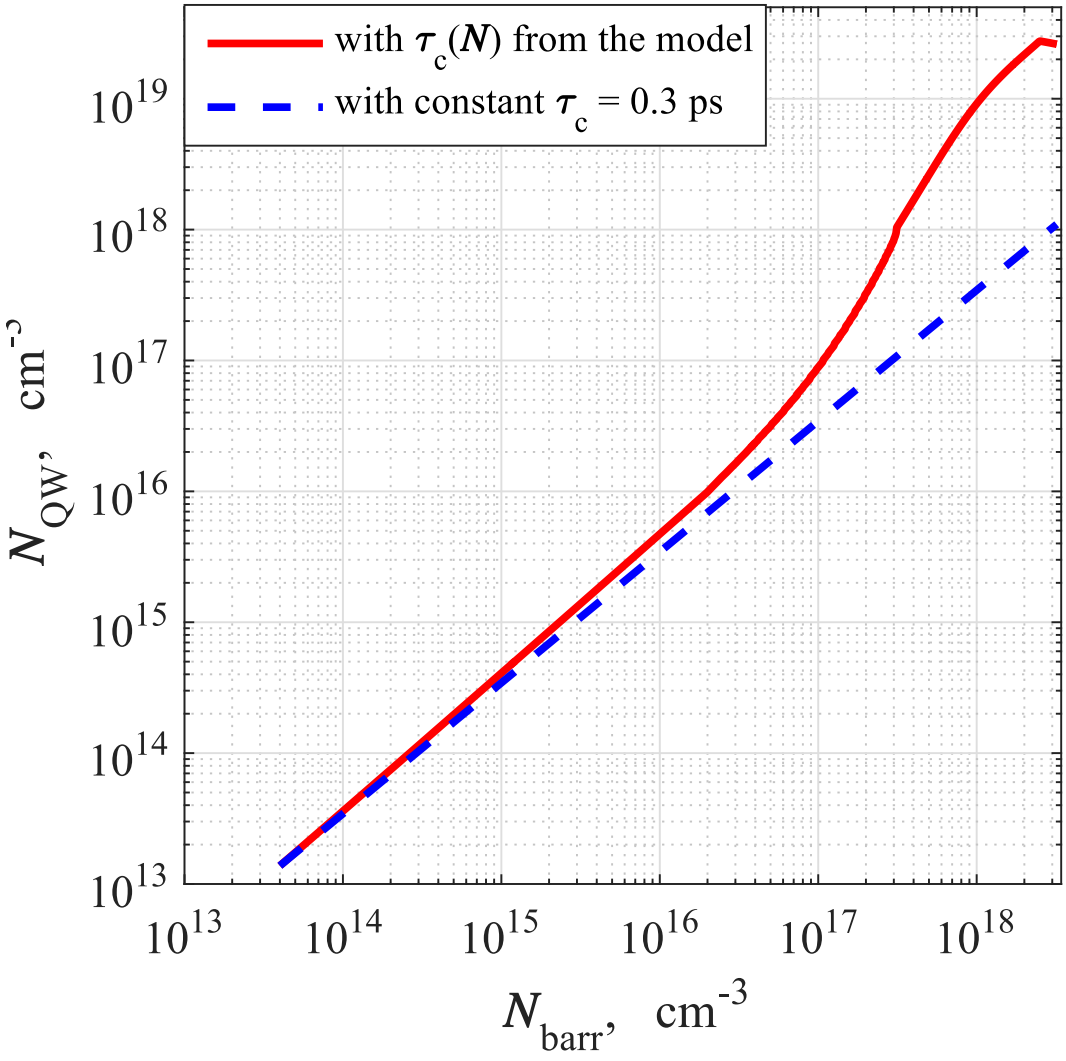

Figure 7.3. Carrier density N_{QW} versus N_{barr} for the QW considered in [137], calculated with the density-dependent capture time $\tau_{c,ph}$ (red solid line) and with a fixed capture time $\tau_{c,0} = 0.3$ ps (blue dashed line).

Appendices

Appendix A

Details about the implementation of the HgCdTe software library in TCAD Sentaurus

The interface to *Sentaurus* consists in a set of ASCII files `HgCdTe_Tnn.par`, one for each desired value of temperature T (for example, for $T = 80$ K, the file is `HgCdTe_Tnn.par`), written in a proprietary format. The `HgCdTe_Tnn.par` file starts with a declaration of the material:

```
  Material = "HgCdTe" {
 ...
}
```

Inside this declaration, all parameters are defined as they were routines. Each material property must appear in a particular format. The easiest case is represented by properties that are not molar-fraction dependent, but just a number. In this case Sentaurus needs just to load its correct value.

Molar-fraction (x) dependent properties requires a more challenging work: first, it is needed to obtain a **cubic fitting** of the expression $f(x)$ of interest (it can represent the energy-gap, a dielectric function, a mobility, etc.) with $P_3(x)$, for x in $[0, 1]$, having fixed the required temperature (x is the Cd mole fraction of Hg$1 - x$Cd$_x$Te). Then, the four obtained coefficients must be written in the format expected by Sentaurus and described in each of the sections that follow.

A.1 Energy gap and electron affinity

For the energy gap and electron affinity, we have a routine where the considered interval of x and the cubic fitting coefficients are stored as described here (please notice that energy gap and electron affinity must be loaded together by the *Sentaurus* interface):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array $Xmax$;
- $Eg0$ = 2-elements array $[Eg(x_{min}), Eg(x_{max})]$;
- in $B(Eg0(1))$ and $C(Eg0(1))$ are stored the quadratic and cubic coefficient of the fit;
- the parameters α and β are unused.

The results, for a given value of T, is:

```
* Energy gap  and affinity
*---------------------------------
*
* Eg(x) = F + A x + B x^2 + C x^3
* Chi(x) = F + A x + B x^2 + C x^3
*
* 0 <= x <= 1
*---------------------------------
 Bandgap{
      Xmax(0)     = 0
      Chi0(0)     = 4.5231
      Eg0(0)      = -0.27749
      alpha(0)    = 0
      beta(0)     = 0
      Xmax(1)     = 1
      Chi0(1)     = 2.976
      Eg0(1)      = 1.6255
      alpha(1)    = 0
      beta(1)     = 0
      B(Chi0(1)) = 0.65853
      C(Chi0(1)) = -0.67642
      B(Eg0(1))   = -0.81
      C(Eg0(1))   = 0.832
  }
}
```

Similar procedures are built for all materials and will be reviewed briefly for each of them. For them, the syntax is similar.

We developed a set of Matlab scripts that, starting from a given analytic formula, automatically builds the file `HgCdTe_Tnn.par` according to this format.

A.2 Effective masses

Interface to *Sentaurus*: polynomial fitting with $P_3(x)$, for $x \in [0,1]$, dividing the interval in 10 sub–intervals, having fixed the required temperature to the desired value. The fine subdivision of the composition interval is done in order to better approximate the $x-$dependence.

The file ".par" section relative to effective masses consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array $Xmax$;
- mm = 2-elements array $[m_e(x_{min}), m_e(x_{max})]$;
- in $B(mm(1))$ and $C(mm(1))$ are stored the quadratic and cubic coefficient of the fit.

```
*-----------------------------------
* electron effective mass
*-----------------------------------
*
* mE(x) = F + A x + B x^2 + C x^3
*
*-----------------------------------
  eDOSMass{
      Formula= 1
      a= 0
      ml= 0
      *-----------------------------------
      * 0 <= x <= 0.1
      *-----------------------------------
      Xmax(0) = 0
      mm(0) = -0.026671
      Xmax(1) = 0.1
      mm(1) = -0.0080161
      B(mm(1)) = -0.63616
      C(mm(1)) = 1.585
      *-----------------------------------
      * 0.1 <= x <= 0.2
      *-----------------------------------
     ...
     ...    and so on, for each interval of x,
            as produced by the Matlab routine.
}
```

```
*-----------------------------------
* Hole effective mass
*-----------------------------------

  hDOSMass{
       Formula= 1
       a= 0
       b= 0
       c= 0
       d= 0
       e= 0
       f= 1
       g= 1
       h= 1
       i= 1
       mm= 0.55      *  for holes, a constant value is considered
  }
```

A.3 Electron and hole mobility

Interface to *Sentaurus*: polynomial fitting with $P_3(x)$, for $x \in [0,1]$, dividing the interval in 51 sub–intervals, having fixed the required temperature to the desired value. The fine subdivision of the composition interval is done in order to better approximate the x–dependence.

