Politecnico di Torino

Dipartimento di Elettronica e Telecomunicazione - DET

Master's degree in Nanotechnologies for the ICTs

Carrier injection modelling in buried tunnel junction VCSELs

Author:
Alberto Gullino

Relatori:
Prof. Francesco Bertazzi
Prof. Michele Goano
Dr. Alberto Tibaldi
Dr. Pierluigi Debernardi

October, 2019

Abstract

After conquering the data-communication and sensing markets, vertical-cavity surface-emitting lasers (VCSELs) are now playing a key role in imaging applications in smartphones and automotive contexts. In fact, among their "competitors", these lasers exhibit unique features in terms of power consumption, reliability, testability, packaging costs and array-oriented manufacturability, making them the ideal light source for portable applications. In order to export these advantages to the widest range of upcoming applications, the VCSEL community is exploring novel concepts and materials, aiming to *move* the emission from the consolidated 850 nm wavelength towards mid-infrared and/or blue ranges.

A criticality of VCSEL design is the injection of holes in the active region, which is controlled, in *classical* AlGaAs VCSELs, by an aperture obtained through wet oxidation processes. A pitfall of the to-be-explored material systems is the technological unavailability of such processes, combined with poor hole transport processes, which demand for alternative strategies. A very interesting approach concerns the employment of a buried tunnel junctions (BTJ). Tunnel junctions consist of heavily doped *pn* junctions where, under reverse bias operation, the conduction and valence bands are overlooking, and carrier transport can take place through band-to-band tunneling (BTBT). In addition to replacing the oxide aperture, BTJs allow to eliminate *p*-doped contacts, with huge benefits to the VCSEL electrical conductivity and giving room to better thermal control and enhanced modulation speed overcoming the state of the art. Despite the great interest in this subject, the frantic times-to-market discourage investments such a revolution of the VCSEL architecture, which would require extensive and expensive prototyping campaigns. In this view, computer-aided design (CAD) tools play a starring role as alternative prototyping frameworks for research and development departments.

The purpose of this Master's thesis is to include the BTBT quantum corrections within the drift-diffusion carrier transport picture implemented in the multiphysics Vcsel Electro-opto-thermal NUmerical Simulator VENUS, developed by CNR-IEIIT and Politecnico di Torino. Two BTJ-VCSELs are investigated as case studies: an InGaAsP long-wavelength device, and an AlGaAs device manufactured and characterized at Chalmers University of Technology. In the first part of the thesis Hurkx's work has been reviewed and implemented, which is based on local generation-recombination (GR) rates accounting for the junction electric field. Even though this model is quite consolidated for silicon electronics and implemented in commercial simulators such as *Sentaurus Device* by Synopsys, it doesn't appear to be applicable to these material

systems and devices. Therefore, the second part of the thesis explores a non-local GR model derived from a 4-band non-equilibrium Green's functions (NEGF) formalism. It is demonstrated that the ballistic nature of BTBT allows to apply the NEGF solver in the coherent limit, lowering its staggering computational cost. This observation is at the basis of the self-consistent NEGF-VENUS simulation framework, enabling to couple a semiclassical picture of the bulky sections of the device with a genuine quantum description of BTBT.

Contents

CHAPTER 1

Introduction

The main goal of the whole thesis is the extensive analysis of the buried tunnel junctions (BTJ) working principles and their introduction inside state-of-the-art vertical-cavity surface-emitting laser (VCSEL) structures to improve their performances. In fact, as presented afterwards, due to the several issues caused by the p-doped contact of a VCSEL structure and by the oxide aperture production process, BTJs have been demonstrated to be a good solution, able to mitigate some of the problems. This work is mainly focused on carrier transport and electrical results, while optical and thermal simulations are left out from the analysis, even though they also represent a crucial portion of the overall problem, that in the end needs to be tackled by means of a multiphysics simulator able to solve simultaneously and in a self-consistent fashion the electrical, thermal and optical problems.

Throughout this introduction, first a brief description of both short and long-wavelength VCSELs together with their corresponding applications is presented. Of course, also their structure and the problems related to them are described. Then, a section is devoted to a preliminary description of the buried tunnel junctions structure and working principle, and especially the reasons they are implemented for. A great relevance is obviously given to the physical and mathematical concepts behind all the obtained results. Therefore, chapter 2 is completely dedicated to their presentation and to how they are implemented in the exploited softwares. Of course, all the formal justifications are presented for the use of drift-diffusion model, together with the main assumptions introduced to make the equations of such a model more easily manageable by a numerical simulator. As a matter of fact, the action of the tunnel junctions is initially investigated with a quantum *local* model, originally born to describe silicon systems and then adapted to other semiconductors, which is presented in details in chapter 3; then tunneling issue is tackled with a more sophisticated mathematical and physical approach, based on non-equilibrium Green's functions (NEGF) formalism, able to extract a *non-local* generation/recombination model. The latter is briefly

described at the end of chapter 4.

Given that two BTBT models are actually exploited, the whole analysis involves three distinct simulators. The main one is D1ANA, a drift-diffusion (DD) based MATLAB code developed by CNR-IEIIT and Politecnico di Torino, that provides the carrier transport picture for the multiphysics Vcsel Electro-opto-thermal NUmerical Simulator (VENUS) [1], which is able to simulate both short and long structures, such as a complete VCSEL devices. Furthermore, the simplest but most critical and interesting parts of these devices (i.e., the BTJs themselves) are extensively explored by means of the commercial software *Sentaurus Device* TCAD by Synopsys (where carrier tranport is based on drift-diffusion model as well) and by a NEGF code developed by Prof. Francesco Bertazzi and dr. Alberto Tibaldi, which relies on a completely different physical theory. The former was first introduced to me by Professors Simona Donati Guerrieri and Michele Goano during two of their courses in two semesters of the Master Degree, but I also have to thank dr. Alberto Tibaldi for providing me some basic files to start my work with. Such a TCAD (Technology computer-aided design) suite has proved to be very useful to validate the correct implementation of the local band-to-band tunneling quantum corrections into the original D1ANA code (which are introduced for the first time during the thesis work), even though it has not been used to perform calculations on the whole structure, that would have been beyond the scopes of the thesis work.

The intentions when starting using NEGF in the first place were to further validate some results coming from drift-diffusion including Hurkx's model (see section 4.3). On the contrary, this actually sorted an opposite effect, revealing that something in the simple local BTBT model was missing, and that a more detailed model should be taken into account to obtain results closer to the experimental ones. Therefore, the task demands for a genuine quantum approach to the tunneling problem, and a good candidate is NEGF formalism itself. However, NEGF is here used as an instrument to extract data from which compute a GR rate to be plugged inside the previously mentioned drift-diffusion code, and allows for simple comparisons with experimental results and already performed simulations. In fact, the last step of the work is the inclusion of the computed non-local GR rates in the D1ANA code, that actually leads to reproduce more useful and reliable results on the studied structures (in particular for the short wavelength device), with respect to the simpler local tunneling model.

All the results coming from simulations (DD and NEGF) on both short and long-wavelength VCSELs are showed in the last chapter, together with the description of all the unsolved issues. In particular, the two case studies of the thesis are an InGaAsP long-wavelength VCSEL and an AlGaAs short-wavelength VCSEL: extensive analysis of their BTJs are also provided for a comprehensive understanding of the matter.

1.1 VCSELs

Before entering into the details and describing accurately how a tunnel junction works, let's start from addressing the primary reason to investigate and find models able to simulate them properly: their inclusion inside vertical cavity surface emitting lasers, electrically- or optically-pumped *pin* diodes.

VCSELs have been developed as an alternative to the traditional in-plane lasers, also called *edge emitters*. In fact, due to the aperture field pattern of the latter, the emitted radiation diagram is not circular, making difficult the fiber coupling process. This issue is readily solved in a VCSEL, since its circular emitting area (placed perpendicularly with respect to the active layers growth) translates into a circular far-field beam [2], thus providing high coupling efficiencies, high modulation bandwidths at low current levels, single-mode operation (due to its short cavity length), and low power consumption. A vertical cavity surface emitting laser is able to deliver single mode optical power as large as few mW, limited by the small aperture required for maintaining sufficient discrimination against higher order transverse modes [3]. Furthermore, it allows for high-volume, low-cost manufacturing, because of its compatibility with low-cost wafer scale fabrication (such as ion implantation) and easy testing methods [4].

However, the "standard" VCSEL structure presents some issues related to power dissipation (i.e., self heating of the device) and to the confinement of the emitted light, which need to be addressed. Such problems are mainly related to the low-resistivity p-side contacts and to the oxide aperture technique, as described in the following section. For this reasons, it has been found that the use of buried tunnel junctions may provide the needed improvements, as presented in section 1.2.

1.1.1 Some applications

Depending on the output light beam wavelength a VCSEL may have different applications. **Short**-wavelength VCSELs are typically made in AlGaAs/GaAs due to its wide gap, and are used for several purposes regarding laser products (computer mice, face ID, fiber optics communications - FC), ethernet applications (10-100 m connections, for $\lambda = 850 - 980$ nm), optical trace-gas sensing (by means of tunable diode laser spectroscopy, e.g. 759 nm laser, for O_2 detection), atomic clocks (780-895 nm) and even as light sources for laser printers. As a matter of fact, Gigabit Ethernet and fiber channel are currently major markets for VCSELs [5, 6].

On the other hand, **long**-wavelength VCSELs are typically based on alloys of GaAsP and InGaAsP lattice matched with InP substrates, due to their small band gap, and are exploited for optical trace-gas sensing (e.g. H_2O - 1800 nm, CO_2 - 2 μm, CH_4 - 1680 nm) as well, illumination and bio-medical (such as eye-safe laser beams) applications, but their main purpose are the long range optical communication systems

(10-20 km). The latter require 1310-1550 nm of wavelength, particularly interesting for single-mode fiber metropolitan area and wide area networks, since they may drive significant cost reduction in high-speed links of over several kilometers with single-mode fibers [5]. However, ternary and quaternary alloys used for these applications show some really poor properties with respect to the simpler AlGaAs systems used for short-wavelength emission. In particular, InP-based materials are characterized by low thermal conductivity due to alloy scattering of phonons, especially when carriers (holes) are injected from p-doped contacts. This is actually the main reason for which buried tunnel junctions have been introduced in first place in such devices. Furthermore, their refractive index modulation is limited: this increases the amount of number of layers needed to build a semiconductor DBR, which in turn increases the losses of the cavity leading to a larger power dissipation.

1.1.2 Structure

Let's now discuss about the general structure of a VCSEL. A scheme of a VCSEL including DBR only at one facet is reported in fig. 1.1, from which it is possible to recognize the main building blocks of such a laser. The vertical emission of coherent light in a VCSEL requires mirrors parallel to the junction plane. This is realized by exploiting DBR (distributed Bragg's reflectors) technique, which makes use of stacks of quarter-wavelength layers (to ensure photons constructive interference), making the cavity design more critical than in conventional edge-emitting lasers. In fact, VCSELs have very short cavity (with $L_{cav} = \lambda/n$, where λ is the lasing wavelength and n is the effective refractive index of the cavity material) which therefore needs larger gains and very low mirror losses (e.g. a reflectivity larger than 99.5%) to achieve lasing threshold [2]. In particular, DBRs are used to make the reflectivity depending on the wavelength and to make easier a fine tune of the frequency position of the lasing mode. The DBR mirror is composed by quarter-wavelength semiconductor or dielectric material layers with different refractive index. In the latter case, few pairs of isolating dielectric material with high-index contrast need to be epitaxially grown.

The active region, i.e. where the cavity gain is attained by radiative recombination from spontaneous emission, is made of a stack of quantum wells (MQWs) separated by barriers, whose material and size has to be properly engineered to match the geometry of the cavity: active region must be designed to have a lower band gap energy than the surrounding layers. This is due to the fact that in a VCSEL the emission wavelength mainly depends on the cavity design itself, rather than on QW characteristics, given the very short length of the former, which is on the order of magnitude of λ (while in a simple laser diode L_{cav} is few hundreds of microns). Electrons and holes are injected inside the active region through the DBR mirrors or by using intracavity contacts. In all the performed simulations, I assume to exploit the latter method: in fact, the former would have required highly doped DBR mirrors, which are not easily and efficiently

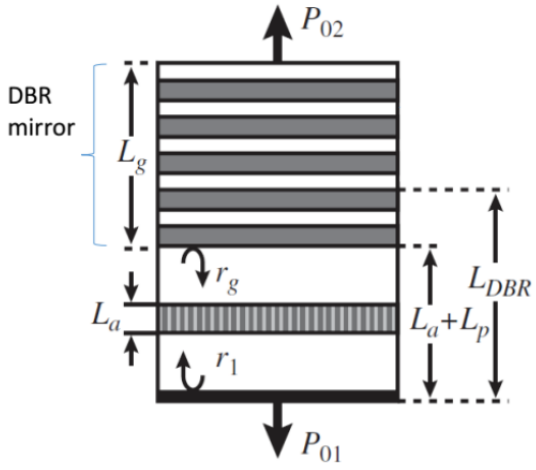


Figure 1.1: Scheme of a VCSEL with a DBR mirror on the top facet. L_a and L_g are the active region and the grating region lengths, respectively, while $L_a + L_p$ is equal to the cavity length L_{cav}; r_g is the grating reflectivity and r_1 is the bottom mirror reflectivity.

achieved.

On the other hand, intracavity contacts are far from ideal. Lateral confinement of both optical mode and electric current is needed to achieve optimal gain-mode overlap, thus providing low threshold current and high efficiency. Typically, electrical and optical confinement is obtained by incorporating buried oxide aperture together with top mesa. This leads to a low threshold and high power-conversion efficiency, over 50%. In fact, a poor lateral hole spreading into p-type contact makes the bare annular contact geometry inefficient for achieving current injection in the central part of the devices, since emission results to be only localized under the contact. Furthermore, state-of-the-art oxide aperture techniques show some major criticalities, especially in InGaAsP-based devices. Again, a solution to overcome these problems is the use of buried tunnel junctions, which are presented in the followings.

1.2 Buried tunnel junctions

As stated in the previous section, "standard" VCSEL structure shows some issues related to the use of p-doped layers as electric contacts and to oxide aperture realization technique. Let's start from the former. Since transport in acceptor-doped layers is inherently dominated by holes motion (which typically are much heavier than electrons), their thermal and electrical conductivity is significantly lower with respect to the donor-doped ones, due to the microscopic Ohmic law:

$$\sigma_{e,h} = \frac{1}{\rho_{e,h}} = \frac{n_{e,h} q^2 \lambda_{e,h}}{m^*_{e,h} v_F} \tag{1.1}$$

with $n_{e,h}$ electrons (holes) density, $\lambda_{e,h}$ electrons (holes) mean free path, v_F Fermi velocity and $m^*_{e,h}$ electrons (holes) effective mass, where typically $m^*_h \gg m^*_e$. As a consequence, **resistivity** of p-doped layers is much greater than n-doped ones ($\rho^*_h \gg \rho^*_e$). The straightforward consequence is a larger amount of dissipated electrical power ($P = RI^2$) at the p-side of the structure, and thus to a greater **self-heating process**, eventually limiting the maximum operating temperature of the device. Furthermore, low ohmic resistance R contacts result to be crucial also for **high speed modulation**, since it directly acts on the parasitic time constant $\tau = RC$ of the laser.

Eventually, the other reason that pushes to the use of BTJs is related to **standard oxide aperture** technique. Even though it is proved to be a very good solution to improve the efficiency of VCSELs by properly confining optical mode and current, it may also lead to lateral loss issues when the aperture, put in the proximity of the active region, reaches very small diameters [7]. The inclusion of buried tunnel junctions in the structure helps to solve such a problem [8]. In addition to this issue, the practical realization of an oxide aperture requires a wet oxidation process of Al compound (in order to obtain Al_xO_y layers), and this is not easily obtained in InP-based systems ([9, 10]), such as the long-wavelength VCSEL presented in section 4.1. For all these reasons, alternative carrier injection methods should be explored, in order to keep the device resistance as low as possible and avoid also the reduction of the wall-plug efficiency of the VCSEL.

Let's recall that a VCSEL is a device that produces both an electrical and an optical output during its working operations. The latter is highly dependent from **temperature** of the device, and this dependence is twofold. In first place, the emitted light wavelength depends on the round trip (RT), which should fulfil the first Barkhausen condition inside its optical cavity: this is reached when the optical gain in the cavity is great enough to compensate for the losses the field experiences during a round trip:

$$|RT| = \left| r_1 r_g e^{-2j\beta L_{cav}} e^{(\Gamma g - <\alpha_i>)L_{cav}} \right| = 1 \tag{1.2}$$

where $r_{1,g}$ are the reflectivity of the DBRs on both sides of the VCSELs, Γg is the gain of the cavity and L_{cav} is the cavity length. Moreover, the emitted light resonant frequencies are given by [2, p. 289]:

$$f_n = n\frac{c_0}{2n_{eff}L_{cav}} \tag{1.3}$$

where n is the mode order, c_0 is the speed of light in vacuum and n_{eff} is the effective **refractive index** of the cavity. The latter is modified by temperature (due to band gap modification), according to the linear law

$$\Delta n_T = n_0 \cdot \alpha_T (T - T_0) \tag{1.4}$$

where $\alpha_T > 0$ is the thermo-optic coefficient. Notice that the refractive index also plays a role into the material losses $< \alpha_i >$. As a consequence, the cavity modes undergo a

shift and output power is reduced as the temperature rises. Eventually, an increase in temperature leads to a **variation of the threshold** current. This is due to the fact that high temperatures favour non-optical recombination processes, meaning that gain decreases and losses are stronger: higher currents are required, and dissipated power rises accordingly [2, p. 320].

The lasing condition is directly related to the round trip, and is reached for:

$$\Gamma g_{th} = \frac{1}{L_{cav}} \log\left(\frac{1}{r_1 r_g}\right) + < \alpha_i > \tag{1.5}$$

with Γg_{th} threshold gain of the cavity. As a consequence a VCSEL requires a lasing cavity mode wavelength that is red-shifted with respect to that of the active material (QWs) gain peak at room temperature. In fact, under current injection, internal temperature increases (and this is particularly true in case of lossy p-contacts): due to (1.5) the threshold condition is reached for a lower wavelength [11].

To solve or mitigate part of these problems, buried tunnel junctions (BTJs) represent a good solution, even though their use have been mainly demonstrated on light emitting diodes (LEDs, [12–14]) and they may generate some difficulties in reaching lasing condition. As a matter of fact, BTJs have some problems related to a non-uniform current injection, due to local composition and doping inhomogeneities. However, the crucial advantage that they provide is that n-contacts can be used instead of p-doped ones if a **reverse-biased** tunnel junction is introduced, such that p-doped layers are located only near the active region, thus reducing optical absorption. A BTJ is a degenerately doped pn-junction which enable band-to-band tunneling for strong electric fields ($> 10^6$ V/cm), first discovered by Esaki in 1958 [15]. A schematic of a tunnel junction band diagram is reported in fig. 1.2, where valence and conduction bands are overlooking in a reverse bias condition. Furthermore, this solution allows to work with a much more laterally uniform injection current (due to the large electron mobility), which simplifies current [9] and light confinement [16] to any shape and size, by means of state-of-the-art photolithography techniques.

Let's remark that if the TJ is grown across the whole cross section of the structure, a current blocking layer would be required, and this is sometimes very difficult to realize. The simplest solution to this is the removal of the tunnel junction outside the active region and overgrow a lightly doped n-doped layer (namely, a mesa), such that blocking is naturally realized by a n/p-junction beneath the active region itself. Such a pattering is useful because it allows to obtain both a current confining layer and optical guiding (due to refractive index step), and to realize without further steps large arrays of VCSELs. Moreover, low series resistance is obtained because only n-contacts and n-layers are used for lateral current flow [9]. Another procedure to place the TJ only in the required region is briefly described in [3, 11], where tunnel junctions are patterned on the VCSELs array before they are overgrown; then, a double wafer fusion

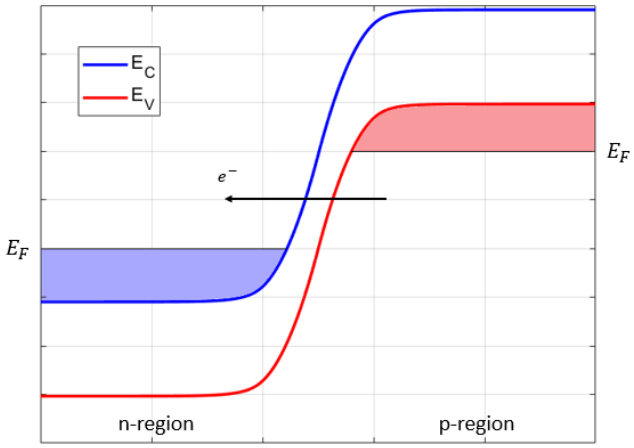


Figure 1.2: Schematic of a reverse-biased tunnel junction. E_c and E_v denotes the conduction and the valence band edges, while E_F denotes the Fermi energy. Only electrons tunneling is sketched here.

of the DBRs is performed. In such a way, the position of the tunnel junctions actually defines the pattern of the VCSELs array.

CHAPTER 2

Modelling and Simulation

As mentioned at the beginning of the introduction, in this chapter the used physical and mathematical models are described. This initial presentation of the physical approximations is followed by the descriptions of how they are actually implemented in the simulators.

Let's start from the fact that optoelectronic devices such as VCSELs need to be simulated from both temporal and three-dimensional (3D) spatial dynamics standpoint. This is realized by coupling a full-wave optical solver to a semiclassical carrier transport model derived from the Boltzmann transport equation (BTE) [17, p. 35], generalized for carrier transport in semiconductors, which deals with the probability time evolution:

$$\frac{\partial f(\mathbf{k},\mathbf{r},t)}{\partial t} + v(\mathbf{k})\nabla_r f(\mathbf{k},\mathbf{r},t) + \frac{F}{\hbar}\nabla_k f(\mathbf{k},\mathbf{r},t) = \frac{df(\mathbf{k},\mathbf{r},t)}{dt}\Big|_{coll} \tag{2.1}$$

where $f(\mathbf{k},\mathbf{r},t)$ is the carrier distribution function describing the evolution in the phase space of carriers treated as point-like particles with well-defined geometrical position $\mathbf{r}$ and wave vector $\mathbf{k}$ (as they were classical particles); $v(\mathbf{k}) = \nabla_k E(\mathbf{k})/\hbar$ is the carrier group velocity (which includes information about the electronic band structure) and $F = \hbar dk/dt$ is the electric field force. The right hand side (RHS) of (2.1) takes into account the scattering phenomenon the carriers may undergo during transport:

$$\frac{df(\mathbf{k},\mathbf{r},t)}{dt}\Big|_{coll} = -\frac{|f(\mathbf{k},\mathbf{r},t) - f_0(\mathbf{k})|}{\tau(\mathbf{k})} \tag{2.2}$$

in which $f_0(\mathbf{k})$ is the distribution function at equilibrium and $\tau(\mathbf{k})$ is the average lifetime due to scattering processes assuming relaxation-time approximation, that takes into account the rate at which a particle scatters from an initial to a final state.

Since (2.1) is an integro-differential equation, nonlinear because the collision term depends on the distribution function, a direct statistical solution is provided by the Monte Carlo methods. This consists in simulating the motion of a carrier ensemble, subject to the action of electric and magnetic fields and of given scattering mechanisms.

In this scheme, the states of the carriers after a scattering process are determined stochastically according to the scattering probabilities themselves. For this reason, this is the most general way to solve the BTE, and it allows to include sophisticated physical models. Unfortunately, such an approach has several drawbacks. First of all, it is not computationally efficient for a self-consistent implementation, i.e. when Poisson's equation is solved self-consistently with the charge distribution, so that different time constants are involved in the calculations (this is the case of bipolar devices). Furthermore, Monte Carlo method results to be inaccurate in treating minority carriers: as a consequence, usually it is used only for non-self-consistent approximations or to unipolar devices [17, pp. 37, 38].

In order to overcome such limitations and simulate more complex systems, alternative paths can be followed to reach acceptable solutions of BTE. A simple one assumes homogeneous conditions and small electric fields, such that closed-form approximate solutions of (2.1) can be obtained by linearizing it with respect to the field itself (relaxation-time approximation). Another possible way to solve the Boltzmann transport equation is the expansion of the unknown $f(\mathbf{k}, \mathbf{r}, t)$ into spherical harmonics and then solving the obtained equation for the expansion coefficients (expansion method).

However, the most common approach to a transport problem is the derivation of semiclassical models directly from the BTE, starting from reasonable assumptions and approximations: hydrodynamic, energy balance and drift-diffusion models. In the next section a focus on the last one is presented.

2.1 Drift-Diffusion model

In this section are presented some details of the model actually used in all the simulations. Since we are dealing with devices that are small with respect to the operating wavelength, Maxwell's equations can be simplified, such that a simpler model (with respect to BTE) is derived. This is called drift-diffusion model, whose constitutive equations are the following [18]:

$$-\nabla_r^2 \phi = \frac{q}{\epsilon}[p - n + N_D^+ - N_A^-] \tag{2.3a}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\nabla_r \underline{J}_n + U_n = 0 \tag{2.3b}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\nabla_r \underline{J}_p + U_p = 0 \tag{2.3c}$$

where (2.3a) is the Poisson's equation: $\epsilon = \epsilon_0 \epsilon_r$ is the dielectric permittivity, ϕ is the electrostatic potential, p and n are the hole and the electron densities, respectively; eventually, N_A^+ and N_D^- are the ionized acceptor and donor densities (a focus on this is provided in section 2.1.1). On the other hand, (2.3b) and (2.3c) are the electron and hole continuity equations, where $\underline{J}_{n,p}$ is the electron (hole) current density and

$U_{n,p} = G_{n,p} - R_{n,p}$ is the net electron (hole) recombination rate. Notice that all the quantities in (2.3) explicitly depend on space and time, e.g. $n = n(\mathbf{r}, t)$ and $p = p(\mathbf{r}, t)$. In turn, the current densities are made of two contributions (i.e., drift + diffusion), derived from BTE [18]:

$$\underline{J}_n = -qn\mu_n \nabla_r \phi + qD_n \nabla_r n \tag{2.4a}$$

$$\underline{J}_p = -qn\mu_p \nabla_r \phi - qD_p \nabla_r p \tag{2.4b}$$

where electrons (holes) diffusivity is given by a generalized Einstein relation:

$$D_{n,p}(\mathcal{E}) = \mu_{n,p}(\mathcal{E}) \frac{k_B T}{q} \tag{2.5}$$

that depends on electric field $\mathcal{E}$ through the electrons (holes) mobility $\mu_{n,p}$.

Notice that at equilibrium condition, both (2.4a) and (2.4b) are equal to 0, such that the system (2.3) reduces to Poisson's equation (2.3a), since continuity equations provide no additional information. However, let's remark that the equilibrium condition simulation is a crucial step for a correct implementation of the DD model, because it provides a proper initial guess needed by the self-consistent loop to correctly converge far from equilibrium.

2.1.1 Charge density

As mentioned, some clarifications are needed for the RHS of (2.3a), because it is a general form of the charge density:

$$\rho = q[p - n + N_D^+ - N_A^-] \tag{2.6}$$

that includes the carrier densities n and p, and the ionized doping charges N_D^+ and N_A^-. Generally, the former are described by the **Fermi** distribution:

$$n = N_C \mathcal{F}_{\frac{1}{2}} \left(\frac{E_F - E_C}{k_B T} \right) \tag{2.7a}$$

$$p = N_V \mathcal{F}_{\frac{1}{2}} \left(\frac{E_V - E_F}{k_B T} \right) \tag{2.7b}$$

where $\mathcal{F}_{\frac{1}{2}}$ is the integral or order 1/2 of the Fermi-Dirac statistics (or simply "Fermi function") ([19, 20]) and $N_{C,V}$ effective densities of states (DOS).

However, in case of non-degenerate semiconductor ($n < N_C$, $p < N_V$), an approximation of the Fermi statistics can be used, the so called **Boltzmann** statistics, which is much simpler and does not require the computation of any integral:

$$n \simeq N_C \exp \left(\frac{E_F - E_C}{k_B T} \right) \tag{2.8a}$$

$$p \simeq N_V \exp \left(\frac{E_V - E_F}{k_B T} \right) \tag{2.8b}$$

Furthermore, such a distribution obeys to the *law of mass action*, stating that the product of the carrier concentrations at thermodynamic equilibrium is constant:

$$n_i^2 = np = N_C N_V \exp\left(-\frac{E_g}{k_B T}\right) \tag{2.9}$$

which also implies that the product of concentrations only depends on material and temperature.

Beside the two distribution models presented in (2.7) and (2.8), some additional attention must be paid for the evaluation of the ionized doping densities. A first, simple approximation takes $N_D^+ = N_D$ and $N_A^- = N_A$, thus assuming a **full** ionization of doping impurities. On the other hand, a more realistic simulation must take into account that not all the dopant species are activated after implantation/diffusion inside another semiconductor: **incomplete** ionization of dopants. In fact, a dopant results to be activated only in case it generates a *substitutional* impurity, and such a phenomenon actually does not occur for every atom. To consider the effect of this, a model including donor and acceptor activation energies should be used:

$$N_D^+ = \frac{N_D}{1 + \frac{2}{n_1} n} \tag{2.10a}$$

$$N_A^+ = \frac{N_A}{1 + \frac{4}{p_1} p} \tag{2.10b}$$

where parameters n_1 and p_1 are given by

$$n_1 = N_C \exp\left(-\frac{\Delta E_D}{k_B T}\right) \tag{2.11a}$$

$$p_1 = N_V \exp\left(-\frac{\Delta E_A}{k_B T}\right) \tag{2.11b}$$

in which ΔE_D and ΔE_A are the donor and acceptor's activation energies (few meV), respectively, and are material and dopant dependent.

2.1.2 Generation-Recombination models

A further approximation involves the simplified collision term of the DD model, included in the carriers continuity equations, i.e., the net recombination rates $U_{n,p} = G_{n,p} - R_{n,p}$. In particular, the generation term $G_{n,p}$ takes into account the phenomenon during which an *e-h* pair is generated: this happens when an electron (hole) is promoted from the valence (conduction) band to the conduction (valence) band, starting to play a role in the transport mechanisms. On the other hand, recombination process is described by the term $R_{n,p}$, and is simply the reverse process with respect to generation: an electron (hole) in the conduction (valence) band fills an empty state in the

valence (conduction) band. As a consequence, the *e-h* pair is annihilated and energy is released in the form of phonons or photons. Both these process can be phonon-assisted (thermal), photon-assisted (optical) or assisted by other electrons or holes (Auger). Furthermore, GR transitions can be either interband (direct) or assisted by intermediate trap levels in the forbidden band (indirect or Shockley-Read-Hall mechanisms). Let's remark that in stationary conditions $U_n = U_p = U$, while the same does always not hold true in case of time-varying conditions (e.g. for GR occurring through intermediate traps or recombination centers acting as reservoirs). The main GR mechanisms that take place in bulk regions of the device are spontaneous emission, Auger, and Shockley-Read-Hall, each of them characterized by a specific rate: U^{sp}, U^{Aug}, and U^{SRH} [17, pp. 48, 49].

Spontaneous emission

Spontaneous emission is the radiative recombination actually exploited to generate photons in a semiconductor. In a VCSEL, the quantum wells in the active region are designed to obtain a high rate of spontaneous emission: electrons and holes are injected in the QWs where they radiatively recombine, transferring their energy to a photon, whose energy depends on the energy band diagram of the quantum well itself. In this way, gain is obtained and lasing condition is reached. Spontaneous emission rate is modelled by the following:

$$U^{sp} = B_{rad}(np - n_i^2) \tag{2.12}$$

where B_{rad} in (2.12) is the radiative recombination coefficient (cm^3/s) and n_i is the intrinsic carrier concentration of the semiconductor; n and p are the usual carrier densities. Of course, at equilibrium conditions U^{sp} is equal to zero, due to (2.9).

Auger recombination

Radiative recombination has an unwanted competitor when a device works in high-injection condition: the Auger recombination. This is an electron or hole-assisted (three particles) process which involves an *e-h* pair and an additional electron or hole. The expression describing it is:

$$U^{Aug} = (C_n n + C_p p)(np - n_i^2) \tag{2.13}$$

where C_n and C_p are the Auger coefficients for electrons and holes (cm^6/s). It is worth noticing that (2.13) takes into account the third particle by including two coefficients which in turn are multiplied by the populations ($\propto n^2p$ or np^2). This because the proportionality of the Auger recombination rate on the carrier densities is related to both their collisions and their energies, since they supply some energy to the process. Notice that the inverse process of the Auger recombination is the *generation by impact ionization*, and is not taken into account in my simulations.

Shocley-Read-Hall recombination

As mentioned, SRH recombination deals with the trap-assisted processes. These are possible when a semiconductor has some traps introducing, in the forbidden band, a discrete energy level. The latter increases the probability of a interband transition, because the energetic "jump" a carrier has to face becomes smaller. The net trap-assisted recombination rate can be expressed as follows:

$$U^{SRH} = \frac{np - n_i^2}{\tau_p(n + n_t) + \tau_n(p + p_t)} \tag{2.14}$$

where $n_t = n_i \exp(E_{trap}/k_B T)$, $p_t = n_i \exp(-E_{trap}/k_B T)$ and τ_p, τ_n in (2.14) are the trap level population densities (E_{trap} is the difference in energy between the trap level and the intrinsic level; particularly important are the *recombination centers*, i.e., midgap traps, that maximize the SRH rate) and the average carrier lifetimes, respectively. Lifetimes are doping level independent and of course they are smaller for larger trap densities.

Eventually, band-to-band tunneling (BTBT) is presented in section 3.1, since it requires a deeper description.

2.2 D1ANA implementation

The drift-diffusion model presented in section 2.1 represents a huge simplification with respect to BTE, but it hides some difficulties. In fact, it consists of a system of partial differential equations that needs to be solved in a self-consistent fashion in order to reach meaningful results. The problem is that it is not possible to directly solve this kind of mathematical problems by means of a software. The only way to perform some calculations on (2.3) requires a spatial discretization (by means of a proper mesh grid) together with an iterative solving method (e.g., the Newton's scheme), able to solve numerically all the involved equations. This allows to reduce the complexity of the problem, because the unknowns quantities are computed only at the node of the spatial grid (*finite box* discretization). As a consequence, the mathematical problem becomes a simpler system of linear equations, which is easily manageable by standard commercial softwares.

2.2.1 Spatial discretization

An important remark is needed before starting: D1ANA takes into account only a 1D DD model, so all the space gradients reduce to simple 1D space derivatives (z-direction has been chosen as the transport one), such that (2.3a) can be written as follows:

$$\frac{d}{dz}\Big[\epsilon(z)\frac{d}{dz}\phi(z)\Big] = -q[p(z) - n(z) + N_D^+(z) - N_A^-(z)] \tag{2.15}$$

This is done in order to further reduce the complexity of the problem, allowing for faster simulations. Such an approximation is allowed because we are assuming a symmetry across the transverse directions to the transport one, meaning that the latter is the only one that really matters when the drift-diffusion model is used, while some quantities such as the net generations/recombination rates and the carriers and current densities can be obtained with simple volume integrations.

According to the mentioned finite box discretization, after having defined a spatial 1D mesh, a system of linear equations is obtained by simply integrating the differential equation on a control space box of each region. As a consequence, nn equations are obtained, with nn equal to the number of nodes used to discretize the 1D domain. All the quantities related to each elements are computed by means of a trapezoidal rule for numerical integration, which represents a good approximation for the functions we are dealing with.

Let's introduce a bit of notation:

l_e : length of the l^{th} element. Assuming that element e is identified by node e and $e+1$:

$$l_e = x_{e+1} - x_e \tag{2.16}$$

ne : is the number of elements of the mesh (in 1D, $ne = nn - 1$)

L_i : length of the control box, with i the index identifying a node. In D1ANA, this is defined as:

$$L_i = z_{i+\frac{1}{2}} - z_{i-\frac{1}{2}} \tag{2.17}$$

For $i = 1$, $L_i = z_{i+\frac{1}{2}}$, while for $i = nn$, $L_i = z_{nn} - z_{nn-\frac{1}{2}}$. Notice that L_i is exactly the length used to performed the already mentioned trapezoidal rule

Now, by integrating (2.15) on the control boxes for each i^{th} node:

$$\int_{z_{i-1/2}}^{z_{i+1/2}} \frac{d}{dz}\left[\epsilon(z)\frac{d}{dz}\phi(z)\right] dz = -q\int_{z_{i-1/2}}^{z_{i+1/2}} \left[p(z) - n(z) + N_D^+(z) - N_A^-(z)\right] dz \tag{2.18}$$

with $i = 1, ..., N$. By applying FD/FEM/FB to the LHS, we obtain:

$$\epsilon_{i+\frac{1}{2}}\frac{\phi_{i+1} - \phi_i}{z_{i+1} - z_i} + \epsilon_{i-\frac{1}{2}}\frac{\phi_{i-1} - \phi_i}{z_{i-1} - z_i} = -q\left[p_i - n_i + N_{D,i}^+ - N_{A,i}^-\right] \tag{2.19}$$

where $x_i \triangleq x(z_i)$, being x a generic node quantity. Eq. (2.19) can be very easily implemented in MATLAB assembling matrices in a node-wise fashion, by using loops running on all the nodes. Even more manageable would be an element-wise implementation. In order to realize it, one needs to rewrite (2.19) by considering two cases concerning each i^{th} node:

- $i = 1e$. In this case, the term:

$$\epsilon_{i+\frac{1}{2}} \frac{\phi_{i+1} - \phi_i}{z_{i+1} - z_i} \rightarrow \epsilon_e \frac{\phi_{2e} - \phi_{1e}}{l_e} = -\frac{\epsilon_e}{l_e}\phi_{1e} + \frac{\epsilon_e}{l_e}\phi_{2e} = \begin{bmatrix} M_{11} & M_{12} \end{bmatrix} \begin{bmatrix} \phi_{1e} \\ \phi_{2e} \end{bmatrix} \tag{2.20}$$

  with $M_{11,e} = -\epsilon_e/l_e$ and $M_{12,e} = \epsilon_e/l_e$

- $i = 2e$. In this case, the term:

$$\epsilon_{i-\frac{1}{2}} \frac{\phi_{i-1} - \phi_i}{z_{i-1} - z_i} \rightarrow \epsilon_e \frac{\phi_{1e} - \phi_{2e}}{l_e} = \frac{\epsilon_e}{l_e}\phi_{1e} - \frac{\epsilon_e}{l_e}\phi_{2e} = \begin{bmatrix} M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \phi_{1e} \\ \phi_{2e} \end{bmatrix} \tag{2.21}$$

  with $M_{21,e} = \epsilon_e/l_e$ and $M_{22,e} = -\epsilon_e/l_e$

Implying that LHS of (2.19)=(2.20)+(2.21). Consequently, the matrix notation:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{2.22}$$

indicates simply how to assemble the final matrix using an element-wise notation.