The file ".par" section relative to Mobility consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array named $Xmax$, where each of its elements is a couple of values (e,h), the first one referring to electrons, the last to heavy holes ;
- $mumax$ = 2-elements array of couples (e,h) $[\mu_{e,h}(x_{min}), \mu_{e,h}(x_{max})]$;
- in $B(mumax(1))$ and $C(mumax(1))$ are stored the quadratic and cubic coefficient of the fit, still ordered as couples (e,h).

```
*-----------------------------------
* Mobility
*-----------------------------------
*
* mu(x) = F + A x + B x^2 + C x^3
*
*-----------------------------------
```

```
  ConstantMobility{
      *-----------------------------------
      * 0 <= x <= 0.02
      *-----------------------------------
      Xmax(0) = 0,0
      mumax(0) = 427264.6091 , 4272.6461
      exponent(0) = 0 , 0
      Xmax(1) = 0.02 , 0.02
      mumax(1) = 561078.6757 , 5610.7868
      exponent(1) = 0 , 0
      B(mumax(1)) = 38942108.7991 , 389421.088
      C(mumax(1)) = 33879166.6127 , 338791.6661
      *-----------------------------------

     ...    and so on, for each interval of x,
            as produced by the Matlab routine.
```

A.4 Low– and high–frequency dielectric constant

Interface to *Sentaurus*: polynomial fitting of the given formula with $P_3(x)$, for $x \in [0, 1]$, having fixed the required temperature (see the motivations given for E_g).

The file ".par" section relative to low-frequency dielectric constant ϵ_0 consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array $Xmax$;
- $epsilon$ = 2-elements array $[\epsilon_0(x_{min}), \epsilon_0(x_{max})]$;
- in $B(epsilon(1))$ and $C(epsilon(1))$ are stored the quadratic and cubic coefficient of the fit.

```
\
*-----------------------------------
* Dielectric Constant - Low Freq.
*-----------------------------------
*
* ep(x) = F + A x + B x^2 + C x^3
*
* 0 <= x <= 1
*-----------------------------------
  Epsilon{
      Xmax(0)     = 0
```

```
      epsilon(0) = 20.5
      Xmax(1)    = 1
      epsilon(1) = 10.7
      B(epsilon(1)) = 5.7
      C(epsilon(1)) = 1.0162e-013
  }
```

The high-frequency dielectric constant is *not used* in the simulations and is not modified from the default value.

A.5 Intrinsic carrier density

Interface to *Sentaurus*: this parameter is not directly passed to *Sentaurus*, entering only indirectly, through the Auger coefficient calculation (see the relative Section).

A.6 Electron and hole Auger recombination coefficients

Interface to *Sentaurus*: polynomial fitting with $P_3(x)$, for $x \in [0.18, 1]$, dividing the interval in 51 sub–intervals, having fixed the required temperature to the desired value. The fine subdivision of the composition interval is done in order to better approximate the x–dependence.

The file ".par" section relative to Auger recombination coefficients C_n, C_p consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array named $Xmax$, where each of its elements is a couple of values (e, h), the first one referring to electrons, the last to heavy holes ;
- A, B and C are 2-elements array of couples (e, h), that is referring to electrons and holes, for the linear, quadratic and cubic coefficient of the polynomial fit:

```
*-----------------------------------
* Auger Coefficients.
*-----------------------------------
*
* C(x) =  F + A x + B x^2 + C x^3
*
*-----------------------------------
  Auger{
      *---------------------------------
      * 0 <= x <= 0.18
```

```
*----------------------------------
Xmax(0) = 0 #, 0
A(0) = 1.5793e-023 , 3.2098e-024
B (0) = 0 , 0
C (0) = 0 , 0
H (0) = 0 , 0
N0(0) = 1 , 1
Xmax(1) = 0.18 #, 0.18
A(1) = 1.5793e-023 , 3.2098e-024
B (1) = 0 , 0
C (1) = 0 , 0
H (1) = 0 , 0
N0(1) = 1 , 1
*----------------------------------
* 0.18 <= x <= 0.1884
*----------------------------------
Xmax(2) = 0.1884 #, 0.1884
A(2) = 8.7006e-024 , 1.7607e-024
B(A(2)) = 1.3033e-018 , 2.7045e-019
C(A(2)) = -2.2753e-018 , -4.7232e-019
B (2) = 0 , 0
C (2) = 0 , 0
H (2) = 0 , 0
N0(2) = 1 , 1
*----------------------------------
* 0.1884 <= x <= 0.1968
*----------------------------------

...     and so on, for each interval of x,
        as produced by the Matlab routine.
```

A.7 Radiative recombination rate

Interface to *Sentaurus*: polynomial fitting of the given formula with $P_3(x)$, for $x \in [0, 1]$, having fixed the required temperature (see the motivations given for E_g).

The file ".par" section relative to Radiative recombination coefficient B_{rad} consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

- first and last element of x, that is x_{min} and x_{max} are stored as a 2-elements array named $Xmax$, where each of its elements is a couple of values (e, h), the first one referring to electrons, the last to heavy holes ;
- A, B and C are 2-elements array of couples (e, h), that is referring to electrons and

holes, for the linear, quadratic and cubic coefficient of the polynomial fit. They are stored in the function specification " RadiativeRecombination" as specified in the lines here below:

```
*-----------------------------------
* Radiative Coefficient.
*-----------------------------------
*
* Gr(x) =  F + A x + B x^2 + C x^3
*
*-----------------------------------
  RadiativeRecombination{
     Xmax(0) = 0
     C(0) = 4.7603e-011 * F is stored here
     Xmax(1) = 0.18
     C(1) = 4.7603e-011 * F is stored here
     Xmax(2) = 1
     C(2) = 1.4527e-009 * A(Xmax(1)) is stored here
     B(C(2)) = -1.1766e-009 * B is stored here
     C(C(2)) = 9.0669e-010 * C is stored here
  }
```

A.8 Shockley–Read–Hall (SRH) recombination rate

Interface to *Sentaurus*: cited parameters (γ, τ_{min}, τ_{max} and the density N_{ref}) are passed to the simulator as (e, h) couples of constants.

The file ".par" section relative to SRH recombination coefficient G_R consists of the following lines, inside the declaration "Material" (see Sec. 2.2 for more details):

```
*-----------------------------------
* SRH Coefficients.
*-----------------------------------
*
* Constant values: electrons, holes
* Note: taumax and taumin must be equal
*
*-----------------------------------
  Scharfetter{
      taumin =  1e-6, 1e-6
      taumax =  1e-6, 1e-6
      Nref   =  1e+16 , 1e+16
      gamma  =  0 , 0
```

```
      Talpha =  0 , 0
      Tcoeff =  0 , 0
      Etrap  =  0
  }
```

A.9 Absorption coefficient and refractive index (real part)

Interface to *Sentaurus*: the formula has been implemented as a C++ routine, linked at runtime, because the Sentaurus syntax employed for all other parameters is not allowed for the absorption coefficient. The parameter file ".par" only choses this form among other possible choices with the statement:

```
Absorption { Formula = 3 HgCdTe }
```

In the command file of “SDEVICE”, the library is linked with the statement.

```
Physics {
...
...

  Optics (
                ComplexRefractiveIndex (
                    WavelengthDep(Real Imag)
                    TemperatureDep
                    CRIModel (Name = "CRIMI_HgCdTe_nk_Nov2015")
                )
}
```

where `CRIMI_HgCdTe_nk_Nov2015.C` is the source file of the compiled routine. In the followings, we report the C++ routine itself. For a given temperature and Cd mode fraction of the HgCdTe alloy at a given point of the mesh, it calculates the real n and imaginary part κ of the refractive index at the wavelength λ. The absorption coefficient α is then given by $\alpha = 4\pi\kappa/\lambda$. This way, if a position-dependent alloy composition is given, the optical properties are evaluated consequently, just linking at run-time this routine.

```
/*
File:          CRIMI_HgCdTe_nk_Nov2015.C

Author:      Benjamin Pinkie (Boston University), Marco Vallone

Function:    Implementation of Complex Refractive Index Model
          Interface for Hg[1-x]Cd[x]Te under variable temperature
          and Cadmium concentration. Generates shared object
```

```
    file which is imported to Synopsys sdevice simulator at
    runtime. Function depends on the specification of a
    'Formula = [1, 2, 3, or 4]' in the _des.cmd file which controls
    whether n [1], k [2], or both [3, default] are calculated.
    Formula      4 uses a more recent absorbtion coefficient model
    (Littler, 2005) for the calculation of the extinction coefficient
    (temperature dependant sigma).

References: C.A. Hougen, "Model for infrared absorption and
         transmission of liquid-phase
    epitaxy HgCdTe," J. Appl. Phys. 66 (8), 15 October 1989
    C.L. Littler, "Temperature, thickness, and interfacial composition
    effects on the absorbtion properties of (Hg,Cd)Te epilayers
    grown by liquid-phase epitaxy on CdZnTe,"
    J. Elec. Mat. 34 (6), 2005

Changelog:  May 25, 2012   -- Original Creation
May 29, 2012   -- Added support for computing only
n, k, or both.
May 30, 2012   -- Added Formula #4 support for Littler
absorbtion model
June 1, 2012   -- Created new functional header and fixed
several errors
     Nov 26, 2015   -- Corrected some minor bugs and introduced
     a new n (real part) definition (P. Capper, ``Mercury Cadmium
     telluride. Growrt, properties and applications'', Wyley, 2011,
     Sec. 9.6) making n dependent also on T.



Compile:    'cmi CRIMI_HgCdTe_nk_Nov2015.C'
*/

#include "CRIModels_DoNotEdit.h"
#include <iostream>
#include <stdlib.h>
#include <math.h>
#include <string.h>
using namespace std;


namespace opto_n_cri {
const double SpeedOfLight = 299792458; // [m/s]
```

```
const double BoltzmannConstant = 6.626176e-34; // [Js]
const double PlanckConstant = 1.380662e-23; // [J/K]
const double ElectronCharge = 1.6021892e-19; // [C]
const double kq = PlanckConstant / ElectronCharge; // [eV/K]
const double Pi = 3.141592654; // [ ]

  class Constant_CRI_Model : public CRI_Model {
private:
string myName;
int formula;
public:
Constant_CRI_Model(const CRI_Environment& env);
~Constant_CRI_Model();
void Compute_n(double& n);
void Compute_k(double& k);
  };

  Constant_CRI_Model::Constant_CRI_Model(const CRI_Environment&
                      env) : CRI_Model(env) {
myName = Name();
formula = InitParameter("Formula",3);

RegisterVariableToRead("xMoleFraction");
  }

  Constant_CRI_Model::~Constant_CRI_Model() {
#ifdef DEBUG
cout << "CRI_Model: delete\n";
#endif
  }

  void Constant_CRI_Model::Compute_n(double& n) {

                double lambda;
                if( strcmp(ReadMaterialName().c_str(),"HgCdTe")==0 )
                {

if (formula == 2) {
n = Read_n();
return;
}
double x = ReadVariableValue("xMoleFraction");
double lambda_read = ReadWavelength();
```

```
double T = ReadTemperature();
double A = 13.173 - 9.852*x + 2.909*x*x + 1.0e-3*(300.0 - T);
double B = 0.83- 0.246*x - 0.0961*x*x + 8.0e-4*(300.0 - T);
double C = 6.706 - 14.437*x + 8.531*x*x + 7.0e-4*(300 - T);
double D = 1.953e-4 - 0.00128*x + 1.853e-4*x*x;

      double eg1 = -0.302 + 1.93 * x  - (0.81 * x * x) + (0.832 *
                          x * x * x);
      double eg2 = 5.35e-4 * ((T * T * T -1822) / (T * T - 255.2)) *
                        (1 - 2 * x);
      double eg = (eg1+eg2);

      if (lambda_read < (1.24 / eg)) {
         lambda = 1.24 / eg; }
      else {
    lambda = lambda_read;
}


double n2 = A + B/(1-(C/lambda)*(C/lambda)) + D*lambda*lambda;
if (n2 < 1) {
cout << "Error: n2 < 1 therefore setting n = 1\n"; // unrealistic
n = 1;
} else {
n = sqrt(n2);
}
 } // end HgCdTe

 else if( strcmp(ReadMaterialName().c_str(),"ZnS")==0 ){
        double lambda_read = ReadWavelength();
        n = 1.6e-07 * pow(lambda_read,6) - 1.2e-05 *
               pow(lambda_read,5) + 0.00033 *
               pow(lambda_read,4) -
               0.0042 * pow(lambda_read,3) + 0.026 *
               pow(lambda_read,2) - 0.081 * lambda_read + 2.4;
} // end ZnS

else if (strcmp(ReadMaterialName().c_str(),"Polymeric_coating")==0 )
{
          n = 1.35;
} // end Polymeric coating

             else{
```

```
                n = 1.0;
              } // end air (default)
  }

  void Constant_CRI_Model::Compute_k(double& k) {


     if( strcmp(ReadMaterialName().c_str(),"HgCdTe")==0 )
     {
if (formula ==1)
k = Read_k();
return;
}
double x = ReadVariableValue("xMoleFraction");
double T = ReadTemperature();
double lambda = ReadWavelength();
double PhotonE = BoltzmannConstant*
          SpeedOfLight/(1e-6*lambda)/ElectronCharge; // [eV]
double AT = 100.0+5000.0*x;
double A0 = exp(-18.88+53.61*x);
double S = 0.0; // for debugging and error catching
if (formula == 4) {
S = 3.05e4*(1.0+x)*(1.0 + (4.02e-4 * T));
} else {
S = 3.267e4*(1.0+x); // Original Hougen model
}
double E0 = -0.3432+1.838*x;
double T0 = 81.9;
double LN = log(AT/A0);
double ST = S/(T+T0);
double phE0 = E0 + LN / ST;
double A = 0.0; // for debug and error catching

if (PhotonE < phE0) {
A = A0 * exp((PhotonE-E0)*ST); // [/cm]
} else {
A = AT * sqrt(2*ST)*sqrt(PhotonE-E0-(LN-0.5)/ST); // [/cm]
}

k = lambda*1.0e-6/(4*Pi)*A/1.0e-2;
    } // end HgCdTe
    else {  // ZnS
       k = 0;
```

```
     } // end ZnS

} // end Compute_k
} //end opto

/* Every CRIMI model must implement a virtual constructor
   which allocates an instance of the derived class
   as well as a virtual destructor which deallocates the
   created object. */

extern "C" {

// Virtual constructor
opto_n_cri::CRI_Model*
new_CRI_Model(const opto_n_cri::CRI_Environment& env) {
return new opto_n_cri::Constant_CRI_Model(env);
}

// Virtual destructor
void
delete_CRI_Model(opto_n_cri::CRI_Model* cri_model) {
delete cri_model;
}

}
```