Similarly, the RHS of (2.19) can be written by means of such a notation. To this purpose, the definition of L_i in (2.17) must be reconsidered. In particular, we need to move to the element quantity $L_e^{(n)}$, defined as the contribution of the e^{th} element on its n^{th} node ($n = 1$ indicates the node on the left, $n = 2$ the one on the right). Eventually, the RHS is assembled.

The same procedure needs to be applied to the continuity equations (2.3b) and (2.3c). Let's start from the stationary condition ($\partial n, p/\partial t = 0$). Considering again a 1D condition and integrating similarly to what has been done in (2.18), we obtain, for the electrons (the same holds for holes):

$$\int_{z_{i-1/2}}^{z_{i+1/2}} \frac{\partial J_n}{\partial z} dz = q \int_{z_{i-1/2}}^{z_{i+1/2}} U_n dz \tag{2.23}$$

with RHS $\approx qU_{n,i}L_i$. On the other hand, the LHS requires non-trivial steps. First of all, notice that continuity equations belong to a system of three partial differential equations, with formally three unknowns (electrostatic potential ϕ, electron density n, and hole density p). Now, it is useful to rewrite equations (2.3b) and (2.3c) by introducing the *quasi*-Fermi levels (E_{fn} and E_{Fp}) as unknowns. This can be realized by exploiting the constitutive relationships:

$$\underline{J}_n = \mu_n n \nabla E_{Fn} \tag{2.24a}$$

$$\underline{J}_p = \mu_p p \nabla E_{Fp} \tag{2.24b}$$

which are formally equivalent to (2.4a) and (2.4b), respectively. In fact, recalling that

$$n = N_C \exp\left(\frac{E_c - E_{Fn}}{k_B T}\right) \Rightarrow E_{Fn} = E_c - k_B T \ln\left(\frac{n}{N_C}\right) \tag{2.25}$$

one could compute the gradient of E_{Fn} and come back to (2.4a). A similar procedure could be applied to holes case. The use of electrons and holes quasi-Fermi levels (also referred to as IMREF) is justified by the fact that even though the *new* formulation is formally identical to the *old* one, numerically the use of IMREFs could result in a more efficient solution.

Eventually, the discretization of (2.3) returns a matrix problem with the form:

$$\begin{bmatrix} \underline{\underline{A}}_{\phi,\phi} & \underline{\underline{A}}_{\phi,n} & \underline{\underline{A}}_{\phi,p} \\ \underline{\underline{A}}_{n,\phi} & \underline{\underline{A}}_{n,n} & \underline{\underline{A}}_{n,p} \\ \underline{\underline{A}}_{p,\phi} & \underline{\underline{A}}_{p,n} & \underline{\underline{A}}_{p,p} \end{bmatrix} \begin{bmatrix} \underline{\phi} \\ \underline{n} \\ \underline{p} \end{bmatrix} = \begin{bmatrix} \underline{\rho} \\ \underline{U}_n \\ \underline{U}_p \end{bmatrix} \tag{2.26}$$

The matrix A in the LHS of (2.26) has a dimension equal to $3 \times nn$, recalling that nn is the number of nodes. In fact, it has as many rows as the number of equations (nn) and as many columns as the number of unknowns (3). The vector of the $3 \times nn$ unknowns is called $\underline{u}$ (the column vector at the left hand side of the system above), while the vector containing the $3 \times nn$ known terms is called $\underline{t}$ (the column vector at the right hand side).

Furthermore, notice that:

- $\underline{\underline{A}}_{\phi,\phi}$ is exactly the matrix obtained by assembling (2.22)
- $\underline{\underline{A}}_{\phi,n}$ is a diagonal matrix containing qL_i as i^{th} element
- $\underline{\underline{A}}_{\phi,p}$ is a diagonal matrix containing $-qL_i$ as i^{th} element

ending up by writing $\underline{\underline{A}}\,\underline{u} + \underline{t} = 0$. Furthermore, the unknowns variables are considered as **independent** (even though they are of course coupled), meaning that:

$$\frac{\partial \phi}{\partial n} = 0, \qquad \frac{\partial n}{\partial \phi} = 0, \qquad \frac{\partial p}{\partial \phi} = 0, \quad ...$$

At this point, it is possible to write down the whole system matrix into two parts:

`Lmat` ($\underline{\underline{L}}$): containing the *linear* part of the system, i.e., the parts related to the discretized differential equations:

```
Lmat*uvet + tvet = fvet
```
(2.27)

where `fvet` is the "function" set to be equal to 0 inside the Newton's method

`Nmat` ($\underline{\underline{N}}$): some parts of the Jacobian are already inserted in `Lmat`: the linear terms, i.e., all those parts which, when directly multiplied by the unknowns, return `fvet` . On the other hand, the derivatives of the non-linear components such as recombinations rates or all those contributions not entering in the computation of `fvet` are inserted in `Nmat`

Then, `Jmat = Lmat + Nmat`. The convergence of the Newton's scheme is reached when the vector containing the unknown terms changes of a quantity smaller than a control one. The variation of `uvet` between a step and the previous one is simply given by:

```
delta_u = -Jmat\fvet
```
(2.28)

Returning now to the LHS of (2.23):

$$\int_{z_{i-1/2}}^{z_{i+1/2}} \frac{\partial J_n}{\partial z} dz = J_{n,i+\frac{1}{2}} - J_{n,i-\frac{1}{2}} \tag{2.29}$$

and applying finite difference (FD) to (2.4a), we get:

$$J_{n,i+\frac{1}{2}} \approx -q\frac{n_i + n_{i+1}}{2}\mu_n \frac{\phi_{i+1} - \phi_i}{z_{i+1} - z_i} + qD_n \frac{n_{i+1} - n_i}{z_{i+1} - z_i},$$

useful to assemble the case $i = 1e$ in both $\underline{\underline{L}}$ and $\underline{\underline{N}}$; and

$$J_{n,i-\frac{1}{2}} \approx -q\frac{n_{i-1} + n_i}{2}\mu_n \frac{\phi_i - \phi_{i-1}}{z_i - z_{i-1}} + qD_n \frac{n_i - n_{i-1}}{z_i - z_{i-1}},$$

useful to assemble the case $i = 2e$, in $\underline{\underline{L}}$, $\underline{\underline{N}}$; similarly to what is done in (2.19) when matrix $\underline{\underline{M}}$ is computed with an element-wise notation.

2.2.2 Stability issues

The formulation just presented results to be rather unstable when implemented, in particular when the number of nodes nn is not high enough. In fact, it is common to observe lack of convergence of Newton's scheme or oscillating quantities where smooth curves should be obtained instead.

The problems of such a formulation come from the discretization used for the electrostatic potential and the carrier densities. In fact, by using eq. (2.7a) or even the simpler (2.8a), the dependence of the density on ϕ is exponential. Moreover, some issues may arise from the fact that (2.4a) and (2.4b) are made up of two contributions, namely drift and diffusion. These are two competing physical phenomena, which are likely to lead to convergence problems [21]. In order to solve the stability issues, some manipulations to the current density expressions are needed. Let's focus on the electron case, and rewrite its equation as follows:

$$\frac{1}{qD_n} J_n = -\frac{n}{V_T}\frac{\partial \phi}{\partial z} + \frac{\partial n}{\partial z} \tag{2.30}$$

in which is possible to discretize ϕ by using finite differences. In fact, the electrostatic potential varies rather mildly through the 1D domain so it becomes reasonable to approximate (2.30) by means of a linear interpolation (FD):

$$\frac{1}{qD_n}J_{n;i,j} \approx -\frac{n}{V_T}\frac{\phi_j - \phi_i}{z_j - z_i} + \frac{\partial n}{\partial z} \tag{2.31}$$

where $J_{n;i,j}$ has the meaning of current between nodes i and j. Let's define $l = z_j - z_i$ as the distance between adjacent nodes. Now, taking a definition from the fluid dynamics, let:

$$\mathcal{R} = \frac{\phi_j - \phi_i}{V_T} \tag{2.32}$$

be the *Reynolds number* of the element delimited by nodes i and j. Such a number is a measure of how big is the variation of the electrostatic potential between two nodes with respect to the thermal voltage V_T. Now, it is necessary to understand how n varies in a cell, i.e., between consecutive nodes. To this purpose, let's impose the two conditions $z_i = 0$ and $z_j = l$, in order to simplify the problem:

$$\frac{\partial n}{\partial z} - \frac{\mathcal{R}}{l}n = \frac{J_{n;i,j}}{qD_n} \tag{2.33a}$$

$$n(0) = n_i, \qquad \text{B.C. at } z = 0 \tag{2.33b}$$

$$n(l) = n_j, \qquad \text{B.C. at } z = j \tag{2.33c}$$

If a proper mesh is used, then $J_{n;i,j}$ can be assumed as constant; of course, from (2.32) $\mathcal{R}$ is constant. Therefore, let K be:

$$K = \frac{J_{n;i,j}}{qD_n}$$

constant as well. It is straightforward to deduce that $J_{n;i,j}$ depends on both $n(0)$ and $n(l)$, and it could be interpreted as the "boost" that makes carrier density change from $n(0)$ and $n(l)$ across the element. Consequently, K can be considered as a degree of freedom to enforce the boundary conditions.

A solution of the system (2.33) for the electron density can be demonstrated to be:

$$n(z) = c_1 \exp\left(\mathcal{R}\frac{z}{l}\right) - A\frac{l}{\mathcal{R}} \tag{2.34}$$

Then, by enforcing the boundary conditions, one can write:

$$n(z) = \frac{n_j - n_i}{e^{\mathcal{R}} - 1}e^{\mathcal{R}\frac{z}{l}} - \frac{n_j - n_i e^{\mathcal{R}}}{e^{\mathcal{R}} - 1} \tag{2.35}$$

which, according to [22], can be rewritten as:

$$n(z) = n_j\frac{1 - e^{\mathcal{R}\frac{z}{l}}}{1 - e^{\mathcal{R}}} + n_i\frac{e^{\mathcal{R}\frac{z}{l}} - e^{\mathcal{R}}}{1 - e^{\mathcal{R}}}$$

By setting a function $g(z, \mathcal{R})$ as:

$$g(z, \mathcal{R}) \triangleq \frac{1 - e^{\mathcal{R}\frac{z}{l}}}{1 - e^{\mathcal{R}}}$$

it becomes possible to write more compactly (2.35):

$$n(z) = (1 - g(z, \mathcal{R}))n_i + g(z, \mathcal{R})n_j \tag{2.36}$$

which is exactly the expression we look for to investigate stability more into the details. In fact, from an analysis of the magnitude of $\mathcal{R}$, one may obtain three cases (also depicted in fig. 11.1, [21, p. 275]). In practice, for $\mathcal{R} \approx 0$, $n(z)$ varies linearly from node i to node j; if $\mathcal{R} \ll 0$, $n(z)$ is almost equal to n_j throughout the whole element; eventually, if $\mathcal{R} \gg 0$, $n(z)$ is mostly similar to n_i. Same reasoning may be applied to holes population, of course.

Such a result implies that a linear interpolation for the carrier densities is usually wrong, possibly leading to instabilities, unless $\mathcal{R} \approx 0$. Recalling that Reynolds number (2.32) depends on how large is $\phi_j - \phi_i$ with respect to V_T, by using a proper (very fine, nn large) mesh, linear interpolation assumption may hold true. Let's remark that while all the previous reasoning can be applied to 1D case, when moving to 2D and 3D domain, such instabilities are not solvable.

2.2.3 Scharfetter-Gummel method

Now we have a rough idea on how carrier concentrations change in a grid element, assuming a linear variation of the potential. Following this path, Scharfetter and Gummel in 1998 provided an alternative discretization scheme for the drift-diffusion constitutive relations (2.4), not affected by numerical instabilities. The idea, presented in [23], is reported in the followings.

Again, let's reformulate eq. (2.4a) as:

$$\frac{1}{q}J_n = -n\mu_n\frac{\partial\phi}{\partial z} + D_n\frac{\partial n}{\partial z}$$

Then, by writing:

$$\frac{\partial n}{\partial z} = \frac{\partial n}{\partial \phi}\frac{\partial \phi}{\partial z}$$

it is straightforward to write:

$$\frac{\partial\phi}{\partial z}\left[-n\mu_n + D_n\frac{\partial n}{\partial\phi}\right] = \frac{1}{q}J_n$$

Now, we know from the previous section that ϕ can be approximated with a linear interpolation between nodes i and j:

$$\frac{\partial\phi}{\partial z} \approx \frac{\phi_j - \phi_i}{z_j - z_i} \triangleq \frac{\Delta\phi}{l}$$

leading to:

$$\frac{1}{q}J_{n;i,j} \approx \frac{\Delta\phi}{l}\left[-n\mu_n + D_n\frac{\partial n}{\partial\phi}\right]$$

By dividing both members of the latter by $-\mu_n\Delta\phi/l$ and rewriting as it was a separated variables ODE:

$$n + \frac{1}{q}J_{n;i,j}\frac{l}{\mu_n\Delta\phi} = \frac{k_BT}{q}\frac{\partial n}{\partial\phi} \Rightarrow \frac{1}{n + \frac{1}{q}J_{n;i,j}\frac{l}{\mu_n\Delta\phi}}\frac{\partial n}{\partial\phi} = \frac{q}{k_BT} \tag{2.37}$$

Whose solution is obtained by integrating both members with respect to the electrostatic potential, leading to:

$$n + \frac{1}{q}J_{n;i,j}\frac{l}{\mu_n\Delta\phi} = \exp\left(\frac{q}{k_BT}\phi + c\right) \tag{2.38}$$

where c is the constant arising from the integration of the RHS of (2.37). By setting:

$$\alpha = \frac{1}{q}J_{n;i,j}\frac{l}{\mu_n\Delta\phi}, \qquad \beta = \frac{1}{V_T} = \frac{q}{k_BT}$$

it is possible to avoid the evaluation of c. In fact, it is straightforward writing:

$$n_j + \alpha = \exp(\beta\phi_j + c)$$

$$n_i + \alpha = \exp(\beta\phi_i + c)$$

which leads to:

$$\frac{n_j + \alpha}{n_i + \alpha} = \exp(\beta\Delta\phi) \Rightarrow n_j - n_i\exp(\beta\Delta\phi) = \alpha(\exp(\beta\Delta\phi) - 1).$$

After some manipulations:

$$\frac{1}{\beta\Delta\phi} = \left[n_j\frac{\beta\Delta\phi}{\exp(\beta\Delta\phi)} - n_i\frac{-\beta\Delta\phi}{\exp(-\beta\Delta\phi) - 1}\right] = \frac{1}{q}J_{n;i,j}\frac{l}{\mu_n\Delta\phi} \tag{2.40}$$

Then, by introducing the Bernoulli function $B(x)$, defined as:

$$B(x) = \frac{x}{\exp(x) - 1}, \tag{2.41}$$

thus obtaining:

$$\frac{1}{q}J_{n;i,j} = \frac{\mu_n}{l\beta}\left[n_jB\left(\frac{\phi_j - \phi_i}{V_T}\right) - n_iB\left(\frac{\phi_j - \phi_i}{V_T}\right)\right] \tag{2.42}$$

The latter can be applied to (2.29), which still holds true because only the definition of the constitutive relation of J_n has been manipulated, while the electrons continuity

equation has not been changed. Therefore, the expression for $J_{n,i+\frac{1}{2}} - J_{n,i-\frac{1}{2}}$ is modified as follows:

$$J_{n,i+\frac{1}{2}} - J_{n,i-\frac{1}{2}} = \underbrace{q\frac{D_n}{z_{i+1}-z_i}\left[n_{i+1}B\left(\frac{\phi_{i+1}-\phi_i}{V_T}\right) - n_iB\left(\frac{\phi_i-\phi_{i+1}}{V_T}\right)\right]}_{\text{useful for } i=1e} + \underbrace{-q\frac{D_n}{z_i-z_{i-1}}\left[n_iB\left(\frac{\phi_i-\phi_{i-1}}{V_T}\right) - n_{i-1}B\left(\frac{\phi_{i-1}-\phi_i}{V_T}\right)\right]}_{\text{useful for } i=2e} \quad (2.43)$$

In this way, it is possible to proceed and compute the matrix elements useful to reach the solution for the **electron** (and similarly for hole) current density:

- $\underline{\underline{L}}_{n,n}$: including the Jacobian of the linear part. More specifically, this term contains the derivative of the electron density with respect to electron density itself.
  - for $i = 1e$:

$$q\frac{D_n}{l}\left[n_{2e}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - n_{1e}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right] = \underbrace{-q\frac{D_n}{l}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)}_{M_{11}}n_{1e} + \underbrace{q\frac{D_n}{l}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right)}_{M_{12}}n_{2e} \quad (2.44)$$

  - for $i = 2e$:

$$\underbrace{-q\frac{D_n}{l}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right)}_{M_{22}}n_{2e} + \underbrace{q\frac{D_n}{l}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)}_{M_{21}}n_{1e} \quad (2.45)$$

- $\underline{\underline{N}}_{n,\phi}$: Jacobian for the non-linear part. In this element are inserted the derivatives of the electron current continuity equation with respect to the electrostatic potential ϕ. This is done starting from (2.29), and it could have been done also without introducing the Scharfetter-Gummel scheme. Basically, eq. (2.43) is Taylor expanded, and only the first order terms are inserted in $\underline{\underline{N}}_{n,\phi}$, while the 0^{th} order contribution is already included in $\underline{\underline{L}}_{n,n}$.
  - for $i = 1e$, the "function" results to be:

$$f_{1e} = q\frac{D_n}{l}\left[n_{2e}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - n_{1e}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

    such that

$$M_{11} = \frac{\partial f_{1e}}{\partial \phi_{1e}} = q\frac{D_n}{V_T l}\left[-n_{2e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - n_{1e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

$$M_{12} = \frac{\partial f_{1e}}{\partial \phi_{2e}} = q\frac{D_n}{V_T l}\left[n_{2e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) + n_{1e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

– for $i = 2e$, the "function" results to be:

$$f_{2e} = -q\frac{D_n}{l}\left[n_{2e}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - n_{1e}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

such that:

$$M_{21} = \frac{\partial f_{2e}}{\partial \phi_{1e}} = -q\frac{D_n}{V_T l}\left[-n_{2e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - n_{1e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

$$M_{22} = \frac{\partial f_{2e}}{\partial \phi_{2e}} = -q\frac{D_n}{V_T l}\left[n_{2e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) + n_{1e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

For what concerns the **hole** current continuity equation, the derivation of the equation is very similar to the electron one. Therefore, the matrix elements are computed in the exact same way:

- $\underline{\underline{L}}_{p,p}$:
  - for $i = 1e$:

$$q\frac{D_p}{l}\left[p_{1e}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - p_{2e}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right] = \underbrace{q\frac{D_p}{l}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)}_{M_{11}}p_{1e} + \underbrace{q\frac{D_p}{l}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right)}_{M_{12}}p_{2e} \quad (2.46)$$

  - for $i = 2e$:

$$\underbrace{-q\frac{D_p}{l}B\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right)}_{M_{21}}p_{1e} + \underbrace{q\frac{D_p}{l}B\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)}_{M_{22}}p_{2e} \quad (2.47)$$

- $\underline{\underline{N}}_{p,\phi}$:
  - for $i = 1e$:

$$M_{11} = \frac{\partial f_{1e}}{\partial \phi_{1e}} = q\frac{D_p}{V_T l}\left[-p_{1e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - p_{2e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

$$M_{12} = \frac{\partial f_{1e}}{\partial \phi_{2e}} = q\frac{D_p}{V_T l}\left[p_{1e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) + p_{2e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

  - for $i = 2e$:

$$M_{21} = \frac{\partial f_{2e}}{\partial \phi_{1e}} = -q\frac{D_p}{V_T l}\left[-p_{1e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) - p_{2e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

$$M_{22} = \frac{\partial f_{2e}}{\partial \phi_{2e}} = -q\frac{D_p}{V_T l}\left[p_{1e}B'\left(\frac{\phi_{2e}-\phi_{1e}}{V_T}\right) + p_{2e}B'\left(\frac{\phi_{1e}-\phi_{2e}}{V_T}\right)\right]$$

Notice that also the **recombination** terms introduce some contributions to the known vector `tvet`, because they enter the carriers continuity equations and to the Jacobian `Jmat`. In particular, in `tvet` the contribution of each rate is added to matrix $\underline{\underline{M}} = qRL_i$ (where R is the overall recombination rate at one node) and assembled in the electron and hole continuity equations. A very similar procedure is performed for the Jacobian, where derivatives with respect to electrons and holes are assembled to the continuity equations.