A.10 Band to band tunneling

Interface to *Sentaurus*: the formula described in Sec. 2.2.11 has been implemented as a C++ routine, linked at runtime.

```
/*
File:         BTBT_MM_MV.C

Author:     Marco Mandurrino, Marco Vallone

Function:   Implementation of the band to band tunneling (BTBT).

References: the Kane's original formulation implemented by
M. Mandurrino for his Master Thesis has been revised and adapted
according to the description of Chapter\,\ref{ch:4}.
```

```
*/

#include <stdlib.h>
#include <iostream>
#include <cmath>
#include "PMIModels.h"

class NewBand2Band_MV : public PMI_Recombination {

  double C;

protected:
    double pi, q, m0, hbar, D;

public:
NewBand2Band_MV (const PMI_Environment& env);
~NewBand2Band_MV();

        void Compute_r
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& r) ;      // recombination rate

        void Compute_drdt
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdt) ;    // derivative of recombination rate
                           // with respect to lattice temperature

        void Compute_drdn
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdn) ;     // derivative of recombination rate
                           // with respect to electron density
```

```
        void Compute_drdp
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdp) ;     // derivative of recombination rate
                            // with respect to hole density

        void Compute_drdnie
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdnie) ;   // derivative of recombination rate
                // with respect to effective intrinsic density

        void Compute_drdf
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdf) ;     // derivative of recombination rate

};



NewBand2Band_MV ::
NewBand2Band_MV (const PMI_Environment& env) :
  PMI_Recombination (env)

{ C = InitParameter ("C", 1e-30);
}

NewBand2Band_MV::
~NewBand2Band_MV ()

{
}
```

```
void NewBand2Band_MV::
Compute_r
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& r)       // recombination rate
{ double mt, eg, A, B, eg1, eg2, eg_eV;
  double pi, q, m0, hbar, D;

      const PMIBaseParam* gamma = ReadParameter ("gamma");
      double g = *gamma;

      const PMIBaseParam* Temp = ReadParameter ("Temp");
      double T = *Temp;

      double x = ReadxMoleFraction();

  //    std::cout << "x = " << x << std::endl;
  //    std::cout << "T = " << T << std::endl;

      pi = 3.14159265358979;
      q = 1.602e-19;
      m0 = 9.109e-31;
      hbar = 1.054e-34;
      mt = 0.017 * m0;

      D = 1;

  //    eg = 0.244 * q;

  // calcola energy gap usando x e T


      eg1 = -0.302 + 1.93 * x  - (0.81 * x * x) +
                            (0.832 * x * x * x);
      eg2 = 5.35e-4 * ((T * T * T -1822) / (T * T - 255.2))
                * (1 - 2 * x);
```

```
      eg_eV = (eg1+eg2);
      eg = eg_eV * q;
   //   std::cout << "eg = " << eg_eV << std::endl;




   A = -(1e-2) * sqrt(2 * mt) * pow(q,2) / (4 * pow(pi,3) *
         pow(hbar,2) * sqrt(eg));
   B = (1e-2) * pi * sqrt(mt * pow(eg,3)) / (2 * sqrt(2) * q * hbar);


   //   std::cout << "A = " << A << std::endl;
   //   std::cout << "B = " << B << std::endl;

      if (f == 0) {
        r = 0;
      }
      else  {
        r = A * pow(g,(D-1)) * pow(f,((D + 1) / 2)) *
                    exp(- B / (g * sqrt(f)));
      }
}

void NewBand2Band_MV::
Compute_drdf
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdf)       // derivate of recombination rate

{ double mt, eg, A, B, eg1, eg2, eg_eV;
  double pi, q, m0, hbar;
         double r;

      const PMIBaseParam* gamma = ReadParameter ("gamma");
      double g = *gamma;

      const PMIBaseParam* Temp = ReadParameter ("Temp");
      double T = *Temp;
```

```
    double x = ReadxMoleFraction();

    pi = 3.14159265358979;
    q = 1.602e-19;
    m0 = 9.109e-31;
    hbar = 1.054e-34;
    mt = 0.017 * m0;

    D = 1;

//    eg = 0.244 * q;

// calcola energy gap usando x e T
    eg1 = -0.302 + 1.93 * x  - (0.81 * x * x) +
               (0.832 * x * x * x);
    eg2 = 5.35e-4 * ((T * T * T -1822) / (T * T - 255.2))
             * (1 - 2 * x);
    eg_eV = (eg1+eg2);
    eg = eg_eV * q;
  //  std::cout << "2nd eg = " << eg_eV << std::endl;

    A = -(1e-2) * sqrt(2 * mt) * pow(q,2) /
                (4 * pow(pi,3) * pow(hbar,2) * sqrt(eg));
    B = (1e-2) * pi * sqrt(mt * pow(eg,3)) /
                (2 * sqrt(2) * q * hbar);

    if (f == 0) {
      r = 0;
    }
    else  {
      r = r = A * pow(g,(D-1)) * pow(f,((D + 1) / 2)) *
                   exp(- B / (g * sqrt(f)));
    }

    if (r == 0) {
      drdf = 0;
    }
    else  {
      drdf = A *  pow(g,D) * pow(f,((D * 0.5)-1)) *
               ( (B / (2 * g * g)) +
               (sqrt(f) / (2 * g)) * (D+1) ) *
              exp(- B / (g * sqrt(f))) ;
```

```
      }
}


void NewBand2Band_MV::
         Compute_drdt
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdt)     // derivative of recombination rate
                            // with respect to lattice temperature
        {;}


void NewBand2Band_MV::
         Compute_drdn
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdn)     // derivative of recombination rate
                            // with respect to electron density
        {;}


void NewBand2Band_MV::
         Compute_drdp
       (const double t,        // lattice temperature
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdp)     // derivative of recombination rate
        {;}

                            // with respect to hole density

void NewBand2Band_MV::
         Compute_drdnie
       (const double t,        // lattice temperature
```

```
        const double n,        // electron density
        const double p,        // hole density
        const double nie,      // effective intrinsic density
        const double f,        // absolute value of electric field
        double& drdnie)   // derivative of recombination rate
               // with respect to effective intrinsic density

        {;}




extern "C" {

   PMI_Recombination* new_PMI_Recombination (const PMI_Environment& env)
    { return new NewBand2Band_MV (env);
   }

}
```

Bibliography

[1] Synopsys, Inc., Mountain View, CA, *Sentaurus Device User Guide. Version K-2015.06*, June 2015.

[2] M. Vallone, M. Goano, F. Bertazzi, G. Ghione, R. Wollrab, and J. Ziegler, "Modeling photocurrent spectra of single-color and dual-band HgCdTe photodetectors: Is 3D simulation unavoidable?," *J. Electron. Mater.*, vol. 43, no. 8, pp. 3070–3076, 2014.

[3] M. Vallone, M. Mandurrino, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, and H. Figgemeier, "Numerical modeling of SRH and tunneling mechanisms in high-operating-temperature MWIR HgCdTe photodetectors," *J. Electron. Mater.*, vol. 44, no. 9, pp. 3056–3063, 2015.

[4] M. Vallone, M. Goano, F. Bertazzi, G. Ghione, W. Schirmacher, S. Hanna, and H. Figgemeier, "Comparing FDTD and ray tracing models in the numerical simulation of HgCdTe LWIR photodetectors," *J. Electron. Mater.*, accepted for publication, forthcoming 2016.

[5] M. Vallone, "Quantum well electron scattering rates through longitudinal optic-phonon dynamical screened interaction: An analytic approach," *J. Appl. Phys.*, vol. 114, p. 053704, Aug. 2013.

[6] M. Vallone, F. Bertazzi, M. Goano, and G. Ghione, "Model for carrier capture time through phonon emission in InGaN/GaN quantum wells," *Phys. Status Solidi B*, vol. 252, no. 5, pp. 971–976, 2015.

[7] P. R. Norton, "Infrared detectors in the next millennium," in *Infrared Technology and Applications XXV* (B. F. Andresen and M. Strojnikb, eds.), vol. 3698, Proceedings of the SPIE, pp. 652–665, Apr. 1999.

[8] A. Rogalski, *Infrared detectors*. Boca Raton, FL: CRC Press, 2nd ed., 2011.

[9] D. J. Lovell, "The development of lead salt detectors," *Am. J. Phys.*, vol. 37, no. 5, pp. 467–478, 1969.

[10] P. R. Norton, "Infrared image sensors," *Opt. Eng.*, vol. 30, no. 11, pp. 1649–1663, 1991.

[11] A. Rogalski, "Infrared detectors: Status and trend," *Progress Quantum Electron.*, vol. 27, pp. 59–210, 2003.

[12] W. Lawson, S. Nielsen, E. Putley, and A. Young, "Preparation and properties of

HgTe and mixed crystals of HgTe-CdTe," *J. Phys. Chem. Solids*, vol. 9, no. 3, pp. 325–329, 1959.

[13] J. Ziegler, D. Eich, M. Mahlein, T. Schallenberg, R. Scheibner, J. Wendler, J. Wenisch, R. Wollrab, V. Daumer, R. Rehm, F. Rutz, and M. Walther, "The development of 3^{rd} gen IR detectors at AIM," in *Infrared Technology and Applications XXXVII* (B. F. Andresen, G. F. Fulop, and P. R. Norton, eds.), vol. 8012, Proceedings of the SPIE, p. 801237, Apr. 2011.

[14] A. Rogalski, J. Antoszewski, and L. Faraone, "Third-generation infrared photodetector arrays," *J. Appl. Phys.*, vol. 105, no. 9, p. 091101, 2009.

[15] G. Ghione, *Semiconductor Devices for High-Speed Optoelectronics*. Cambridge, U.K.: Cambridge University Press, 2009.

[16] M. A. Kinch, *State-of-the-Art Infrared Detector Technology*. Bellingham, WA: SPIE, 2014.

[17] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, 1992.

[18] M. A. Kinch, *Fundamentals of Infrared Detector Materials*. Bellingham, WA: SPIE, 2007.

[19] K. Jóźwikowski, M. Kopytko, and A. Rogalski, "Numerical estimations of carrier generation-recombination processes and the photon recycling effect in HgCdTe heterostructure photodiodes," *J. Electron. Mater.*, vol. 41, pp. 2766–2774, 2012.

[20] M. A. Kinch, "Fundamental physics of infrared detector materials," *J. Electron. Mater.*, vol. 29, no. 6, pp. 809–817, 2000.

[21] T. Ashley and C. T. Elliott, "Model for minority carrier lifetimes in doped HgCdTe," *Electron. Lett.*, vol. 21, no. 10, pp. 451–452, 1985.