2.2.4 Material parameters and geometry

D1ANA code allows to easily insert new material parameters, by simply adding a section in the proper function. The **parameters** that need to be provided to the code are the following:

- $m^*_{n,p}$: bulk electrons and holes effective mass, m_0
- $\mu_{n,p}$: electron and hole mobility, $cm^2/V{\cdot}s$
- $N_{C,V}$: bulk conduction and valence density of states, cm^{-3}
- E_g: energy band gap, eV
- n_i: intrinsic carrier concentration, cm^{-3}
- χ: electron affinity, eV
- ϵ_r: relative dielectric constant
- $\Delta E_{A,D}$: acceptors and donors activation energies, used in the incomplete ionization model seen in (2.10) , eV
- B_{rad}: radiative recombination coefficient (see eq. (2.12)), cm^3/s
- C_{nnp}, C_{ppn}: electron and hole Auger coefficients (see eq. (2.13)), cm^6/s
- $\tau_{n,p}$: SRH electron and hole capture time (see eq. (2.14)), s
- E_{trap}: position in energy of the intermediate SRH trap level, eV

All these are associated to the corresponding node or element. The dependence of the carrier mobilities on the electric field is not taken into account for these simulations; furthermore, the effect of the polarization charges can be neglected, because no nitride is used in any of the investigated structures. For each material not already inserted in the archive of D1ANA, I've looked at the sections in NSM Archive [24].

Eventually, for what concerns the **geometry** of each layer composing the simulated structure, it is required to insert:

- length of the layer
- material the layer is made of
- doping type and level (a graded doping is also feasible), associated to each mesh node
- molar fractions for ternary and quaternary alloys, associated to each mesh node, needed to compute the alloy parameters by means of linear interpolations

2.2.5 Solver

After having generated the mesh and associated all the quantities to the corresponding nodes or elements, D1ANA starts with the simulation. The first step performed is aimed to solve the Poisson's equation at thermodynamic equilibrium condition. This is a crucial step, because it provides an accurate initial guess for the drift-diffusion solver, before starting with the voltage sweep. Simulation starts from the boundary conditions: **neutrality** is imposed at contacts. While analytical solutions are available if Boltzmann statistics and full ionization (PBFI) are assumed, when Fermi statistics and/or incomplete ionization of doping concentration (PBII, PFFI, PFII) are taken into account neutrality at contacts can be solved only numerically. To do it, in D1ANA Netwon's method is adopted. In particular, it begins from the analytical solution of PBFI at each point of the device mesh; then, if one of the other three cases is required, Newton's scheme is used only at the contacts mesh nodes, i.e., the first and the last one. At each Newton's loop, charge concentrations are computed starting from the potential. The derivatives are introduced inside the Jacobian matrix `Jmat`, such that the linear system `Jmat*delta_u = -rvet` is solved (where `delta_u` is the variation `uvet` at each step and is used to determine convergence, while `rvet` is the "residual vector").

Once thermodynamic equilibrium solution is computed, drift-diffusion far from equilibrium simulation starts. Again, this is based on a generalized Newton's method; the difference with respect to the previous case is that now electron and hole continuity equations are solved in addition to Poisson's, with an overall of $3 \times nn$ equations, in a self-consistent fashion. After taking the result from the equilibrium condition, the voltage sweep loop can start. At each step, all the needed matrix are assembled as described in the previous section. Again, optical generation model and quantum corrections are not included in the simulations, so are not described. Also here, convergence is reached once the norm of `delta_u` takes a value lower than a control value.

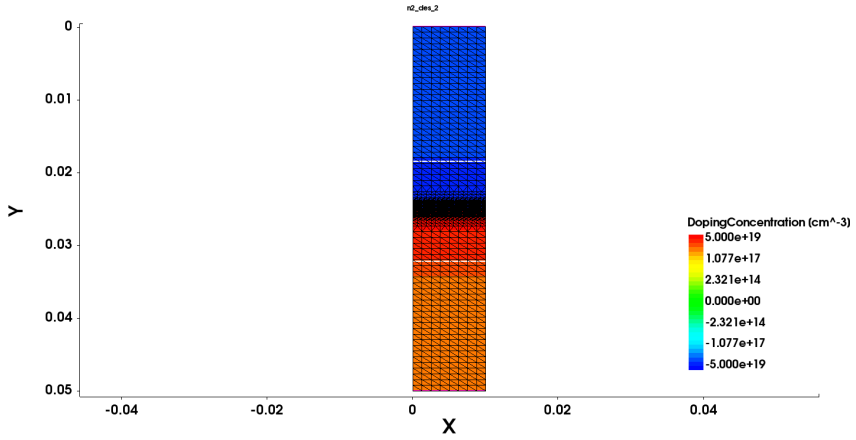
A final remark must be done: in order to properly converge, D1ANA exploits a **scaling** of some quantities during their computation: this is done for numerical reasons, but it does not affect units or order of magnitude of the final results.

2.3 Synopsys TCAD

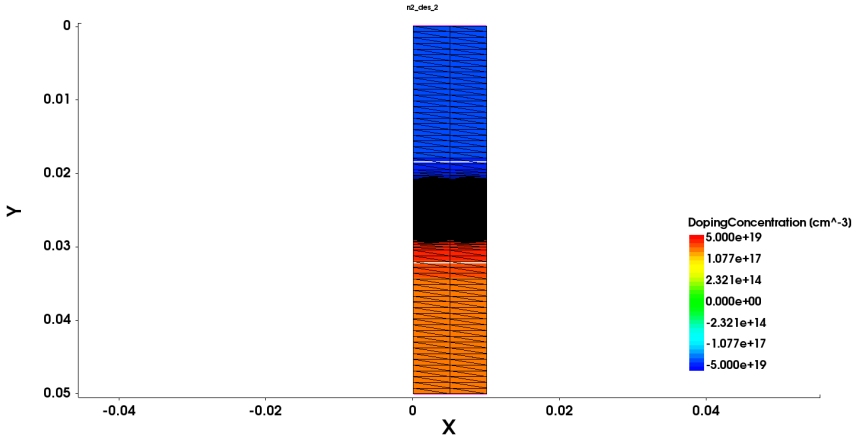
The core of all the simulations performed throughout the whole thesis work involves the use of D1ANA. Nonetheless, another simulator is used beside it: a software based on technology computer-aided design (TCAD) developed by Synopsys called *Sentaurus Device*. More into the details, it is a multidimensional device simulator that is used to simulate electrical characteristics of some compound semiconductor structures. Just like D1ANA, drift-diffusion model is the carrier transport model used by *Sentaurus Device*. By simply changing some options in the `physics` input files it is possible to make the software using the very same physical models of D1ANA, i.e., Fermi-Dirac statistics or Boltzmann approximation; full or incomplete ionization.

As described in several sections of chapter 4, such a software is used to test only the tunnel junctions included in the investigated VCSELs. In fact, *Sentaurus Device* is a very good benchmark for my implementation of the local BTBT model in D1ANA, for the reasons described at the end of section 3.1. Furthermore, also in this simulator some materials are not included, so that the generation of proper sections to fully describe them is a required step. In order to perform a fair comparison with respect to D1ANA, the very same parameters of the materials are used in the two simulators. Of course, also the physical models activated in the TCAD simulations are the very same already included in D1ANA, in order to avoid discrepancies not due to implementation errors which would be difficult to debug if different parameters were employed.

Considering that also *Sentaurus Device* has to work with a drift-diffusion model, it is worth spending some words to talk about the used mesh. First of all, let's remark that while D1ANA deals with a 1D spatial domain, in a TCAD this is not allowed. As a matter of fact, at least a 2D spatial grid must be defined when even a very simple structure like a tunnel junction is realized. Therefore, if the mesh is not properly designed, significant longer simulation times may occur with respect to a simpler 1D case, without obtaining improved results. In order to realize an efficient mesh, first one should recognize the direction along which the results are more important: in such a way, the grid across the other direction can be made more coarse, and the simulation time can be shortened. This is exactly what is done when working with *Sentaurus Device*. An example is reported in figures 2.1, where *pn* junction at equilibrium is shown. In fig. 2.1a, a fine mesh is adopted in both the directions. Nevertheless, along x direction this is not required, since none of the parameters change in that direction (material, doping, applied bias, ...). On the other hand, the junction is generated along y direction and the bias is applied between the top and the bottom contacts, so this is the only relevant direction. As a consequence, it is possible to reduce down to few points along x, as represented in fig. 2.1b. This allows to add points along y, thus improving the result of the interpolation on the points in that direction across the whole length. Or, similarly, it becomes possible to improve the mesh in the zone



(a) The mesh is fine in both x and y direction, with a small refinement region in the depletion region.



(b) A very coarse mesh is used along x, since the relevant results are all on y direction. This allows for a wider refinement region where is needed.

Figure 2.1: Examples of 2D mesh generated in *Sentaurus Device*.

around the depletion region, which is the most important one in a *pn* junction (this is what is actually done in the figure).

A little remark on the obtained results from the TCAD. As mentioned, pure 2D simulations are not possible, but all the quantities are extracted by performing a vertical cut in the mid point of the x direction, such that they are compatible between each other. These quasi-1D results are then exported to MATLAB where they are plotted. This makes easier a fast comparison with the results coming from D1ANA.

CHAPTER 3

Band-to-band tunneling

The buried tunnel junctions analysed during the thesis all work at equilibrium or in **reverse** bias condition: for the latter, the main transport mechanisms across the *pn* junction are represented by the direct interband band-to-band tunneling (BTBT) and the trap-assisted tunneling (TAT). However, TAT plays only a marginal role in case of heavily doped ($N_{A,D} > 10^{18}$ cm^{-3}) homojunctions in reverse bias [25], while it is crucial in case of tunneling through oxide barriers and heterojunctions [26]. On the contrary, BTBT must be added to the recombination mechanisms (SRH, Auger, radiative) already incorporated in the drift-diffusion solver to get significant and reliable results about BTJs.

This is very important because, in principle, no significant current should be able to pass through a reversely biased junction, due to the presence of a very high potential barrier arising due to the different type of doping at the two sides and the effects of the applied reverse bias. Nevertheless, such a mechanism is able to bring a remarkable amount of charges from one side to the other one of the junction by means of quantum tunneling. In particular, it is able to move electrons from the top of the valence band (VB) on the *p*-side to the bottom of the conduction band (CB) on the *n*-side of the junction. Therefore, it could be exploited in VCSEL structure to inject holes in the active region by using only a very thin *p*-doped layer near the active region, instead of large *p* contacts, which otherwise could lead to some issues, for the reasons already discussed in section 1.2.

3.1 BTBT modelling

In the following sections an overview about the main band-to-band tunneling models is presented. This is done to provide a brief physical justification to all the performed numerical simulations.

The simplest model dealing with quantum tunneling mechanisms is based on the Wentzel-Kramers-Brillouin (**WKB**) theory, which allows to evaluate bound-state en-

ergies and tunneling rates through a potential barrier. The basic idea behind this semiclassical approximation is to avoid the application of the perturbation theory to reach a solution of the one-dimensional time independent Schrödinger equation. This is possible with the assumption of a particle an energy E moving in a *slowly varying* potential $V(x)$. In general, the particle wavevector $k(x)$ in the 1D Schrödinger equation:

$$\frac{d^2\psi(x)}{dx^2} + k(x)^2\psi(x) = 0 \tag{3.1}$$

takes the following form, according to the magnitude of E with respect to $V(x)$:

$$k(x) = \sqrt{\frac{2m}{\hbar^2}(E - V(x))}, \qquad E > V(x) \tag{3.2a}$$

$$k(x) = i\sqrt{\frac{2m}{\hbar^2}(V(x) - E)} = i\kappa(x), \qquad E < V(x) \tag{3.2b}$$

In case $k(x)$ is constant, a simple solution of (3.1) is easily found to be $\psi(x) = Ae^{\pm ikx}$, where the amplitude is given by A and its wavelength is given by $\lambda = 2\pi/k$. On the other hand, when $k(x)$ changes (this is actually the case of non-constant potential $V(x)$) another solution can be used: $\psi(x) = e^{\pm i\int k(t)dt}$. In particular, in case of slowly varying potential (slowly with respect to λ), the varying part of k is much smaller than its constant part, allowing to exploit a first-order approximation. Therefore, the wavefunction can be rewritten as:

$$\psi_1(x) = \frac{1}{\sqrt{k_z(x)}} \exp\left(\pm i \int_{x_1}^{x_2} k_z(x)dx\right) \tag{3.3}$$

where $k_z(x)$ is the imaginary transverse component to the potential profile (i.e., the tunneling direction z), while the integration extrema are the *classical turning points*, namely the points where $E = V(x)$, thus the boundary points identifying the regions in which both classical and quantum tunneling are allowed.

In a reversely biased heavily doped *pn* junction, i.e., in a tunnel junction, the energies available for quantum tunneling span from the valence band edge energy at the *p*-side to the conduction band edge energy at the *n*-side. Such energies can be used to identify the classical turning points x_1 and x_2 needed for the WKB approximation [27]. In this framework, the tunneling probability along z direction across the forbidden band gap is described by the following expression:

$$T_{WKB} = \exp\left(-\int_{x_1}^{x_2} 2k_z(x)dx\right) \tag{3.4}$$

Notice that the maximum value of tunneling is reached at the midgap point of the junction, whereas at the turning points $T = 0$, since $k_z(x_1) = k_z(x_2) = 0$.

Solving WKB approximation assuming a triangular potential barrier leads to the Fowler-Nordheim tunneling transmission expression [28, 29]:

$$T_{WKB}^{FN} = \exp\left(-\frac{4}{q\mathcal{E}\frac{\sqrt{2m_t^* E_g^{3/2}}}{3\hbar}}\right) \tag{3.5}$$

where q is the elementary charge, $\mathcal{E}$ is the electric field in the tunneling region, E_g is the semiconductor band gap and m_t^* is the tunneling effective mass, whose value depends on the investigated structure.

3.1.1 Kane's Model

WKB approximation is useful to compute states and tunneling rate of simple potential profiles. Therefore, the crucial step of the thesis is the research for a simple model able to grasp the main details of the BTBT phenomenon across a BTJ and yet easily implementable inside the drift-diffusion simulator. To this purpose, *local* tunneling models are taken into account in first place, meaning that *non-local* properties of the process are adapted to obtain local variables. In literature, a great work has been done by Kane in its inspiring works [30, 31], and here some of the main details about them are reported.

The starting point of the model is the evaluation of the moving particle energy eigenfunctions, needed to provide a rigorous treatment of the phenomenon. To this purpose, the time-independent Schrödinger equation of an electron in case of uniform electric field in one direction (hereinafter is assumed to be the z-direction) must be solved:

$$\left[E_n(k) - i\mathcal{E}\frac{\partial}{\partial k_z} - E_{tot}\right]\phi_n(k) - \sum_m \mathcal{E}X_{n,m}\phi_m(k) = 0 \tag{3.6}$$

where n, m are band indexes, $\mathcal{E}$ is the (electric) force acting on the electron, $E_n(k)$ is the electron energy in the semiconductor bulk (with k electron wavevector) and E_{tot} is the total energy. $\phi_{n,m}(k)$ is the wavefunction included in the Bloch approximation:

$$\psi(r) = \sum_{n,k} \phi_n(k)\Psi_{n,k}(r) \tag{3.7}$$

where $\Psi_{n,k}(k)$ is a Bloch function. The transitions are governed by the terms $i\partial/\partial k_z$ and $X_{n,m}$, called *intra* and *inter*-band operators, respectively, obtained by decomposing the position operator. The off-diagonal matrix elements of X, actually including the interband components, are defined as:

$$X_{n,m} = i\int u_{n,k}^*(z)\frac{\partial}{\partial k}u_{m,k}(z)dz \tag{3.8}$$

where $u(z)$ is the periodic part of the Bloch function and the integration in space includes the volume of the unit cell.

Kane computes the intraband matrix elements by first extracting the eigenfunctions from (3.6) neglecting the operator $i\partial/\partial k_z$, thus obtaining:

$$M_{n,m}^{(1)}(k_z) = -\frac{\mathcal{E}}{\mathbf{K}} \int X_{n,m}(k) \exp\left\{ \frac{i}{\mathcal{E}} \int_0^{k_z} \left[E_n^{(1)}(k', \mathcal{E}) - E_m^{(1)}(k', \mathcal{E}) \right] dk'_z \right\} dk_z \tag{3.9}$$

where superscript 1 denotes the term related to transitions with electric field at the first order (even though Kane demonstrated that it is equivalent to the expression for second order case) and $\mathbf{K}$ is the principal vector of the reciprocal lattice. The probability per unit time, w, that a transition from band n to band m occurs is obtained by means of the Fermi Golden's rule:

$$w = \frac{2\pi}{\hbar} \left| M_{n,m}^{(1)} \right|^2 \rho(E) \tag{3.10}$$

where $\rho(E)$ is the density of states.

Now, if one wants to evaluate the transition rate between conduction and valence band, first needs to take into account the dispersion relations. Thus, let's assume a two-band dispersion relation, as done by Mandurrino [32, pp. 114, 115]:

$$E_{c,v}(k) = \frac{E_g}{2} + \frac{\hbar^2 k^2}{2m_{n,p}^*} \pm \frac{\sqrt{E_g^2 + E_g \frac{\hbar^2 k^2}{m_r^*}}}{2} \tag{3.11}$$

where E_g is the energy band gap and $m_{n,m}^*$ are the electron and hole effective masses; $m_r^* = (m_n^* \cdot m_p^*)/(m_n^* + m_p^*)$ is the reduced mass.