[22] H. F. Schaake, M. A. Kinch, D. Chandra, P. K. Liao, D. F. Weirauch, C.-F. Wan, and H. D. Shih, "High operating temperature MWIR detectors," in *Quantum Sensing and Nanophotonic Devices VII*, vol. 7608, Proceedings of the SPIE, 2010.

[23] A. M. Itsuno, J. D. Phillips, and S. Velicu, "Mid-wave infrared HgCdTe nBn photodetector," *Appl. Phys. Lett.*, vol. 100, p. 161102, Apr. 2012.

[24] A. M. White, "Infrared detectors," in *U.S. Patent 4,679,063*, 1983.

[25] P. Martyniuk, J. Antoszewski, M. Martyniuk, L. Faraone, and A. Rogalski, "New concepts in infrared photodetector designs," *Apr.*, vol. 1, no. 4, p. 041102, 2014.

[26] A. F. Franz, G. A. Franz, S. Selberherr, C. Ringhofer, and P. Markowich, "Finite boxes - a generalization of the finite-difference: method suitable for semiconductor device simulation," *IEEE Trans. Electron Devices*, vol. 30, pp. 1070–1082, Sept. 1983.

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

[28] M. A. Kinch, "The future of infrared; III-Vs or HgCdTe?," *J. Electron. Mater.*, vol. 44, no. 9, pp. 2969–2976, 2015.

[29] D. G. Seiler, J. R. Lowney, C. L. Litter, and M. R. LoLoee, "Temperature and composition dependence of the energy gap of $Hg_{1-x}Cd_xTe$ by two-photon magnetoabsorption technique," *J. Vac. Sci. Technol. A*, vol. 8, no. 2, pp. 1237–1244, 1990.

[30] G. L. Hansen, J. L. Schmit, and T. N. Casselman, "Energy gap versus alloy composition and temperature in $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 53, pp. 7099–7100, Oct. 1982.

[31] P. Capper, ed., *Properties of Narrow Gap Cadmium-based Compounds*, vol. 10 of *EMIS Datareviews Series*. London: INSPEC, 1994.

[32] R. Legros and R. Triboulet, "Photoluminescence of Cd–rich $Hg_{1-x}Cd_xTe$ alloys (0.7<x<1)," *J. Cryst. Growth*, vol. 72, no. 1–2, pp. 264–269, 1985.

[33] F. M. Tong, H. Yuan, X. Yang, and N. M. Ravindra, "HgCdTe photovoltaic detectors and some related aspects," in *Infrared Detectors and Focal Plane Arrays II* (E. L. Dereniak and R. E. Sampson, eds.), vol. 1685, Proceedings of the SPIE, pp. 182–192, Sept. 1992.

[34] M. H. Weiler, "Magnetooptical properties of $Hg_{1-x}Cd_xTe$ alloys," in *Defects, (HgCd)Se, (HgCd)Te* (R. K. Willardson and A. C. Beer, eds.), vol. 16 of *Semiconductors and Semimetals*, ch. 3, pp. 119–191, New York: Academic Press, 1981.

[35] J. P. Laurenti, J. Camassel, A. Bouhemadou, B. Toulouse, R. Legros, and A. Lusson, "Temperature dependence of the fundamental absorption edge of mercury cadmium telluride," *J. Appl. Phys.*, vol. 67, pp. 6454–6460, May 1990.

[36] J. Wenus, J. Rutkowski, and A. Rogalski, "Two-dimensional analysis of double-layer heterojunction HgCdTe photodiodes," *IEEE Trans. Electron Devices*, vol. 48, pp. 1326–1332, July 2001.

[37] J. P. Rosbeck, R. E. Starr, S. L. Price, and K. J. Riley, "Background and temperature dependent current–voltage characteristics of HgCdTe photodiodes," *J. Appl. Phys.*, vol. 53, pp. 6430–6441, Sept. 1982.

[38] W. M. Higgins, G. N. Pultz, R. G. Roy, R. A. Lancaster, and J. L. Schmit, "Standard relationships in the properties of $Hg_{1-x}Cd_xTe$," *J. Vac. Sci. Technol. A*, vol. 7, pp. 271–275, Mar. 1989.

[39] J. D. Patterson, W. A. Gobba, and S. L. Lehoczky, "Electron mobility in n-type $Hg_{1-x}Cd_xTe$ and $Hg_{1-x}Zn_xTe$ alloys," *J. Mater. Res.*, vol. 7, pp. 2211–2218, Aug. 1992.

[40] J. R. Lowney, D. G. Seiler, C. L. Littler, and I. T. Yoon, "Intrinsic carrier concentration of narrow-gap mercury cadmium telluride based on the nonlinear temperature dependence of the band gap," *J. Appl. Phys.*, vol. 71, pp. 1253–1258, Feb. 1992.

[41] G. L. Hansen and J. L. Schmit, "Calculation of intrinsic carrier concentration in $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 53, pp. 1639–1640, Mar. 1983.

[42] F. L. Madarasz, F. Szmulowicz, and J. R. McBath, "Intrinsic carrier concentrations and effective masses in $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 58, no. 1, pp. 361–365, 1985.

[43] V. C. Lopes, A. J. Syllaios, and M. C. Chen, "Minority carrier lifetime in mercury cadmium telluride," *Semiconductor Sci. Tech.*, vol. 8, pp. 824–841, June 1993.

[44] A. R. Beattie and P. T. Landsberg, "Auger effect in semiconductors," *Proc. Royal Soc. London A*, vol. 246, pp. 16–29, Jan. 1959.

[45] J. S. Blakemore, *Semiconductor Statistics*, vol. 3 of *International Series of Monographs on Semiconductors*. New York: Pergamon Press, 1962.

[46] M. A. Kinch, F. Aqariden, D. Chandra, P.-K. Liao, H. F. Schaake, and H. D. Shih, "Minority carrier lifetime in *p*-HgCdTe," *J. Electron. Mater.*, vol. 34, no. 6, pp. 880–884, 2005.

[47] M. A. Kinch, M. J. Brau, and A. Simmons, "Recombination mechanisms in 8–14–μ HgCdTe," *J. Appl. Phys.*, vol. 44, no. 4, pp. 1649–1663, 1973.

[48] F. Bertazzi, M. Goano, and E. Bellotti, "Calculation of Auger lifetime in HgCdTe," *J. Electron. Mater.*, vol. 40, no. 8, pp. 1663–1667, 2011.

[49] S. Krishnamurthy and T. N. Casselman, "A detailed calculation of the Auger lifetime in *p*-type HgCdTe," *J. Electron. Mater.*, vol. 29, no. 6, pp. 828–831, 2000.

[50] M. C. Chen, L. Colombo, J. A. Dodge, and J. H. Tregilgas, "The minority carrier lifetime in doped and undoped *p*-type $Hg_{0.78}Cd_{0.22}Te$ liquid phase epitaxy films," *J. Electron. Mater.*, vol. 24, no. 5, pp. 539–544, 1995.

[51] S. Velicu, C. H. Grein, P. Y. Emelie, A. Itsuno, J. D. Phillips, and P. S. Wijewarnasuriya, "MWIR and LWIR HgCdTe infrared detectors operated with reduced cooling requirements," *J. Electron. Mater.*, vol. 39, no. 7, pp. 873–881, 2010.

[52] S. E. Schacham and E. Finkman, "Recombination mechanisms in *p*-type HgCdTe: Freezeout and background flux effects," *J. Appl. Phys.*, vol. 57, pp. 2001–2009, Mar. 1985.

[53] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, 1952.

[54] R. N. Hall, "Electron–hole recombination in germanium," *Phys. Rev.*, vol. 87, no. 2, p. 387, 1952.

[55] C. A. Hougen, "Model for infrared absorption and transmission of liquid-phase epitaxy $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 66, no. 8, pp. 3763–3766, 1989.

[56] E. Finkman and S. Schacham, "The exponential optical absorption band tail of $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 56, pp. 2896–2900, Nov. 1984.

[57] P. Capper and J. Garland, eds., *Mercury Cadmium Telluride. Growth, Properties and Applications*. Chichester, U.K.: John Wiley & Sons, 2011.

[58] E. O. Kane, "Theory of tunneling," *J. Appl. Phys.*, vol. 32, no. 1, pp. 83–89, 1961.

[59] R. Adar, "Spatial integration of direct band-to-band tunneling currents in general device structures," *IEEE Trans. Electron Devices*, vol. 39, pp. 976–981, Apr. 1992.