Then, the expression for the off-diagonal elements of the interband matrix X is also computed:

$$X_{c,v} = i \frac{\hbar E_g^{2/3}}{2\sqrt{m_r^*}(E_c - E_v)^2} \tag{3.12}$$

Before calculating the transition probability by inserting (3.11) and (3.12) in (3.9) and then in (3.10), a crucial assumption is needed in order to simplify the calculations. In fact, the assumption of sufficiently small and constant electric field $\mathcal{E}$ allows to simplify the integral in (3.9) and obtain [31]:

$$M_{c,v}(k_z) = \frac{\pi \mathcal{E}}{3\mathbf{K}} \exp\left(-\frac{\pi \sqrt{m_r^*} E_g^{3/2}}{2q\hbar\mathcal{E}} \right) \tag{3.13}$$

Furthermore, such an approximation allows to write the DOS as $\rho(E) \simeq \mathbf{K}/2\pi\mathcal{E}$. Transmission coefficient can be now computed in WKB approximation as $T(E) = \hbar\mathbf{K}w/\mathcal{E}$ [31]:

$$T = \frac{\pi^2}{9} \exp\left(-2 \int_{x_1}^{x_2} ikdz \right) \tag{3.14}$$

$$= \frac{\pi^2}{9} \exp\left(-\frac{\pi \sqrt{m_r^*} E_g^{3/2}}{2\sqrt{2}\hbar\mathcal{E}} \right) \exp\left(-2E_\perp / \bar{E} \right) \tag{3.15}$$

with $E_\perp = \hbar^2(k_x^2+k_y^2)/2m_r^*$ perpendicular component of the electron energy (such that $E = E_\perp + E_z$) and $\bar{E} = \sqrt{2}\hbar\mathcal{E}/\pi\sqrt{m_r^* E_g}$ a constant which depends on the material through the effective masses and the energy band gap. It is clear that (3.15) contains two contributions to the tunneling: one coming from the electron motion in the same direction of electric field, the other coming from its perpendicular components, which are responsible of reducing such a probability.

At this point, the expression of the BTBT current density is simply given by [33]:

$$J_{btbt} = \int T(E)(f_{FD}(E, E_c) - f_{FD}(E, E_v))dE \tag{3.16}$$

which contains an integration over energy and the Fermi-Dirac statistics at the n and p-regions, to take into account the occupancy of the bands. Putting all the results together [31], the following expression is obtained:

$$J_{btbt} = \frac{qm_r^*}{18\hbar^3}\exp\left(-\frac{\pi\sqrt{m_r^* E_g^3}}{2\sqrt{2}\hbar\mathcal{E}}\right)\int\left[f_{FD}\left(E, E_c\right) - f_{FD}\left(E, E_v\right)\right]\exp\left(-\frac{2E_\perp}{\bar{E}}\right)dEdE_\perp =$$

$$= \frac{qm_r^*}{18\hbar^3}\exp\left(-\frac{\pi\sqrt{m_r^* E_g^3}}{2\sqrt{2}\hbar\mathcal{E}}\right)\left(\frac{\bar{E}_\perp}{2}\right)\cdot$$
$$\cdot\int\left[f_{FD}\left(E, E_C\right) - f_{FD}\left(E, E_V\right)\right]\left[1-\exp\left(-\frac{2E_{min}}{\bar{E}}\right)\right]dE \tag{3.17}$$

In order to get a more manageable expression, two further assumptions are made by Mandurrino [32, p. 112]: the applied reverse bias should be strong enough to have $qV_{bias} \gg \bar{E}_\perp$, which assures discontinuous quasi-Fermi levels at the borders of the depletion region; the contribution coming from the electron perpendicular energy should be neglected, such that:

$$\int e^{-2_\perp/\bar{E}}dE_\perp \simeq \int\left(1-e^{-2E_{min}/\bar{E}}\right)dE \approx 1 \tag{3.18}$$

where E_{min} is the minimum energy between $E - E_{C,min}$ at the n-side and $E_{V,max} - E$ at the p-side. The final result, retrieved using WKB approximation, can be written as:

$$J_{btbt} \approx \frac{q^3\sqrt{2m_r^*\mathcal{E}}}{4\pi^3\hbar^2\sqrt{E_g}}V_{bias}\exp\left(-\frac{\pi\sqrt{m_r^* E_g^3}}{2\sqrt{2}q\hbar\mathcal{E}}\right) \tag{3.19}$$

3.1.2 Hurkx's Model

The model obtained in the previous section only provides a value of tunnel current density in a reversely biased tunnel junction. Unfortunately, this is not suitable for a D1ANA implementation, since it needs an expression able to take into account the

quantities at the mesh nodes or elements (like charge densities, electrostatic potential or electric field) of the spatial grid and return the corresponding value of current density in each point. However, it is important to highlight that mainly reversely biased *pn* junctions are investigated: this actually is the condition in which Hurkx's model [34] is applicable. In practice, it simplifies a lot all the formulas coming from Kane's papers, treating the band-to-band tunneling simply as a net local recombination/generation GR rate which has to be added to the other ones, like Auger, radiative, SRH rates (already included in D1ANA code). It is important to underline the fact that such a model is originally intended for silicon-based systems simulations, therefore any other material requires some adjustments to the model.

More into the details, Hurkx extracts a band-to-band tunneling GR rate, which in case of a reversely biased tunnel junction actually is a generation rate, from the expression of the current density coming from [31]. This is done by simply using the relation:

$$R_{btbt} = -\frac{1}{q}\nabla \cdot J_{btbt} = -\frac{1}{q}\frac{dJ_{btbt}}{d\psi} \cdot \nabla\psi = -\frac{dJ_{btbt}}{dE} \cdot \mathcal{E} \tag{3.20}$$

where J_{btbt} is the tunnel current density presented in the previous section. This step leads to:

$$\mathcal{R}_{tunn} = \frac{q^2\sqrt{m_r^*}}{18\pi\hbar^2\sqrt{E_g}}\mathcal{E}^2 D\left(\mathcal{E}, E, E_{fn}, E_{fp}\right) \exp\left(-\frac{\pi\sqrt{m_r^*}E_g^{3/2}}{2q\hbar\mathcal{E}}\right) \tag{3.21}$$

whose units are the typical ones of a recombination rate: $\mathrm{cm}^{-3}\mathrm{s}^{-1}$. Notice that the function $D\left(\mathcal{E}, E, E_{fn}, E_{fp}\right)$ accounts for the relative position of the electron and hole quasi-Fermi levels E_{fn} and E_{fp} in the neutral regions and for the dependence of electrons tunneling on the electric field $\mathcal{E}$, and is better described afterwards. The terms containing fundamental physical constant and/or material parameters can be grouped into **Kane's coefficients** A and B, defined as:

$$A = \frac{q^2\sqrt{m_r^*}}{18\pi\hbar^2\sqrt{E_g}} \tag{3.22}$$

with units $\mathrm{V}^{-2}\mathrm{cm}^{-1}\mathrm{s}^{-1}$, and

$$B = \frac{\pi\sqrt{m_r^*}E_g^{3/2}}{2q\hbar} \tag{3.23}$$

that has the units of an electric field (V/cm). In fact, B can be seen as the *critical* electric field at which $\mathcal{R}_{tunn}$ starts to play a significant role, typically few tens of MV/cm large. The Kane's coefficients can be used to rewrite (3.21) in a simpler way [34]:

$$\mathcal{R}_{btbt} = A\mathcal{E}^2 D\left(\mathcal{E}, E, E_{fn}, E_{fp}\right) \cdot \exp\left(-B/\mathcal{E}\right) \tag{3.24}$$

which is the expression actually implemented in D1ANA to include band-to-band tunneling contribution. Notice that in a reversely biased tunnel junction, the transition

probability described in (3.21) is a generation rate, and this is the power of Hurkx's model.

Before explaining how (3.21) has been included in D1ANA together with the other recombination mechanisms, let's return to the function D. As a matter of fact, such a function comes from the fact that in his calculations Hurkx has not neglected the contribution coming from the perpendicular energy integral, such that:

$$D\left(\mathcal{E}, E, E_{fn}, E_{fp}\right) \simeq \int 1 - e^{-2E_{\perp}/\bar{E}} dE_{\perp}$$

while in Kane it is approximated to 1 (see eq. (3.18)).

In forward bias, D is able to model the tunneling current in an Esaki diode. However, Hurkx in [34] also provides an expression for it which is suitable for implementation in a device simulator when dealing with zero and reverse bias conditions, which in turn comes from [31]:

$$D(\psi, E_{fn}, E_{fp}) = \frac{1}{\exp\left(\frac{-E_{fp}-q\psi}{k_B T}\right)+1} - \frac{1}{\exp\left(\frac{-E_{fn}-q\psi}{k_B T}\right)+1} \tag{3.25}$$

This function should be equal to zero outside the depletion region because of lack of final states *into* which an electron can tunnel or initial states *from* an electron can tunnel.

Whereas eq. (3.25) has a problem related to the reference potential, which is not well defined and may lead to a wrong evaluation of the electrostatic potential ψ, however *Sentaurus Synopsys* TCAD user guide [35, p. 453] provides an alternative formulation of the function D:

$$D(n, p) = \frac{np - n_i^2}{(n + n_i^2)(p + n_i^2)} \tag{3.26}$$

This results to be much more easy to be implemented in D1ANA. In fact, the code is able to compute at each point of the mesh the value of the charge concentration, such that the value of D is quickly calculated. Notice that in this framework, BTBT GR does not explicitly depend on tunnel path length or on tunnel window wideness, but only on charge concentrations and electrical field across the system, since it is a *local* model.

The formulation contained in (3.26) is very useful when trying to include (3.21) in D1ANA, as it allows to evaluate the recombination (which actually is a generation) rate due to interband band-to-band tunneling at each node of the mesh. As a matter of fact, R_{btbt} is implemented in a total similar way with respect to the other rates. A slightly deeper analysis is needed to for its derivatives, included in the Jacobian matrix. In fact, while SRH, Auger and radiative recombination show derivatives equal to zero with respect to the potential, the presence of electric field in (3.21) requires also the computation of $\partial R_{btbt}/\partial\phi$:

$$\frac{\partial R_{btbt}}{\partial n} = A\mathcal{E}^2 e^{-B/\mathcal{E}} \frac{n_i}{n + n_i^2}$$

$$\frac{\partial R_{btbt}}{\partial p} = A\mathcal{E}^2 e^{-B/\mathcal{E}} \frac{n_i}{p + n_i^2}$$

$$\frac{\partial R_{btbt}}{\partial \phi} = AD(n,p)\left[\frac{\partial \mathcal{E}^2}{\partial \phi} e^{-B/\mathcal{E}} + \mathcal{E}^2 \left(\frac{\partial}{\partial \phi} e^{-B/\mathcal{E}}\right)\right] =$$

$$= AD(n,p) e^{-B/\mathcal{E}} (2\mathcal{E} + B) \frac{\partial \mathcal{E}}{\partial \phi}$$

where the derivative of the electric field with respect to the electrostatic potential is computed point-wise in D1ANA, without using an analytical expression.

3.2 BTBT in *Sentaurus Device*

One of the reasons leading to the use of *Sentaurus Device* TCAD as a benchmark for some of the results obtained from D1ANA is the fact that the former allows to select among a number of band-to-band tunneling models, one of them being the Hurkx's one. Furthermore, we have actually borrowed the definition of D (3.26) contained in [34] from Sentaurus manual because it provides a simpler way to implement the function in a drift-diffusion code.

Just like it is done in D1ANA, here BTBT is modeled by an additional GR process. The expression for R_{btbt} used by *Sentaurus Device* also includes a basic expression to introduce the effects due to temperature, which acts on the energy gap and thus on the tunneling (higher T typically means larger lattice constant, which in turn reduces the energy band gap: tunneling is enhanced; the converse holds for T $<$ 300 K). The function D can take either positive or negative values: if $D > 0$ recombination is dominating; if $D < 0$, then net generation is larger (and this is the condition we are actually working into in case of reversely biased tunnel junctions). The TCAD uses fixed values for Kane's parameters (3.22 and 3.23), [35, p. 452]:

$$A = 3.4 \cdot 10^{21}\ cm^{-1}s^{-1}V^{-2}, \quad B = 22.6 \cdot 10^{6}\ V/cm$$

These values will be used also in D1ANA instead of the material dependent ones previously described, in order to perform a more significant comparison between the simulators, rather than obtain more accurate results for the investigated structures. However, let's remark that the values computed for each material are very close to these ones, especially the critical electric field, such that there is no loss of information due to this slight modification.

3.2.1 Other band-to-band tunneling models

As mentioned at the beginning of this section, the used TCAD includes several ways to model BTBT. Here there is a brief list of the most interesting ones.

- **Kane**'s model [35, p. 452]: the expression is very similar to (3.21), beside the fact that the function D is dropped:

$$R_{btbt}^{Kane} = A\mathcal{E}^2 \exp\left(-\frac{B}{\mathcal{E}}\right)$$

  As a consequence, such a simple model predicts a nonzero generation rate even at equilibrium, which obviously is far from a real case

- **Schenk** model, for phonon-assisted BTBT in highly doped junctions reaching very high electric field values

- **Non-local** band-to-band tunneling model. In contrast to the previous models, this model introduces real spatial carrier transport through barriers. It is straightforward that whereas in local models e-h pair recombines in the same position, in a non-local model recombination occurs at different physical positions (at the ends of the tunneling path), such that there is an actual transport of carriers [33]. *Senaturus Device* allows to activate a non-local model for tunneling mechanism [36], again relying on WKB approximation. This is usually required in case of devices containing some abrupt and graded heterojunctions [35, p. 454]. The corresponding GR rate takes the following form:

$$R_{CV}(u,l,E) = \frac{A_{CV}}{qk_B} \cdot \mathcal{E} \cdot \Gamma_{CV}(r) \cdot T \cdot \ln\left[\frac{1+\exp\left(\frac{E_{fp}(u)-E}{k_BT}\right)}{1+\exp\left(\frac{E_{fn}(l)-E}{k_BT}\right)}\right] \tag{3.27}$$

  where u and l are the points corresponding to the tunneling window (with E range of energies included in the window). All the other parameters are fitting or material constants.

  This model actually does not bring to much different results, meaning that it must be optimized to return significant improvements. Furthermore, in D1ANA this is not implementable, but an attempt of adding a non-local model is done by moving to NEGF formalism, that will be presented at the end of the next chapter together with the corresponding results.

CHAPTER 4

Investigated structures

Now that the main transport mechanism occurring in a reverse bias pn junction has been described and implemented, it is time to move to the main purpose of the thesis: the complete simulation of VCSELs with buried tunnel junctions included in their layers. Let's remark that all the simulations are only performed from the **electrical** results point of view, while both optical and thermal analysis are not considered here. Two case studies are investigated, including both short and long-wavelength devices. In the following sections are presented all the results obtained for each of them. In both cases, further analysis are performed on the included TJs, in order to fully understand the effects of the junction on the results. This is done in order to comprehensively understand if the models retrieved in literature actually work properly or not for these material systems. Of course, a comparison with literature results is also performed.

4.1 Long-wavelength BTJ-VCSEL

For the reasons already introduced in 1.2, literature is full of works dealing with long-wavelength device, such as VCSELs and LEDs, including a BTJ in their structure. Therefore, it is reasonable to start from the analysis of one of them. After having read several papers [16, 37–39] which however lack of crucial details about doping levels, geometry or molar fractions, the choice for the first investigated device has fallen on the fully described structure presented in Lamy's work [40]. In particular, it deals with a long wavelength $\lambda = 1.55$ μm electrically pumped lattice matched $In_{1-x}Ga_xAs_yP_{1-y}$ quantum well based BTJ-VCSEL grown on InP substrate.

As already mentioned, the buried tunnel junction is exploited here to inject holes into the active region of the device, by using only a very thin layer of p-doped material. All the layers (grown over an InP substrate) embodied in the structure are reported in table 4.1, together with the corresponding molar fractions, thickness, doping type and level. The overall length of the structure is approximately 300 nm. The DBR layers are not included in the latter because they are made of dielectric material (a-

Si ($n = 3.7$) and a-SiN$_x$ ($n = 1.8$) stack realized by means of magnetron sputtering, leading to a reflectivity value of 99.5% at $\lambda = 1.55$ μm): as a consequence, electrical simulations is not performed for such layers, and thus are not inserted in D1ANA code. The active region can be realized by growing by molecular beam epitaxy (MBE) and contains six quantum wells (molar fractions: x = 0.47, y = 1; 7.2 nm of length each); lattice matched alloy ($In_{0.8}Ga_{0.2}As_{0.435}P_{0.565}$ named $Q_{1.18}$ because its gap corresponds to $\lambda = 1.18$ μm) is used to generate the five barriers (length = 10 nm each).

As already highlighted, it is worth noticing that the layers corresponding to the contacts (#1 and #19) are both heavily n-doped ($N_D = 1 \cdot 10^{19}$ cm^{-3}), thanks to the BTJ included in layers #2 and #3, that allows to avoid p contacts. In particular, the tunnel junction mesa is realized with a strongly doped ($N_D = N_A = 5 \cdot 10^{19}$ cm^{-3}) lattice matched alloy called $Q_{1.4}$ (molar fractions: x = 0.34, y = 0.732; each side is 25 nm long), whose energy gap avoids optical absorption at $\lambda = 1.55$ μm (to further limit this phenomenon, BTJ is thin and positioned at a node of the standing wave of the VCSEL) and which due to applied bias allows for strong quantum tunneling probability. While carrier and optical confinements in the active region are provided by the tunnel junction mesa, the current confinement is further ensured by a 30 nm p^+-doped InP layer #5, realized by stopping the MBE growth at the $Q_{1.4}$ and realizing a mesa by chemical etching together with a InP over-etching. Such a layer is important also because it should avoid the change of curvature of active region band diagram in reverse bias conditions. In order to obtain a transversal monomode emission, VCSEL should have a diameter < 10 μm [40].

Before presenting the results of the simulations, it is worth talking about the inclusion of InGaAsP-based layers into D1ANA. In fact, when the thesis work started, the library containing the material parameters did not include InGaAsP ones in the list of materials. Consequently, before starting any simulations, the needed InGaAsP parameters are retrieved from the proper section of NSM Archive [24]. Given that this is a quaternary alloy and Lamy's structure contains layers with different values of molar fractions x and y, some kind of linear interpolation is required to evaluate accurately each layer parameters. This is really useful, because it allows to tailor some crucial parameters (e.g. lattice constant of energy band gap) to obtain alloys with intermediate properties with respect to their components.

A good choice to compute the parameters of $In_{1-x}Ga_xAs_yP_{1-y}$ is to use a global Vegard's law [41], which is a bilinear approximation:

$$P_{InGaAsP} = (1-x)(1-y)P_{GaAs} + (1-x)yP_{GaP} + xyP_{InP} + x(1-y)P_{InAs} \qquad (4.1)$$

where P is a generic property of the alloy (e.g. E_g, ϵ_r, $m^*_{n,p}$, ...). Let's remark that GaP has an indirect band gap, and thus (4.1) could lead to little inaccuracies [2, p. 29]. Electron affinity χ needs a further remark: in fact, by taking a look to the calculated band diagram, it seems clear that Vegard's law is not suited to interpolate

Layer	#	Material	Molar fractions	Thickness [nm]	Doping type	Doping level [cm^{-3}]
Left contact	1	InP	x = 0; y = 0	30	***n***	$1 \cdot 10^{19}$
TJ	2	$Q_{1.4}$	x = 0.34; y = 0.732	**25**	n	$5 \cdot 10^{19}$
TJ	3	$Q_{1.4}$	x = 0.34; y = 0.732	**25**	p	$5 \cdot 10^{19}$
Buffer layer	4	$Q_{1.18}$	x = 0.2; y = 0.435	10	p	$2 \cdot 10^{18}$
EBL	5	InP	x = 0; y = 0	30	p	$2 \cdot 10^{18}$
Cladding	6	$Q_{1.18}$	x = 0.2; y = 0.435	30	/	/
QW 1	7	InGaAs	x = 0.47; y = 1	7.2	/	/
Barrier 1	8	$Q_{1.18}$	x = 0.2; y = 0.435	10	/	/
QW 2	9	InGaAs	x = 0.47; y = 1	7.2	/	/
Barrier 2	10	$Q_{1.18}$	x = 0.2; y = 0.435	10	/	/
QW 3	11	InGaAs	x = 0.47; y = 1	7.2	/	/
Barrier 2	12	$Q_{1.18}$	x = 0.2; y = 0.435	10	/	/
QW 4	13	InGaAs	x = 0.47; y = 1	7.2	/	/
Barrier 4	14	$Q_{1.18}$	x = 0.2; y = 0.435	10	/	/
QW 5	15	InGaAs	x = 0.47; y = 1	7.2	/	/
Barrier 5	16	$Q_{1.18}$	x = 0.2; y = 0.435	10	/	/
QW 6	17	InGaAs	x = 0.47; y = 1	7.2	/	/
Cladding	18	$Q_{1.18}$	x = 0.2; y = 0.435	30	/	/
Right contact	19	InP	x = 0; y = 0	30	***n***	$1 \cdot 10^{19}$

Table 4.1: Summary of the layers which compose the **long**-wavelength VCSEL structure [40] along a vertical cut passing through the buried tunnel junction and the top and bottom contacts. All of them are made of quaternary alloy $In_{1-x}Ga_xAs_yP_{1-y}$, with different molar fractions x and y.

such a parameter. Therefore, Anderson's rule [42] has been used to correctly generate the energy band diagram, by taking χ_{InAs} and $E_{g,InAs}$ as the reference values for the electron affinity and energy band gap, and by setting the conduction band offset (CBO) to 0.65.