[60] K. Jóźwikowski, M. Kopytko, A. Rogalski, and A. Jóźwikowska, "Enhanced numerical analysis of current-voltage characteristics of long wavelength infrared n-on-p HgCdTe photodiodes," *J. Appl. Phys.*, vol. 108, no. 7, p. 074519, 2010.

[61] M. A. Kinch, "HgCdTe: Recent trends in the ultimate IR semiconductor," *J. Electron. Mater.*, vol. 39, no. 7, pp. 1043–1052, 2010.

[62] D. D'Orsogna, S. P. Tobin, and E. Bellotti, "Numerical analysis of a very long-wavelength HgCdTe pixel array for infrared detection," *J. Electron. Mater.*, vol. 37, no. 9, pp. 1349–1355, 2008.

[63] J. Schuster, B. Pinkie, S. Tobin, C. Keasler, D. D'Orsogna, and E. Bellotti, "Numerical simulation of third-generation HgCdTe detector pixel arrays," *IEEE J. Select. Topics Quantum Electron.*, vol. 19, no. 5, p. 800415, 2013.

[64] J. Schuster and E. Bellotti, "Analysis of optical and electrical crosstalk in small pitch photon trapping HgCdTe pixel arrays," *Appl. Phys. Lett.*, vol. 101, no. 26, p. 261118, 2012.

[65] A. K. Sood, J. E. Egerton, Y. R. Puri, E. Bellotti, D. D'Orsogna, L. Becker, R. Balcerak, K. Freyvogel, and R. Richwine, "Design and development of multicolor MWIR/LWIR and LWIR/VLWIR detector arrays," *J. Electron. Mater.*, vol. 34, pp. 909–912, June 2005.

[66] A. M. Itsuno, J. D. Phillips, and S. Velicu, "Predicted performance improvement of Auger-suppressed HgCdTe photodiodes and *p-n* heterojunction detectors," *IEEE Trans. Electron Devices*, vol. 58, pp. 501—507, Feb. 2011.

[67] P. Chakrabarti, A. Krier, and A. F. Morgan, "Analysis and simulation of a mid-infrared P^+-$InAs_{0.55}Sb_{0.15}$ $P_{0.30}$/n^0-$InAs_{0.89}Sb_{0.11}$/N^+-$InAs_{0.55}Sb_{0.15}P_{0.30}$ double heterojunction photodetector grown by LPE," *IEEE Trans. Electron Devices*, vol. 50, no. 10, pp. 2049–2058, 2003.

[68] Y. Tian, B. Zhang, T. Zhou, J. Hong, and Y. Jin, "Theoretical analysis of the detectivity in N-p and P-n GaSb/GaInAsSb infrared photodetectors," *IEEE Trans. Electron Devices*, vol. 47, no. 3, pp. 544–552, 2000.

[69] E. R. Blazejewski, J. M. Arias, G. M. Williams, W. McLevige, M. Zandian, and J. Pasko, "Bias-switchable dual-band HgCdTe infrared photodetector," *J. Vac. Sci. Technol. B*, vol. 10, pp. 1626–1632, July 1992.

[70] R. A. Coussa, A. M. Gallagher, K. Kosai, L. T. Pham, G. K. Pierce, E. P. Smith, G. M. Venzor, T. J. D. Lyon, J. E. Jensen, B. Z. Nosho, J. A. Roth, and J. R. Waterman, "Spectral crosstalk by radiative recombination in sequential-mode, dual mid-wavelength infrared band HgCdTe detectors," *J. Electron. Mater.*, vol. 33, pp. 517–525, June 2004.

[71] K.-M. Mahlein, A. Bauer, H. Bitterlich, M. Bruder, K.-U. Gassmann, M. Haiml, S. Hanna, H.-P. Nothaft, R. Wollrab, and J. Ziegler, "Next generation IR sensor technology for space applications at AIM," in *Sensors, Systems, and Next-Generation Satellites XII* (R. Meynart, S. P. Neeck, H. Shimoda, and S. Habib, eds.), vol. 7106, Proceedings of the SPIE, pp. 71061J–1–71061J–10, Oct. 2008.

[72] E. Bellotti and D. D'Orsogna, "Numerical analysis of HgCdTe simultaneous two-color photovoltaic infrared detectors," *IEEE J. Quantum Electron.*, vol. 42, pp. 418–426, Apr. 2006.

[73] H. B. DeVore, "Spectral distribution of photoconductivity," *Phys. Rev.*, vol. 102, no. 1, pp. 86–91, 1956.

[74] C. Bouchenaki, B. Ullrich, J. P. Zielinger, H. N. Cong, and P. Chartier, "Preparation, characterization, and bistable photoconduction properties of thin CdS layers," *J. Opt. Soc. Amer. B*, vol. 8, no. 3, pp. 691–700, 1991.

[75] B. Pejova, "Analysis of the shape of spectral dependence of absorption coefficient and stationary photoconductivity spectral response in nanocrystalline bismuth(III) sulfide thin films," *Mater. Res. Bull.*, vol. 43, no. 3, pp. 2887–2903, 2008.

[76] B. Ullrich and H. Xi, "Photocurrent theory based on coordinate dependent lifetime," *Opt. Lett.*, vol. 35, no. 23, pp. 3910–3912, 2010.

[77] M. B. Reine, "HgCdTe photodiodes for IR detection: A review," in *Photodetector Materials and Devices VI*, vol. 4288, Proceedings of the SPIE, Jan. 2001.

[78] J. Wenisch, H. Bitterlich, M. Bruder, P. Fries, R. Wollrab, J. Wendler, R. Breiter, and J. Ziegler, "Large-format and long-wavelength infrared mercury cadmium telluride detectors," *J. Electron. Mater.*, vol. 42, no. 11, pp. 3186–3190, 2013.

[79] A. S. Verhulst, D. Leonelli, R. Rooyackers, and G. Groeseneken, "Drain voltage dependent analytical model of tunnel field-effect transistors," *J. Appl. Phys.*, vol. 110, no. 2, p. 024510, 2011.

[80] K. Ahmed, M. M. M. Elahi, and M. S. Islam, "A compact analytical model of band-to-band tunneling in a nanoscale p-i-n diode," in *2012 International Conference on Informatics, Electronics and Vision (ICIEV)*, May 2012.

[81] R. Wollrab, W. Schirmacher, T. Schallenberg, H. Lutz, J. Wendler, and J. Ziegler, "Recent progress in the development of hot MWIR detectors," in *6th International Symposium on Optronics in Defence and Security (OPTRO 2014)*, (Paris), Feb. 2014.

[82] T. Casselman, "Calculation of the Auger lifetime in p-type $Hg_{1-x}Cd_xTe$," *J. Appl. Phys.*, vol. 52, no. 2, pp. 848–854, 1981.

[83] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. Hoboken, NJ: John Wiley & Sons, 3rd ed., 2007.

[84] S. Krishnamurthy, M. A. Berding, Z. G. Yu, C. H. Swartz, T. H. Myers, D. D. Edwall, and R. DeWames, "Model for minority carrier lifetimes in doped HgCdTe," *J. Electron. Mater.*, vol. 34, no. 6, pp. 873–879, 2005.

[85] W. Vandenberghe, B. Sorée, W. Magnus, and M. V. Fischetti, "Generalized phonon assisted Zener tunneling in indirect semiconductors with non-uniform electric fields: a rigorous approach," *J. Appl. Phys.*, vol. 109, no. 12, p. 124503, 2011.

[86] Y. Okuto and C. R. Crowell, "Energy-conservation considerations in the characterization of impact ionization in semiconductors," *Phys. Rev. B*, vol. 6, pp. 3076–3081, Oct. 1972.

[87] M. A. Kinch, J. D. Beck, C.-F. Wan, F. Ma, and J. Campbell, "HgCdTe electron avalanche photodiodes," *J. Electron. Mater.*, vol. 33, pp. 630–639, June 2004.

[88] C. T. Elliott, N. T. Gordon, R. S. Hall, and G. Crimes, "Reverse breakdown in long wavelength lateral collection $Cd_xHg_{1-x}Te$ diodes," *J. Vac. Sci. Technol. A*, vol. 8, no. 2, pp. 1251–1253, 1990.

[89] J. Rothman, L. Mollard, S. Goût, L. Bonnefond, and J. Wlassow, "History-dependent impact ionization theory applied to HgCdTe e-APDs," *J. Electron. Mater.*, vol. 40, no. 8, pp. 1757–1768, 2011.