4.1.1 Results

Once that all the structure layers are inserted in the code, a simulation at **equilibrium** condition is run, where Fermi statistics and incomplete ionization of dopants are used (PFII); all the recombination mechanisms are activated. The obtained energy band diagram is reported in fig. 4.1, which seems to reproduce rather accurately the one proposed in fig. 5 of [40], meaning that the implementation of the layers has been performed correctly. Moreover, notice that when the applied bias is equal to 0, the tunnel junction arising at approximately 50 nm from the left contact is not able to

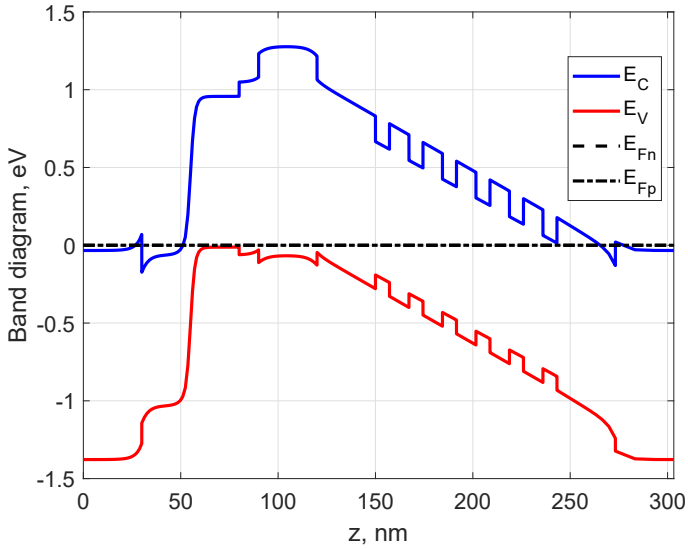


Figure 4.1: Energy band diagram of the VCSEL described in [40], at equilibrium condition.

bring a significant amount of holes in the active region, because the bands are not overlooking so they don't allow for any interband band-to-band tunneling of carriers. As expected from an equilibrium condition, right and left contacts of the VCSEL lie at the same energy, due to the very same doping level; furthermore, the quasi-Fermi levels are superimposed and don't move from 0 eV, since no recombination mechanism is acting to modify them.

Of course, it is more interesting an analysis of the **forward** bias condition of the device, realized by applying a positive tension (cathode) to the n-side of the tunnel junction, i.e., at the left hand side of the VCSEL, in order to bring the BTJ in reverse bias condition. A voltage sweep is thus performed from 0 V to 2 V: when such a value is reached, D1ANA plots the quantities of interest. Convergence of the code is very fast, requiring just three or four iterations at each step of the voltage sweep. The voltage step between each bias must be set to a value small enough to avoid too steep jump in successive simulations (which would lead to a convergence failure), but it should also be large enough to reduce simulation times. Fermi distribution with incomplete ionization of dopants is assumed, with Auger, radiative, SRH, BTBT models activated. For the latter, Kane's coefficients take the following values:

$$A = 2.5 \cdot 10^{20}\, cm^{-1} s^{-1} V^{-2}, \quad B = 13.9 \cdot 10^{6}\, V/cm$$

Again, the first analysed plot is the one containing the energy band diagram, represented in fig. 4.2a. The most interesting detail to highlight is that when a strong reverse bias is applied to the pn junction, the BTJ works as described in the previous chapter: bands get much closer in space with respect to the equilibrium condition, and

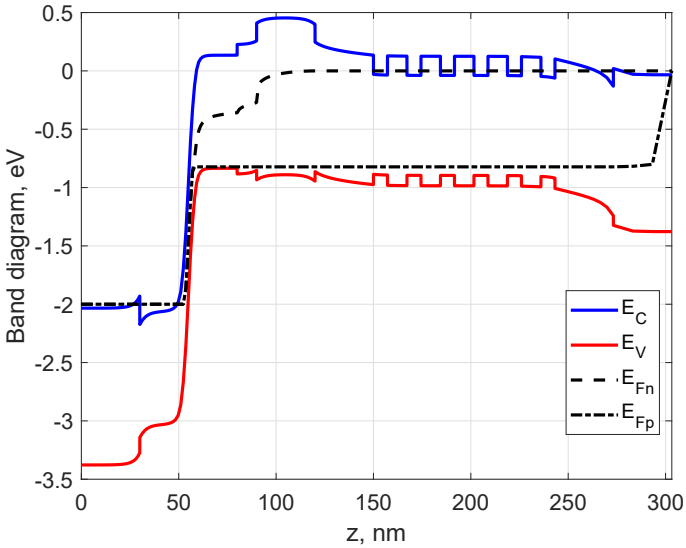
the tunnel window is clearly much wider for what concerns the energy values, thus providing the carriers with a significant tunnel path. However, in the framework of Hurkx's model this does not affect the tunneling current. What actually matters is the shorter depletion region which yields larger electric fields at the interface of the junction sides. As a consequence, local GR rate from Hurkx's model reaches significant values in the central region of the tunnel junction, thus leading to a significant tunnel current. The effect of the applied bias, and consequently of BTBT mechanism, is also clear from the splitting of the electrons and holes quasi-Fermi levels, especially in the active region. The role of the BTJ in hole injection can be appreciated in fig. 4.2b. The presence of n-doped contacts forces zero holes current density to be null at the two ends of the VCSEL, but in correspondence of the BTJ J_p reaches the level of J_n at the contacts, due to the conservation of the current across the structure, providing the correct amount of holes to the active region.

Obviously, D1ANA extracts a value of current at each bias point; from these a current-voltage plot is realized and reported in fig. 4.3. It is worth noticing that a significant current starts to flow through the structure only above a certain value of applied forward bias, similar to the results presented in fig. 5 of [16]. Once the threshold is reached, the dependence of current on the applied voltage is almost linear (instead of the typical parabolic trend displayed in literature), whose slope depends on the sum of the series resistances across the structure, namely the contributions coming from: TJ, pn junctions, bulk layer and contacts (which typically are designed to provide a very small resistance to the overall structure; the use of a BTJ actually allows to reduce it by avoiding highly resistive p-doped contacts). This is due to the nature of the band-to-band tunneling and to the active region layers that limits the current, and eventually by the current blocking layer (layer #5 of tab 4.1). Notice that the reached current level is close to the values obtained for VCSELs emitting at $\lambda = 1.55$ μm [43]. Therefore, the inclusion of a BTJ in a long-wavelength VCSEL based on InGaAsP is demonstrated effectively substitute a p-type contact, without worsening carrier transport meanwhile providing all the advantages presented in the introduction.

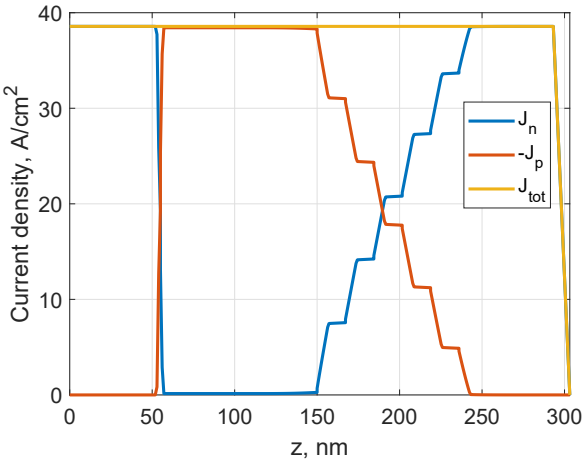
In the following section, a deeper analysis on the InGaAsP-based tunnel junction characteristics is presented, in order to better evaluate the impact of the latter and of BTBT process on the performances of the whole VCSEL.

4.1.2 InGaAsP BTJ

After having investigated the whole VCSEL structure, a deeper analysis of the buried tunnel junction included in the device is performed. This is useful to understand if the structure actually operates as expected, and if the Hurkx's model presented in section 3.1.2 is correctly implemented in D1ANA. To this aim, two different simulators have been exploited. Of course, one is D1ANA; but, for a more comprehensive analysis, also



(a) *Energy band diagram*



(b) *Electrons, holes and total current densities*

Figure 4.2: Energy band diagram and current densities of Lamy's VCSEL in reverse bias condition (2 V applied to the left contact).

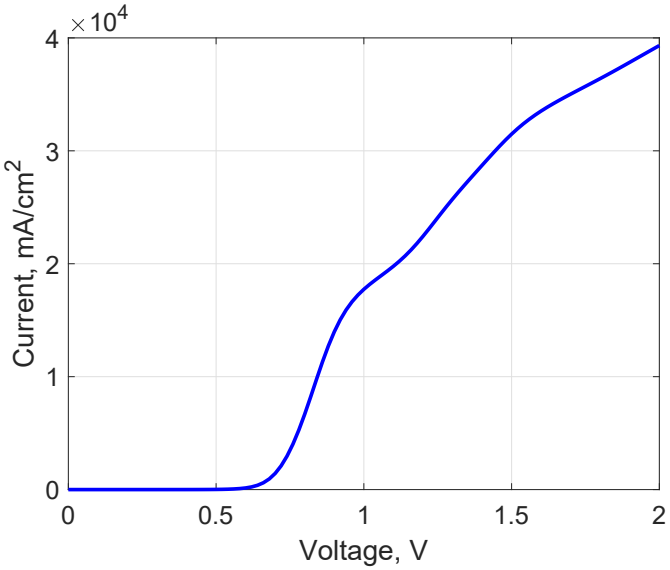


Figure 4.3: Current density vs Voltage plot of Lamy's device, coming from D1ANA simulations in PFII with all the recombination mechanisms turned on.

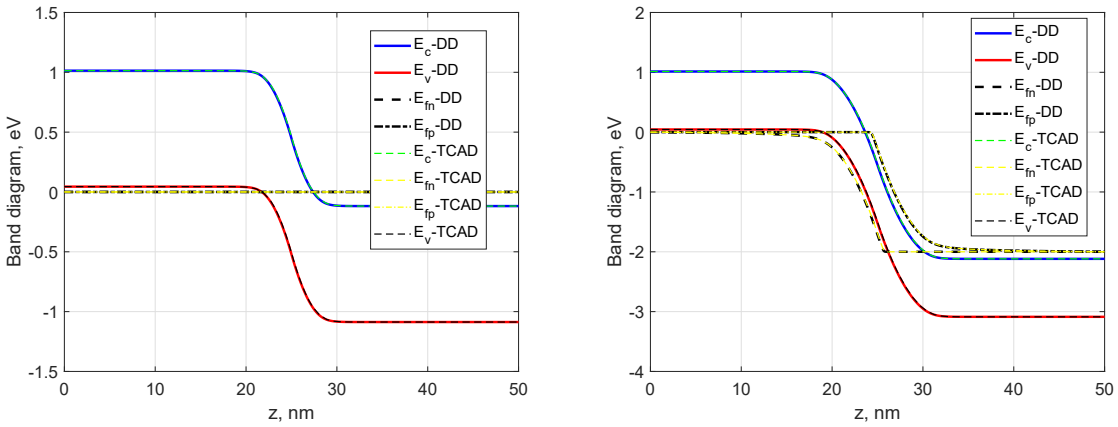
Sentaurus Device TCAD from Synopsys has been used.

Let's recall that the BTJ is 50 nm long, such that each side (with molar fractions $x = 0.34$ and $y = 0.732$) has a length of 25 nm. The doping levels are symmetric and equal to the tunnel junction described in table 4.1, namely $N_D = N_A = 5 \cdot 10^{19}$ cm^{-3}, realized with Te ($\Delta E_D \approx 3$ meV as donor activation energy) and Mg ($\Delta E_A \approx 35$ meV as acceptor activation energy), respectively.

Charge models

Since this is the first analysed BTJ, all the possible charge models described in section 2.1.1 (PBFI, PBII, PFFI, PFII) are investigated and compared, in order to appreciate their impact on the results. Moreover, beside the band-to-band tunneling models, the other three recombination mechanisms included in D1ANA are compared with TCAD results, to understand if they play a significant role in carrier transport. In doing this, it is also taken the chance to understand the relative importance of such processes in a tunnel junction working in reverse bias conditions. A brief remark is needed before going on: Auger, SRH and radiative recombination by spontaneous emission are activated only for PFII model. Furthermore, for a fair comparison they are not switched on all together, but one simulation at a time is performed in order to better understand the contribution coming from each mechanism. On the other hand, BTBT is activated for all the charge density models, while it is turned off in PFII whenever one of the other process is investigated.

As a first step, every recombination mechanism is left out from all the simulations.



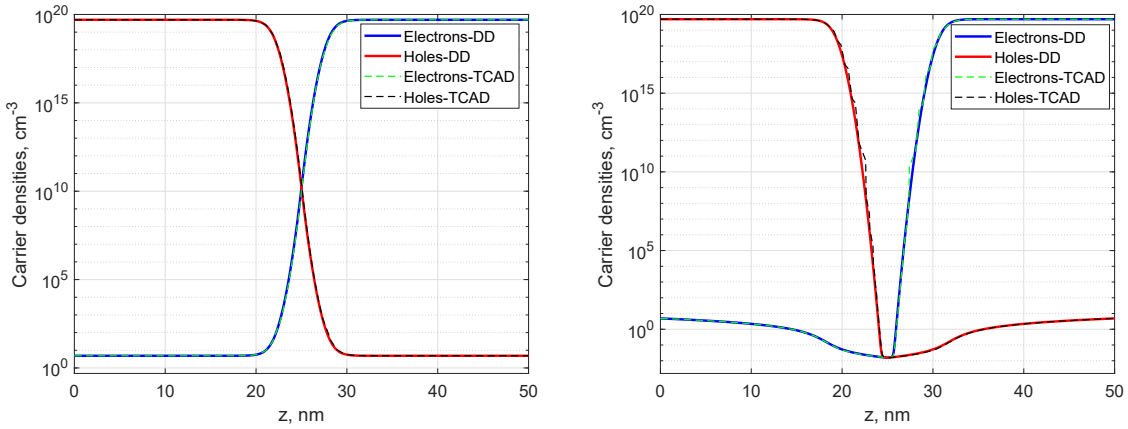
(a) *Equilibrium condition* (b) *Reverse bias condition*

Figure 4.4: Energy band diagram of InGaAsP BTJ at equilibrium and at reverse bias condition with **PBFI** model, without any GR mechanism, in both the simulators.

This is mainly done to ensure a correct implementation of all the material parameters. The simplest case involves Boltzmann approximation (2.8) and full ionization of dopants (PBFI). The band diagrams coming from both D1ANA and *Sentaurus Device* for this case are reported in fig. 4.4. Both equilibrium (fig. 4.4a) and reverse bias condition (fig. 4.4b) the superimposition of the curves coming from the two simulators displays a high degree of agreement. Of course, at equilibrium the quasi-Fermi levels are not split and lie on the zero level, chosen as a reference by both the simulators. It is interesting to notice that in this framework, at equilibrium conduction band and valence band already offer to the carriers a tunneling path, even though tunneling current is still equal to 0, suggesting that PBFI is not suited. On the other hand, when a 2 V reverse bias is applied to the *pn* junction, electron and hole quasi-Fermi levels undergo a separation in the depletion region.

Such a situation is reflected into an another interesting quantity to evaluate: the charge densities across the tunnel junction. The comparison of them between the two bias conditions is plotted in fig. 4.5. In both the figures, the doping levels far from the depletion region are equal to the "nominal" one, since a full ionization model is used here. In reverse bias condition, see fig. 4.5b, mass action law does not hold any more, so that all the recombination mechanisms will surely be different from zero at the center of the structure. Moreover, in that region the overall charge level of both carriers is below 1 cm^{-3}, when BTBT is not activated.

Even though the correspondence between the curves appear to be good, a zoom in the depletion region, where the bands bending is more accentuated, reveals some discrepancies, especially in reverse bias condition. Therefore, in figures 4.6, the zoom of both energy band diagram and carrier densities are plotted in reverse bias (2 V) condition. In particular, in fig. 4.6a a very sall difference of $\approx$ 20 meV between



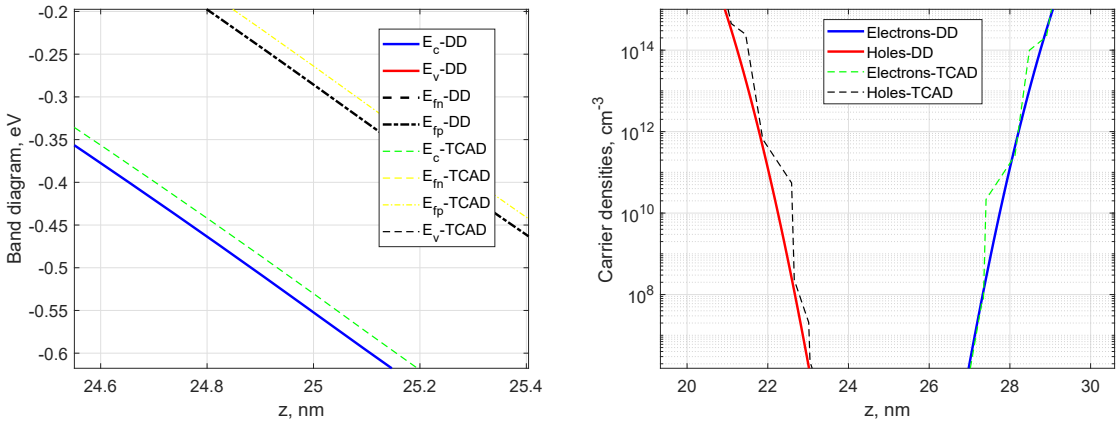
(a) *Equilibrium condition* (b) *Reverse bias condition*

Figure 4.5: Electron and hole densities of InGaAsP-based BTJ at equilibrium and at reverse bias condition with **PBFI** model, without any GR mechanism activated.

D1ANA and *Sentaurus Device* band diagram levels can be highlighted. More evident is the difference arising in the carrier densities curves. In fact, from fig. 4.6b, it is clear that the result coming from the TCAD has a significant difference with respect to D1ANA, since it shows a step-like behaviour that should not be present. Both these two issues are due to the two used different meshes: in fact, by performing some adjustments to the mesh generated by *Sentaurus Device* (as discussed at the end of section 2.3 and in fig. 2.1), a far better correspondence between the curves is obtained. The reason of this is straightforward. As a matter of fact, the spatial discretization of the domain means that all the physical quantities are approximated and associated at each node (or element) of the mesh: if the latter is not properly designed, then the interpolation of the curves by means of the trapezoidal rule between adjacent points of the 1D mesh could result in a lack of smoothness.