[90] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagation*, vol. 14, pp. 302–307, May 1966.

[91] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, vol. 114, no. 2, pp. 185–200, 1994.

[92] D. Vasileska, S. M. Goodnick, and G. Klimeck, *Computational Electronics. Semi-classical and Quantum Device Modeling and Simulation*. Boca Raton, FL: CRC Press, 2010.

[93] M. Salazar-Palma, T. K. Sarkar, L. E. Garcia-Costillo, and T. Roy, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling*. Norwood, MA: Artech House, 1998.

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

[95] M. Born and E. Wolf, *Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge, U.K.: Cambridge University Press, 7th ed., 1999.

[96] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *J. Opt. Soc. Amer.*, vol. 71, pp. 811–818, July 1981.

[97] J.-M. Liu, *Photonic devices*. Cambridge, U.K.: Cambridge University Press, 2005.

[98] G. H. Spencer and M. V. R. K. Murty, "General ray-tracing procedure," *J. Opt. Soc. Amer.*, vol. 52, pp. 672–676, June 1962.

[99] T. Rahman and K. Fobelets, "Efficient tool flow for 3D photovoltaic modelling," *Comp. Phys. Comm.*, vol. 193, pp. 124–130, Aug. 2015.

[100] C. Keasler and E. Bellotti, "Three-dimensional electromagnetic and electrical simulation of HgCdTe pixel arrays," *J. Electron. Mater.*, vol. 40, no. 8, pp. 1795–1801, 2011.

[101] B. Pinkie and E. Bellotti, "Numerical simulation of spatial and spectral crosstalk in two-color MWIR/LWIR HgCdTe infrared detector arrays," *J. Electron. Mater.*, vol. 42, no. 11, pp. 3080–3089, 2013.

[102] R. S. Saxena, N. K. Saini, R. Bhan, and R. Sharma, "A new circuit model of HgCdTe photodiode for SPICE simulation of integrated IRFPA," *Infrared Phys. Tech.*, vol. 67, pp. 58–62, 2014.

[103] R. G. Driggers, R. Vollerhauser, P. Reynolds, J. Fanning, and G. C. Holst, "Infrared detector size: how low should you go?," *Opt. Eng.*, vol. 51, no. 6, p. 063202, 2012.

[104] W. Tennant, D. Gulbransen, A. Roll, M. Carmody, D. Edwall, A. Julius, P. Drieske, A. Chen, W. McLevige, S. Freeman, D. Lee, D. Cooper, and E. Piquette, "Small-pitch hgcdte photodetectors," *J. Electron. Mater.*, vol. 43, no. 8, pp. 3041–3046, 2014.

[105] A. Rogalski, P. Martyniuk, and M. Kopytko, "Challenges of small-pixel infrared detectors: a review," *Rep. Prog. Phys.*, vol. 79, no. 4, p. 046501, 2016.

[106] G. Verzellesi, D. Saguatti, M. Meneghini, F. Bertazzi, M. Goano, G. Meneghesso, and E. Zanoni, "Efficiency droop in InGaN/GaN blue light-emitting diodes: Physical mechanisms and remedies," *J. Appl. Phys.*, vol. 114, p. 071101, Aug. 2013.

[107] D. Saguatti, L. Bidinelli, G. Verzellesi, M. Meneghini, G. Meneghesso, E. Zanoni, R. Butendeich, and B. Hahn, "Investigation of efficiency-droop mechanisms in multi-quantum-well InGaN/GaN blue light-emitting diodes," *IEEE Trans. Electron Devices*, vol. 59, pp. 1402–1409, May 2012.

[108] M. Asada, "Intraband relaxation time in quantum-well lasers," *IEEE J. Quantum Electron.*, vol. 25, pp. 2019–2026, Sept. 1989.

[109] R. Binder, D. Scott, A. E. Paul, M. Lindberg, K. Henneberger, and S. W. Koch, "Carrier-carrier scattering and optical dephasing in highly excited semiconductors," *Phys. Rev. B*, vol. 45, pp. 1107–1115, Jan. 1992.

[110] P. Sotirelis and K. Hess, "Electron capture in GaAs quantum wells," *Phys. Rev. B*, vol. 49, pp. 7543–7547, Mar. 1994.

[111] J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures*. Berlin: Springer-Verlag, 1996.

[112] R. C. Iotti and F. Rossi, "Carrier thermalization versus phonon-assisted relaxation in quantum-cascade lasers: A monte carlo approach," *Appl. Phys. Lett.*, vol. 78, pp. 2902 –2904, May 2001.

[113] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Physics*. New York: Dover Publications, 3rd ed., 2003.

[114] G. D. Mahan, *Many-Particle Physics*. New York: Plenum Press, 2nd ed., 1990.

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

[116] H. Haug and S. W. Koch, *Quantum theory of the optics and electronic properties of semiconductors*. Singapore: World Scientific Publishing, 2004.

[117] H. Haug and S. W. Koch, "Semiconductor laser theory with many-body effects," *Phys. Rev. A*, vol. 39, pp. 1887–1898, Feb. 1989.

[118] P. J. Price, "Polar-optical-mode scattering for an ideal quantum-well heterostructure," *Phys. Rev. B*, vol. 30, pp. 2234–2235, Aug. 1984.

[119] J. A. Brum and G. Bastard, "Resonant carrier capture by semiconductor quantum wells," *Phys. Rev. B*, vol. 33, pp. 1420–1423, Jan. 1986.

[120] P. W. M. Blom, C. Smit, J. E. M. Haverkort, and J. H. Wolter, "Carrier capture into a semiconductor quantum well," *Phys. Rev. B*, vol. 47, pp. 2072–2081, Jan. 1993.

[121] S. A. Levetas and M. J. Godfrey, "Calculation of capture of carriers by quantum wells," *Phys. Rev. B*, vol. 59, pp. 10202–10207, Apr. 1999.

[122] P. Sotirelis, P. von Allmen, and K. Hess, "Electron intersubband relaxation in doped quantum wells," *Phys. Rev. B*, vol. 47, pp. 12744–12753, May 1993.

[123] T. R. Nielsen, P. Gartner, and F. Jannke, "Many-body theory of carrier capture and relaxation in semiconductor quantum-dot lasers," *Phys. Rev. B*, vol. 69, pp. 235314–235327, June 2004.

[124] K. Schuh, P. Gartner, and F. Jahnke, "Combined influence of carrier-phonon and coulomb scattering on the quantum-dot population dynamics," *Phys. Rev. B*, vol. 87, pp. 035301–035308, Jan. 2013.

[125] S. M. Goodnick, S. Limpert, C. Honsberg, and P. Lugli, "Simulation of carrier relaxation in hot carrier solar cells," in *Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE*, pp. 001657–001662, June 2012.

[126] A. Kukharskii, "Plasmon-phonon coupling in GaAs," *Solid State Communications*, vol. 13, no. 11, pp. 1761 – 1765, 1973.

[127] S. D. Sarma, J. K. Jain, and R. Jalabert, "Theory of hot-electron energy loss in polar semiconductors: Role of plasmon-phonon coupling," *Phys. Rev. B*, vol. 37, pp. 6290–6296, Apr. 1988.

[128] M. Takeshima, "Phonon-assisted auger recombination in a quasi-two-dimensional structure semiconductor," *Phys. Rev. B*, vol. 30, pp. 3302–3308, Sept. 1984.

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

[130] J. Piprek, "Efficiency droop in nitride-based light-emitting diodes," *Phys. Status Solidi A*, vol. 207, pp. 2217–2225, Oct. 2010.

[131] M. Snelling, P. Perozzo, D. C. Hutchings, I. Galbraith, and A. Miller, "Investigation of excitonic saturation by time-resolved circular dichroism in GaAs-$Al_xGa_{1-x}As$ multiple quantum wells," *Phys. Rev. B*, vol. 49, pp. 17160–17169, June 1994.

[132] G. D. Mahan, *Many–particle Physics*. New York: Plenum press, 1986.

[133] P. Coleman, *Introduction to many body physics*. Cambridge, USA: Cambridge University Press, 2011.

[134] T. Y. Wang and G. B. Stringfellow, "Strain effects on $Ga_xIn_{1-x}As$/InP single quantum wells grown by organometallic vapor phase epitaxy with $0 \leq x \leq 1$," *J. Appl. Phys.*, vol. 67, no. 1, pp. 344–352, 1990.

[135] S. M. Goodnick and P. Lugli, "Effect of electron-electron scattering on nonequilibrium transport in quantum-well systems," *Phys. Rev. B*, vol. 37, pp. 2578–2588, Feb. 1988.