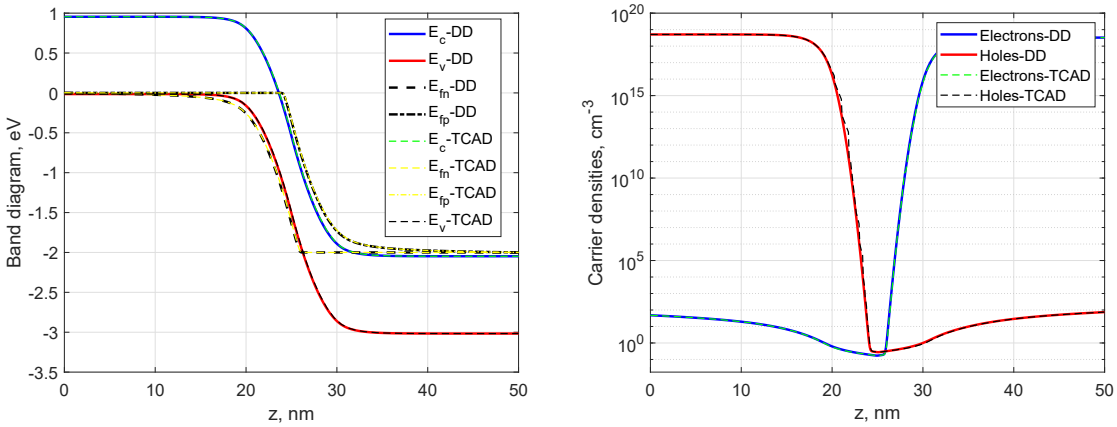
Very similar results are obtained from the PBII model, in which the only difference with respect to the previous case is the activation of a more accurate model for the dopants ionization (see equations (2.10)). Consequently, a slight asymmetry between hole and electron densities can be noticed. To appreciate it, only the results in reverse bias are reported for such a case, with energy band diagram plotted in fig. 4.7a and the carrier density curves plotted in fig. 4.7b, where actually the observations on the discrepancies made right above are even more clear.

Let's now leave Boltzmann approximation to consider the more accurate Fermi-Dirac distribution (equations (2.7)), initially considering a full ionization approximation. The most clear difference with respect to PBFI and PBII is the very strong asymmetry of the band diagram between n and p-sides. In fact, in both the bias conditions reported in fig. 4.8, the bands at the n-side of the tunnel junction reach much



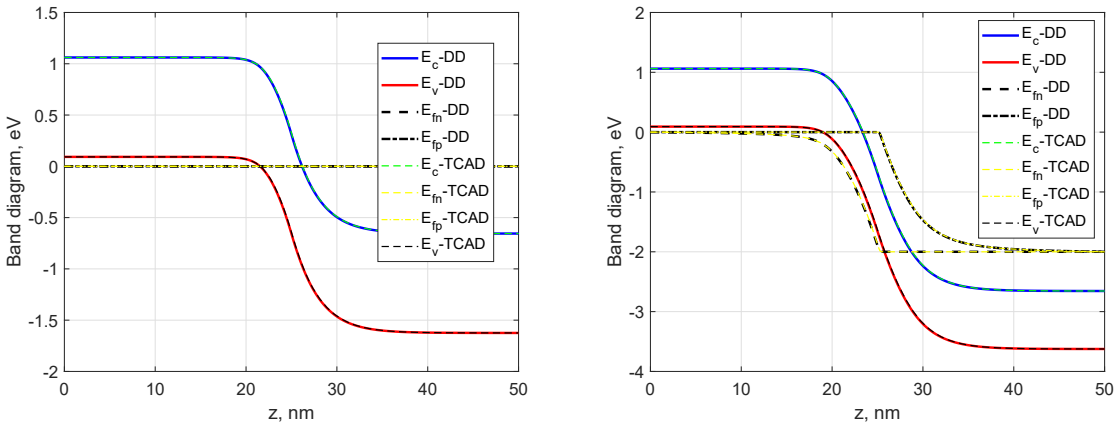
(a) *Energy band diagram.* (b) *Charge densities.*

Figure 4.6: Zoom of energy band diagram (in the left hand side figure) and charge densities (in the right hand side figure) of InGaAsP-based BTJ at reverse bias condition with **PBFI** model.



(a) *Energy band diagram* (b) *Charge density*

Figure 4.7: Energy band diagram and carrier densities of InGaAsP-based BTJ at reverse bias condition with **PBII** model; without any GR mechanism.

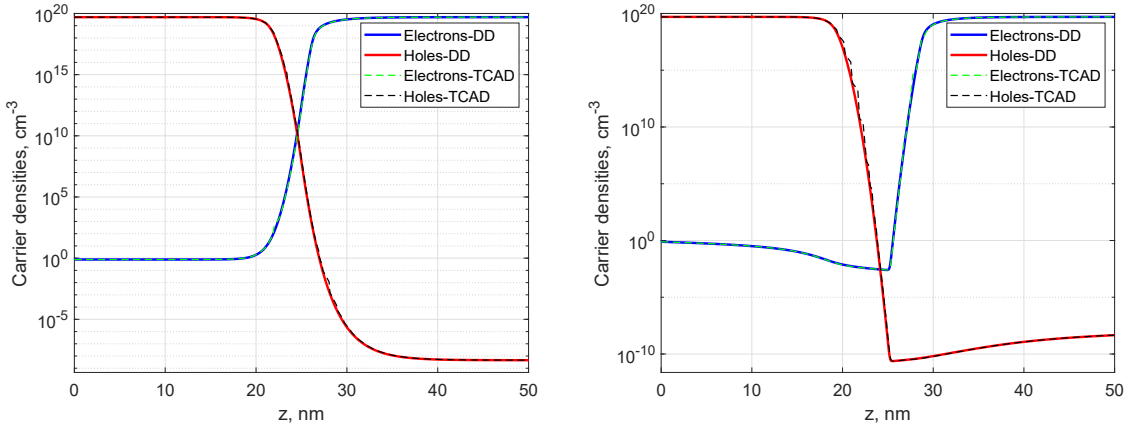


(a) *Equilibrium condition* (b) *Reverse bias condition*

Figure 4.8: Energy band diagram of InGaAsP BTJ at equilibrium and at reverse bias condition with **PFFI** model; without any GR mechanism.

lower ($\approx$ 1 eV smaller) values, significantly below the quasi-Fermi energy levels: this makes Fermi distribution paired with full ionization model a weak choice for simulating such a system. Also the separation between E_{fn} and E_{fp} is greater than in the previous cases, as clear in fig. 4.8b. Of course, the unbalancing arising in the energy bands leads to a significant difference in the charge densities, as reported in fig. 4.9. Beside this, D1ANA and *Sentaurus Device* still have a high degree of correspondence, which allows to state that the tunnel junction is correctly implemented in both the simulators, together with all the material parameters.

Eventually, incomplete ionization together with Fermi distribution (PFII) makes the curves return to reasonable levels in both conditions. The results are reported in fig. 4.10, in which the band diagrams at equilibrium and in reverse bias condition are plotted in fig. 4.10a and 4.10b, respectively, whereas the electron and hole densities are reported in fig. 4.10c. For what concerns the comparison between the band diagrams coming from D1ANA and *Sentaurus Device*, similar observations with respect to the previous cases can be made. Furthermore, the incomplete ionization model makes the energy band levels come back close to E_{fn} and E_{fp} far from the depletion region, meanwhile maintaining the separation between the latter. However, in this case, an improved mesh is also tried in the TCAD, in order to improve the smoothness of the curves. By comparing fig. 4.10d, a little improvement can be appreciated with respect to the original code result. Nevertheless, this modification yields slightly heavier TCAD simulations from the computation time standpoint, and thus it is probably not worth it given the little advantage it provides in the results.



(a) *Equilibrium condition* (b) *Reverse bias condition*

Figure 4.9: Electron and hole densities of InGaAsP BTJ at equilibrium and at reverse bias condition with **PFFI** model; without any GR mechanism.

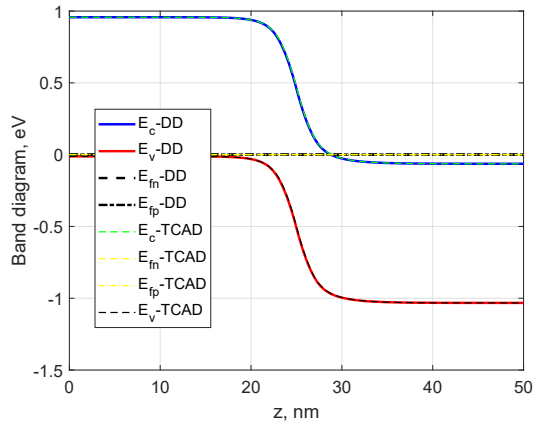
Recombination rates analysis

Given that a strong confidence that the buried tunnel junction is correctly implemented in both the simulators is reached, it is time to start a brief analysis on the recombination rates magnitude in a BTJ at reverse bias condition. Since the definitions shown in 2.1.2 lead to rates equal to zero at equilibrium (due to mass action law (2.9)), the results illustrated afterwards are obtained at a reverse bias of 2 V, and are evaluated both in D1ANA and in *Sentaurus Device*.

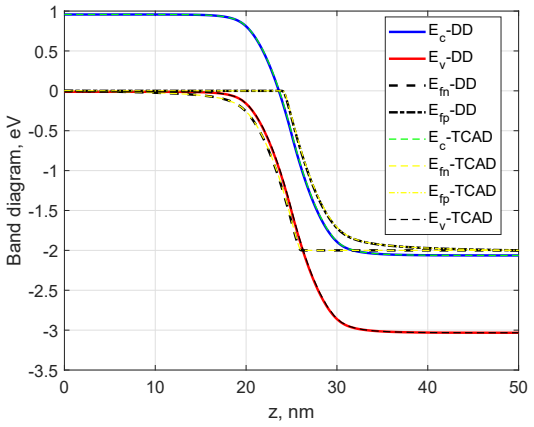
Let's start from **Auger** recombination process. For this, Auger coefficients are set to:

$$C_{nnp} = 1.2 \cdot 10^{-30}\, cm^6/s, \quad C_{nnp} = 8.5 \cdot 10^{-30}\, cm^6/s$$

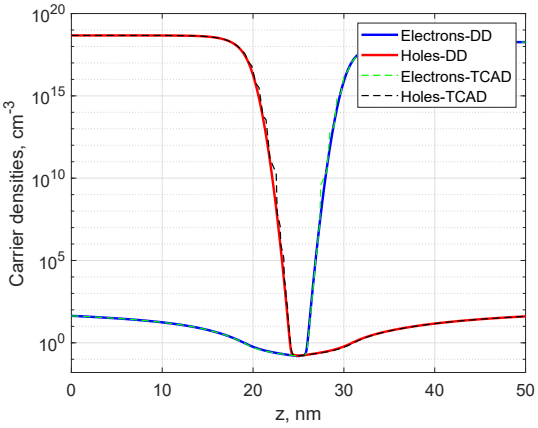
which are the ones inserted in equation (2.13). The obtained rates are plotted in fig. 4.11, where both linear and semi-logarithmic scales are provided in order to perform a more accurate comparison. First of all, it is worth noticing that the order of magnitude of such a rate is rather low to provide a significant contribution to charges generation ($\approx 10^9\, cm^{-3}s^{-1}$) in the TCAD, and almost equal to 0 ($\approx 10^{-9}\, cm^{-3}s^{-1}$) in D1ANA. Nevertheless, by multiplying the latter by a factor $4.1016 \cdot 10^{17}$, a good superimposition of the curves is retrieved. The need of the scaling factor is unexpected, because the used models are exactly the same. An explanation may be sought in the magnitude of the rate itself: since it is low and is acting alone in the junction (remember that no other mechanism other than this is acting), D1ANA struggles in calculating it and the corresponding generated current, therefore it is strongly affected by numerical noise. Beside this, the biggest discrepancies are found at both the ends of the domain, probably because *Sentaurus Device* sets to zero all the recombination rates at the last points before the contacts. On the other hand, the step-like behaviour originally appreciable



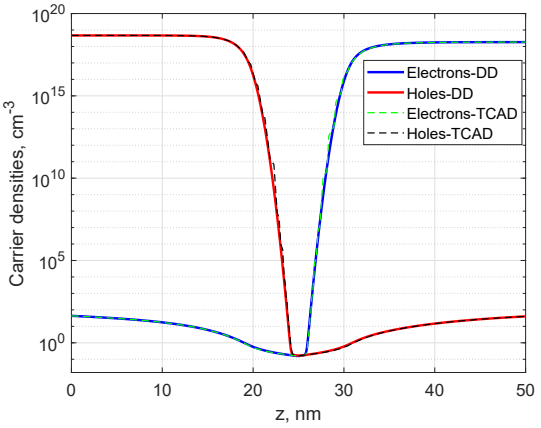
(a) *Energy band diagram at equilibrium*



(b) *Energy band diagram in reverse bias condition (2 V)*

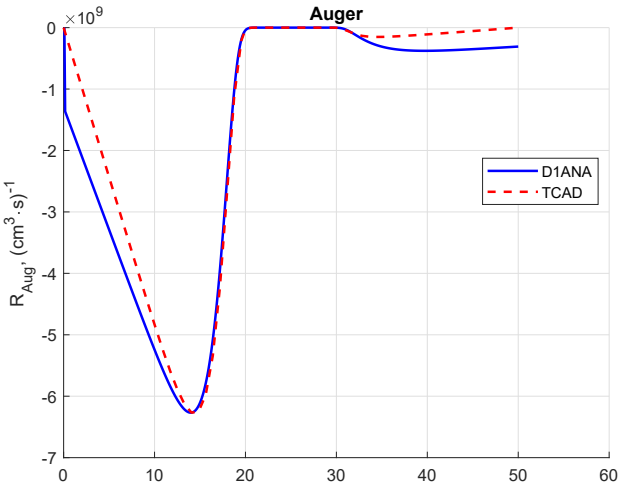


(c) *Charge density*

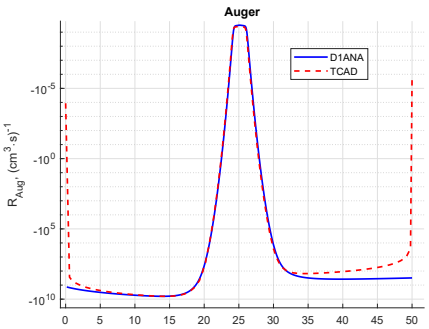


(d) *Charge density with an improved spatial mesh in Sentaurus Device*

Figure 4.10: Energy band diagram at equilibrium and in reverse bias condition, and carrier densities of InGaAsP-based BTJ at reverse bias condition with **PFII** model applied, without any GR mechanism.



(a) *Linear scale*



(b) *Semi-logarithmic scale*

Figure 4.11: **Auger** recombination rate obtained in D1ANA and in *Sentaurus Device* TCAD, in reverse bias condition. Both linear and semi-logarithmic scale are displayed.

in the *semilog* scale was again due to the mesh used in the TCAD, which then is easily improved by following the procedure illustrated at the end of section 2.3, obtaining what is displayed in fig. 4.11b. Eventually, the low level of such a recombination mechanism leads to the fact that both energy band diagram and charge densities are basically not modified with respect to figures 4.10b, reporting the results without any recombination mechanism activated, so they are not reported again here.

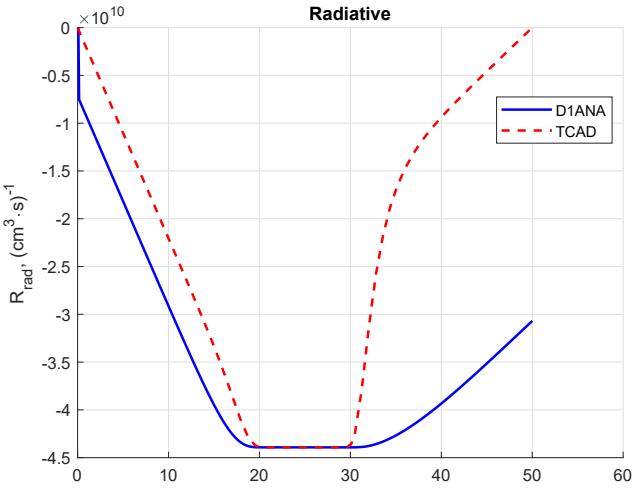
Let's now switch off Auger recombination, and let's move to the analysis of the **radiative** recombination rate. The spontaneous emission coefficient is set to:

$$B_{rad} = 1.1 \cdot 10^{-10}\ cm^3/s$$

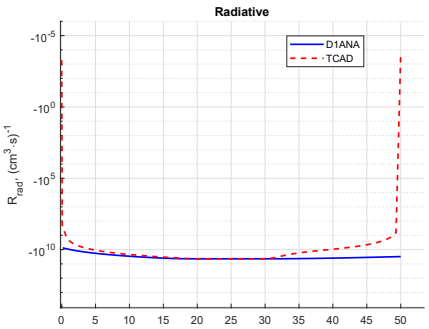
in both the simulators. Again, the results are superimposed in figures 4.12. A perfect overlap of the curves can be appreciated only close to the center of the domain (between 20 and 30 nm), corresponding to the depletion region. The values at the very ends of the junction are again different due to the reason already discussed for Auger case. A more significant difference arises at n-side of the junction, even though the order of magnitude is the same in both the curves. Moreover, just like the Auger recombination, the rate for such a mechanism is on the order of $10^{10}\ cm^{-3}s^{-1}$, such that the energy band diagram and the charge densities are basically not affected by the activation of it.

Shockley-Read-Hall (**SRH**) recombination rate is the next analysed mechanism, for which the following electron and hole lifetimes are used in D1ANA:

$$\tau_n = 5\ ps, \qquad \tau_p = 20\ ps$$



(a) *Linear scale*

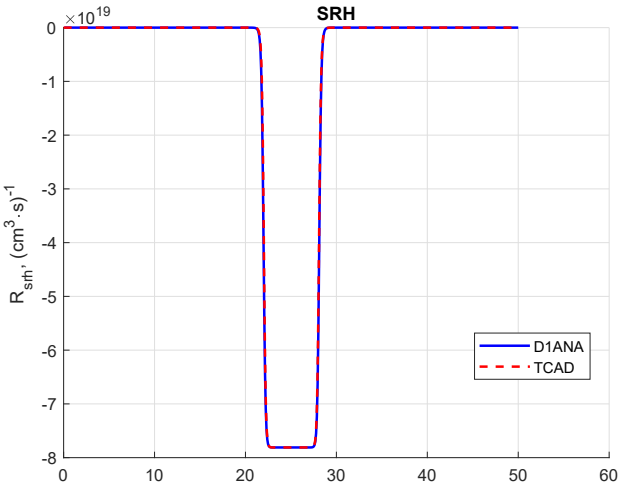


(b) *Semi-logarithmic scale*

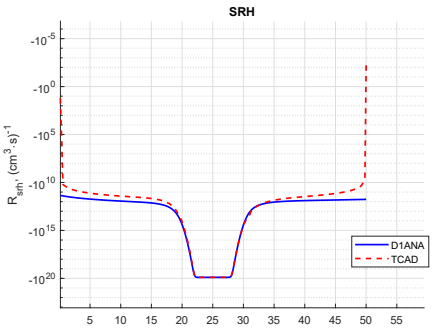
Figure 4.12: **Radiative** recombination rate by spontaneous emission obtained in D1ANA and in *Sentaurus Device* TCAD. Both linear and semi-logarithmic scale are provided.