[136] S.-C. Lee and I. Galbraith, "Intersubband and intrasubband electronic scattering rates in semiconductor quantum wells," *Phys. Rev. B*, vol. 59, pp. 15796–15805, June 1999.

[137] W. H. Fan, S. M. Olaizola, J.-P. R. Wells, A. M. Fox, T. Wang, P. J. Parbrook, D. J. Mowbray, and M. S. Skolnick, "Femtosecond studies of electron capture times in

InGaN/GaN multiple quantum wells," *Appl. Phys. Lett.*, vol. 84, pp. 3052–3054, Apr. 2004.

[138] W. H. Fan, S. M. Olaizola, J.-P. R. Wells, A. M. Fox, T. Wang, P. J. Parbrook, D. J. Mowbray, and M. S. Skolnick

[139] M. H. Crawford, "LEDs for solid-state lighting: Performance challenges and recent advances," *IEEE J. Select. Topics Quantum Electron.*, vol. 15, pp. 1028–1040, July–Aug. 2009.

[140] A. Laubsch, M. Sabathil, W. Bergbauer, M. Strassburg, H. Lugauer, M. Peter, S. Lutgen, N. Linder, K. Streubel, J. Hader, J. V. Moloney, B. Pasenow, and S. W. Koch, "On the origin of IQE-'droop' in InGaN LEDs," *Phys. Status Solidi C*, vol. 6, no. S2, pp. S913–S916, 2009.

[141] Ü. Özgür, H. Liu, X. Li, X. Ni, and H. Morkoç, "GaN-based light-emitting diodes: Efficiency at high injection levels," *Proc. IEEE*, vol. 98, pp. 1180–1196, July 2010.

[142] M. Calciati, M. Goano, F. Bertazzi, M. Vallone, X. Zhou, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, E. Bellotti, G. Verzellesi, D. Zhu, and C. Humphreys, "Correlating electroluminescence characterization and physics-based models of InGaN/GaN LEDs: Pitfalls and open issues," *AIP Adv.*, vol. 4, p. 067118, June 2014.

[143] V. Avrutin, S. A. Hafiz, F. Zhang, Ü. Özgür, E. Bellotti, F. Bertazzi, M. Goano, A. Matulionis, A. T. Roberts, H. O. Everitt, and H. Morkoç, "Saga of efficiency degradation at high injection in InGaN light emitting diodes," *Turk. J. Phys.*, vol. 38, pp. 269–313, Nov. 2014.

[144] D. S. Sizov, R. Bhat, A. Zakharian, K. Song, D. E. Allen, S. Coleman, and C. en Zah, "Carrier transport in InGaN MQWs of aquamarine and green-lased diodes," *IEEE J. Select. Topics Quantum Electron.*, vol. 17, pp. 1390–1401, Sept. 2011.

[145] M. Takeshima, "Simple method of calculating phonon-assisted Auger recombination rate in direct-gap semiconductors," *Japan. J. Appl. Phys.*, vol. 22, pp. 491–498, Mar. 1983.

[146] W. Bardyszewski and D. Yevick, "Compositional dependence of the Auger coefficient for InGaAsP lattice matched to InP," *J. Appl. Phys.*, vol. 58, pp. 2713–2723, Oct. 1985.

[147] E. Kioupakis, P. Rinke, K. T. Delaney, and C. G. Van de Walle, "Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes," *Appl. Phys. Lett.*, vol. 98, no. 16, p. 161107, 2011.

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

[149] N. S. Mansour, K. W. Kim, and M. A. Littlejohn, "Theoretical study of electron transport in gallium nitride," *J. Appl. Phys.*, vol. 77, pp. 2834–2836, Mar. 1995.

[150] D. Bradt, Y. Sirenko, and V. Mitin, "Inelastic and elastic mechanisms of electron capture to a quantum well," *Semiconductor Sci. Tech.*, vol. 10, pp. 260–269, 1995.

[151] M. Abou-Khalil, M. Goano, A. Champagne, and R. Maciejko, "Capture and escape in quantum wells as scattering events in Monte Carlo simulation," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 19–21, Jan. 1996.

[152] M. Moško and K. Kálna, "Carrier capture into a GaAs quantum well with a separate confinement region: comment on quantum and classical aspects," *Semiconductor Sci. Tech.*, vol. 14, pp. 790–796, 1999.

[153] J. Wang, K. W. Kim, and M. A. Littlejohn, "Carrier capture in pseudomorphically strained wurtzite GaN quantum-well lasers," *Appl. Phys. Lett.*, vol. 71, no. 6, pp. 820–822, 1997.

[154] X. Zhang, D. H. Rich, J. T. Kobayashi, N. P. Kobayashi, and P. D. Dapkus, "Carrier relaxation and recombination in an InGaN/GaN quantum well probed with time-resolved cathodoluminescence," *Appl. Phys. Lett.*, vol. 73, pp. 1430–1432, Sept. 1998.

[155] U. Özgür, M. J. Bergmann, H. C. C. Jr., H. O. Everitt, A. C. Abare, S. Keller, , and S. P. DenBaars, "Ultrafast optical characterization of carrier capture times in InGaN multiple quantum wells," *Appl. Phys. Lett.*, vol. 77, pp. 109 –111, July 2000.

[156] H.-C. Wang, Y.-C. Lu, C.-Y. Chen, and C. C. Yang, "Carrier capture times of the localized states in an InGaN thin film with indium-rich nanocluster structures," *Appl. Phys. Lett.*, vol. 89, no. 1, p. 011906, 2006.

[157] P. W. M. Blom, J. E. M. Haverkort, P. J. van Hall, and J. H. Wolter, "Carrier-carrier scattering induced capture in quantum well lasers," *Appl. Phys. Lett.*, vol. 62, pp. 1490–1492, Mar. 1993.

[158] S. Chiaria, E. Furno, M. Goano, and E. Bellotti, "Design criteria for near-ultraviolet GaN-based light-emitting diodes," *IEEE Trans. Electron Devices*, vol. 57, pp. 60–70, Jan. 2010.

[159] M. Meneghini, N. Trivellin, G. Meneghesso, E. Zanoni, U. Zehnder, and B. Hahn, "A combined electro-optical method for the determination of the recombination parameters in InGaN-based light-emitting diodes," *J. Appl. Phys.*, vol. 106, no. 11, p. 114508, 2009.

[160] M. Zhang, P. Bhattacharya, J. Singh, and J. Hinckley, "Direct measurement of Auger recombination in $In_{0.1}Ga_{0.9}N$/GaN quantum wells and its impact on the efficiency of $In_{0.1}Ga_{0.9}N$/GaN multiple quantum well light emitting diodes," *Appl. Phys. Lett.*, vol. 95, no. 20, p. 201108, 2009.

[161] Q. Dai, Q. Shan, J. Wang, S. Chhajed, J. Cho, E. F. Schubert, M. H. Crawford, D. D. Koleske, M.-H. Kim, and Y. Park, "Carrier recombination mechanisms and efficiency droop in GaInN/GaN light-emitting diodes," *Appl. Phys. Lett.*, vol. 97, no. 13, p. 133507, 2010.

[162] A. Laubsch, M. Sabathil, J. Baur, M. Peter, and B. Hahn, "High-power and high-efficiency InGaN-based light emitters," *IEEE Trans. Electron Devices*, vol. 57, pp. 79–87, Feb. 2010.

[163] A. David and M. J. Grundmann, "Droop in InGaN light-emitting diodes: A differential carrier lifetime analysis," *Appl. Phys. Lett.*, vol. 96, no. 10, p. 103504, 2010.

[164] W. G. Scheibenzuber, U. T. Schwarz, L. Sulmoni, J. Dorsaz, J.-F. Carlin, and N. Grandjean, "Recombination coefficients of GaN-based laser diodes," *J. Appl. Phys.*, vol. 109, no. 9, p. 093106, 2011.

[165] D. Schiavon, M. Binder, M. Peter, B. Galler, P. Drechsel, and F. Scholz, "Wavelength-dependent determination of the recombination rate coefficients in single-quantum-well GaInN/GaN light emitting diodes," *Phys. Status Solidi B*, vol. 250, pp. 283–290, Feb. 2013.

[166] P. Kinsler, P. Harrison, and R. W. Kelsall, "Intersubband electron-electron scattering in asymmetric quantum wells designed for far-infrared emission," *Phys. Rev. B*, vol. 58, pp. 4771–4778, Aug. 1998.