On the other hand, *Sentaurus Device* makes use of a particular model for the lifetimes computation, such that it is not possible to retrieve the exact values used for the simulations. However, the trap energy level is set to 0 eV in both the simulators. The corresponding rates are reported in figures 4.13, in which the apparent good superimposition must take into account the fact that there is a factor ~ 125 which separates the value of the rate coming from D1ANA to the TCAD one in each point of the domain. This can actually be explained by the lifetimes issue mentioned above. Beside this, the curves are almost identical across the whole domain (especially at the center of the depletion region), beside the usual discrepancy at both ends. A huge difference with respect to Auger and radiative recombinations is that SRH rate reaches values approximately ten orders of magnitude larger ($\approx 10^{19}\ cm^{-3}s^{-1}$). Therefore, both the energy band diagram and the charge densities are heavily affected by such a mechanism.

For this reason, they are reported in figures 4.14. In particular, in the energy band diagram plotted in fig. 4.14a, both the quasi-Fermi levels are slightly different with respect to the case not including SRH mechanism. This is especially true for holes E_{fp}, for which a hump can be appreciated just near the beginning of the n-side depletion region. For what concerns the comparison between the two simulators, due to the already mentioned factor ≈ 100 between the rates, the TCAD curve shows a more pronounced difference with respect to E_{fp} without SRH. This is reflected on the carrier densities curves, reported in fig. 4.14b: at the center of the structure there is a rather large discrepancy between the results of the two simulators. It is important to notice that inside the depletion region, SRH process makes the population densities several orders of magnitude larger with respect to all the previous cases, even though they are



(a) *Linear scale*



(b) *Semi-logarithmic scale*

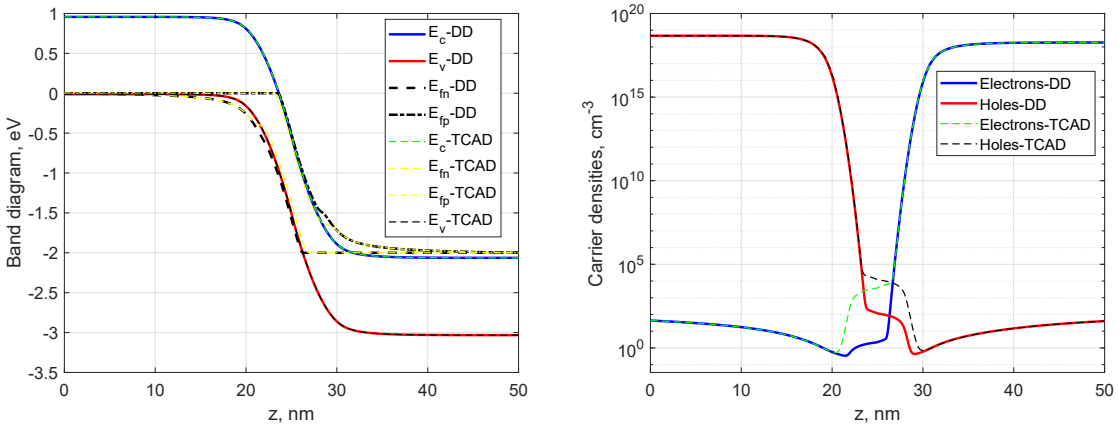
Figure 4.13: Shockley-Read-Hall (**SRH**) recombination rate obtained in D1ANA and in *Sentaurus Device* TCAD. Both linear and semi-logarithmic scale are provided.

not yet large enough to create a significant current into the structure. Since none of the already included mechanism is able to produce large currents in reverse bias condition across the junction, it is even more clear that the introduction of a model including BTBT is needed to correctly model carrier injection in VCSELs by means of a BTJ.

Therefore, let's move to the actual reason for which the tunnel junction is investigated separately: the analysis of the direct interband band-to-band tunneling (**BTBT**). For this reason, the refined spatial mesh plotted in fig. 2.1b is used. The values of the Kane's parameters for Hurkx's model used for the comparison are the ones already mentioned in section 3.2, and not the material dependent ones:

$$A = 3.4 \cdot 10^{21}\ cm^{-1}s^{-1}V^{-2}, \quad B = 22.6 \cdot 10^{6}\ V/cm$$

The resulting rates are plotted in fig. 4.15a, while fig. 4.15b displays the comparison if a coarse mesh were used in *Sentaurus Device*. First of all, notice that the results coming from the two simulators are exactly superimposed. Furthermore, curves show three peaks, corresponding to the regions where the function D becomes significantly different from zero, i.e., when $np \gg n_i^2$, and where the electric field is greater than the critical value given by B. Moreover, BTBT rate is ~ 7 orders of magnitude larger than SRH one. Therefore, when activated in the InGaAsP-based BTJ, it has a huge impact on the junction operations, meaning that all the other mechanisms play a marginal role in a reversely biased tunnel junction. Let's start from the analysis of the band diagram by considering fig. 4.16a. The most clear difference between this case and all the previous ones is that the electron and hole quasi-Fermi levels are no more split, such that they are joined in the whole domain in both D1ANA and *Sentaurus*

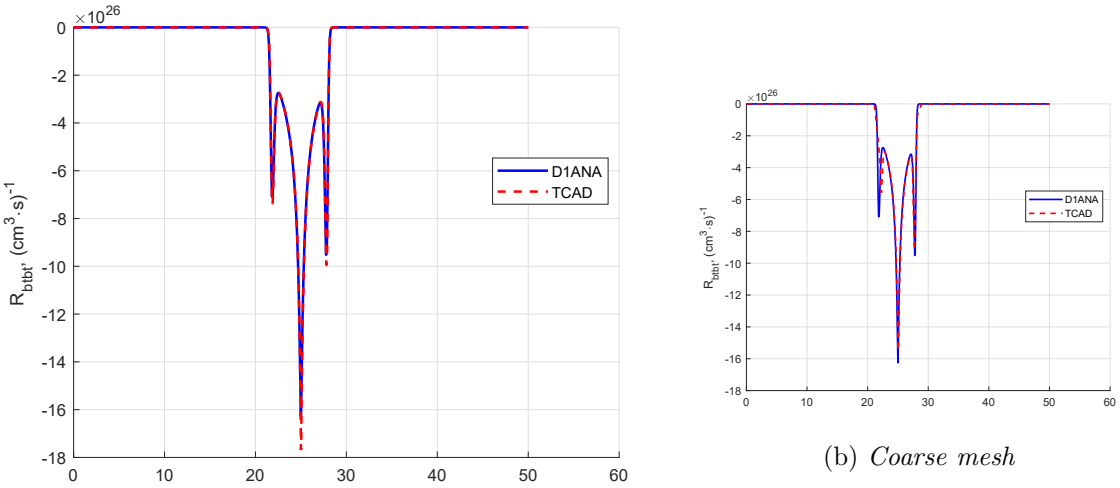


(a) *Energy band diagram.* (b) *Charge density.*

Figure 4.14: Energy band diagram and carrier densities of InGaAsP-based BTJ at reverse bias condition with **PFII** model and SRH recombination rate activated.

Device, including the depletion region. As expected, this has a strong impact on the carrier densities, represented in fig. 4.16b. In fact, inside the depletion region the levels reached by both of them is significantly larger (several orders of magnitude) with respect to the Auger, radiative and SRH cases. In turn, this affects the amount of current able to flow across the tunnel junction in reverse bias. While with the previous mechanisms the current flowing across the BTJ was very close to zero (or anyhow lower than the mathematical noise level), when BTBT is activated the obtained result is different, actually due to the level of density reached around the junction interface. The generated characteristic is plotted in fig. 4.17, where both D1ANA and *Sentaurus Device* TCAD results are plotted and superimposed. The curves are very interesting. They show an initial quick increase in current, which corresponds to the bias point at which valence and conduction bands start to get closer and overlook, such that a significant tunneling window opens, depletion region becomes narrower, electric field increases, eventually overcoming significantly the critical field B. Then, a region of less steep increase begins: this is due to the fact that entering deeper in a reverse bias condition, electrical field increases less, so that BTBT struggles more and more to bringing more charges from one side to the other of the junction.

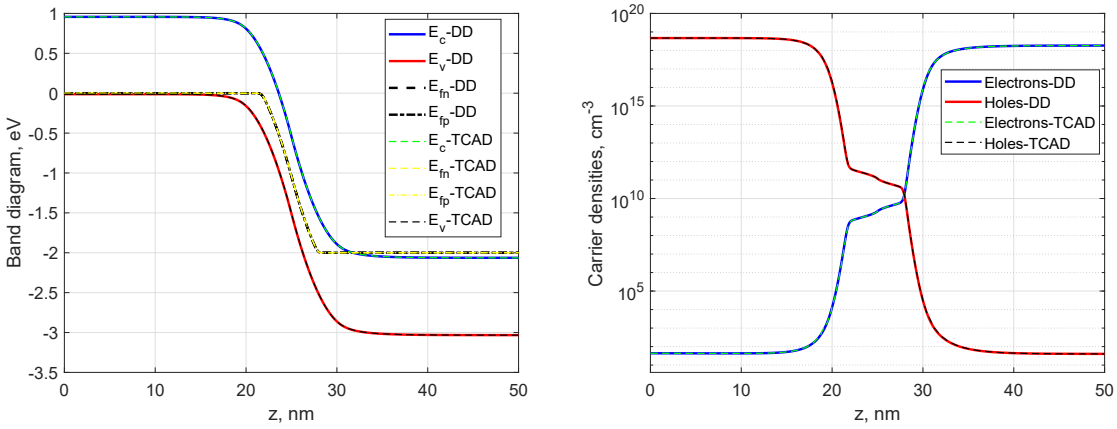
It is worth noticing that the trend of the curve is rather different from the one obtained from the whole VCSEL, already reported in fig. 4.3. In that plot, current remained very close to zero until a certain bias threshold, followed by an almost linear increase in current. This means that the layers of the whole device actually play a role in determining the voltage threshold of the VCSEL, and parasitic resistances strongly change the current profile. Nevertheless, the magnitude of the obtained current is roughly the same in both the plotted characteristics.



(a) *Refined mesh*

(b) *Coarse mesh*

Figure 4.15: Superimposition of the **BTBT** recombination rate curves obtained from D1ANA and *Sentaurus Device* by using "Model `E2`" in reverse bias. The smaller picture reports the rate comparison when a coarse mesh is used in the TCAD.



(a) *Energy band diagram.*

(b) *Charge density.*

Figure 4.16: Energy band diagram and carrier densities of InGaAsP-based BTJ at reverse bias condition with "Model `E2`" for band-to-band tunneling. The curves almost perfectly superimpose because a very fine mesh is used.

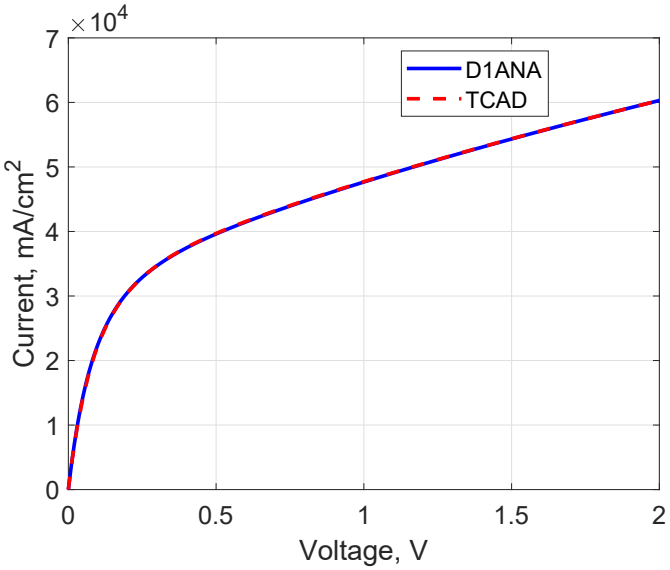


Figure 4.17: Current density vs voltage characteristic of the InGaAsP-based BTJ included in the Lamy's VCSEL. The curves obtained in D1ANA and in TCAD superimpose perfectly.

A final analysis on BTBT local model is then performed, in order to confirm the robustness of the implementation of Hurkx's model inside D1ANA. This involves the use of a different set of values for the Kane's parameters, thus giving a different "weight" to the electric field contribution. In particular, the values used previously are included in the so called "Model `E2`", described in Table 82 in [35, p. 452]; in the same table, another possibility is to choose "Model `E1`", which implies:

$$A = 1.1 \cdot 10^{27}\ cm^{-2}s^{-1}V^{-1}, \quad B = 21.3 \cdot 10^{6}\ V/cm$$

The different units of coefficient A are due to the fact that the exponent of the electric field $\mathcal{E}$ changes and becomes 1, instead of 2. Of course, the same set of parameters is inserted in D1ANA. The corresponding BTBT rate is plotted in fig. 4.18, as always including results from both simulators: the generation peak reached in the TCAD is twice as deep as the D1ANA one, probably due to the different meshes used; more importantly, the order of magnitude $\approx 10^{31}\ cm^{-3}s^{-1}$ is much greater than the one obtained for "Model `E2`". Therefore, energy band diagram (reported in fig. 4.19a) is modified even more strongly than before: in fact, both E_{fn} and E_{fp} are again split and are different even in the neutral sides. As a consequence, the charge densities (fig. 4.19b) take greater values in the depletion region. The correspondence between the curves is very good in both the plots, beside a slight difference in the electron density in the p-side, again confirming that BTBT model is correctly working in D1ANA (band diagram also shows a discrepancy between the e^- quasi-Fermi level). Eventually, the IV curve obtained from this model is reported in fig. 4.20. The results is radically different with respect to "Model `E2`" curve: current now increases parabolically as the applied bias grows, and reaches values up to four orders of magnitude greater than

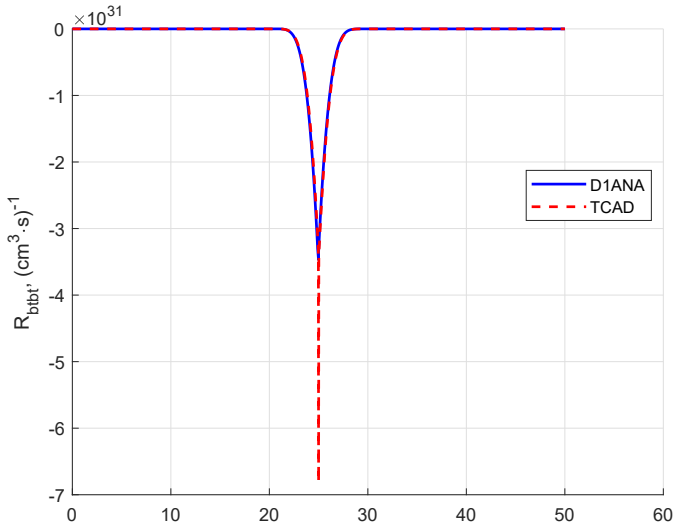


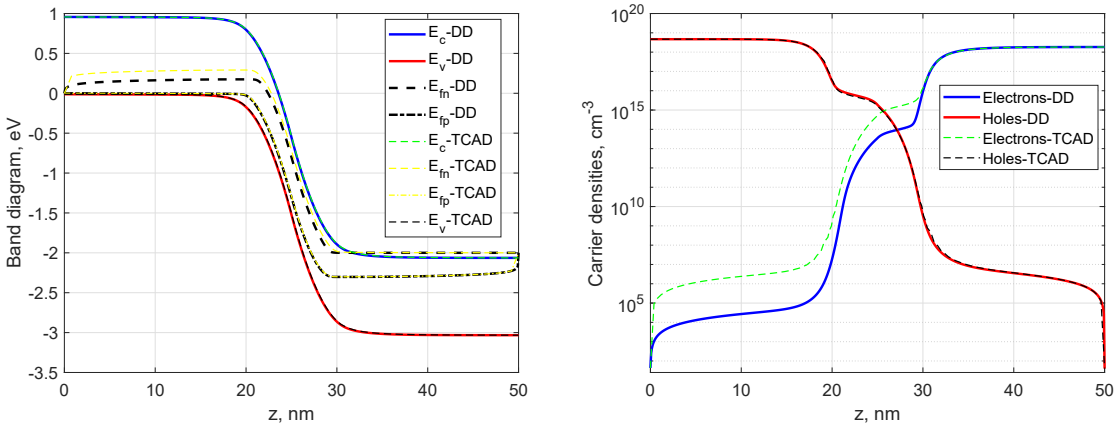
Figure 4.18: Superimposition of the BTBT recombination rate curves obtained from D1ANA and *Sentaurus Device* by using "Model `E1`".

before, which has not any literature confirmation, suggesting that the model is not suited to this case.

4.2 Short-wavelength BTJ-VCSEL

Since the model exploited for the low band gap tunnel junctions returns results in a good agreement with literature for large-wavelength VCSELs, let's now try to investigate short-wavelength devices including BTJ with Hurkx's model. This is done with the aim of producing some new results, since literature is plenty of works talking about BTJ in 1310-1550 nm VCSEL, but lacks of results for shorter wavelength devices (GaAs-based BTJ VCSELs, 850-920 nm). The main difference in the structure of the latter with respect to the former is the energy gap of the used semiconductor alloy, which becomes much wider (in [40], E_g in the BTJ was ≈ 0.97 eV, while GaAs and AlGaAs alloys gap is greater than 1.4 eV). The consequence of such a difference is straightforward: as described in chapter 3, tunneling probability heavily depends on the value of the gap. This means that an AlGaAs BTJ surely has a reduced tunneling current with respect to an InGaAsP one, if Hurkx's model is taken into account. The results presented afterwards agree with this observation.

For the analysis of such devices, data coming from Chalmers University of Technology (contacted by dr. Pierluigi Debernardi) are used as a starting point to generate the investigated structures. Besides the buried tunnel junctions, in these are also included oxide apertures (easily realized here since large portions of Al are included in the device



(a) *Energy band diagram.* (b) *Charge density.*

Figure 4.19: Energy band diagram and carrier densities of InGaAsP-based BTJ at reverse bias condition with "Model `E1`" for band-to-band tunneling.

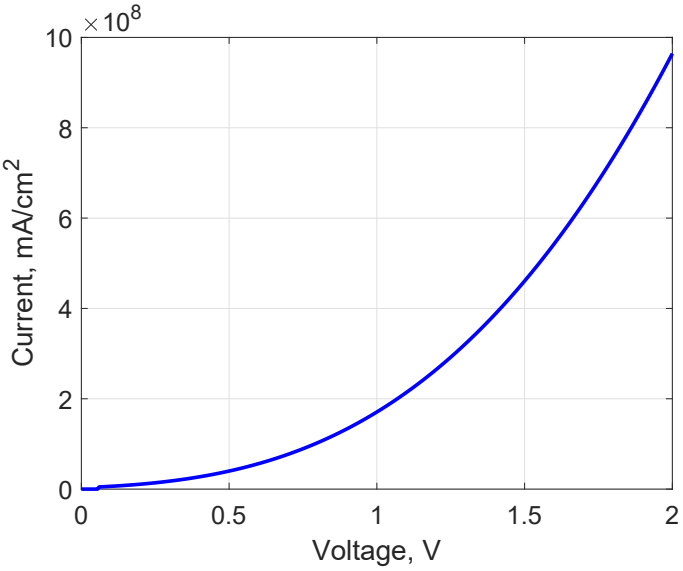


Figure 4.20: IV plot of InGaAsP-based BTJ, exploiting "Model `E1`".

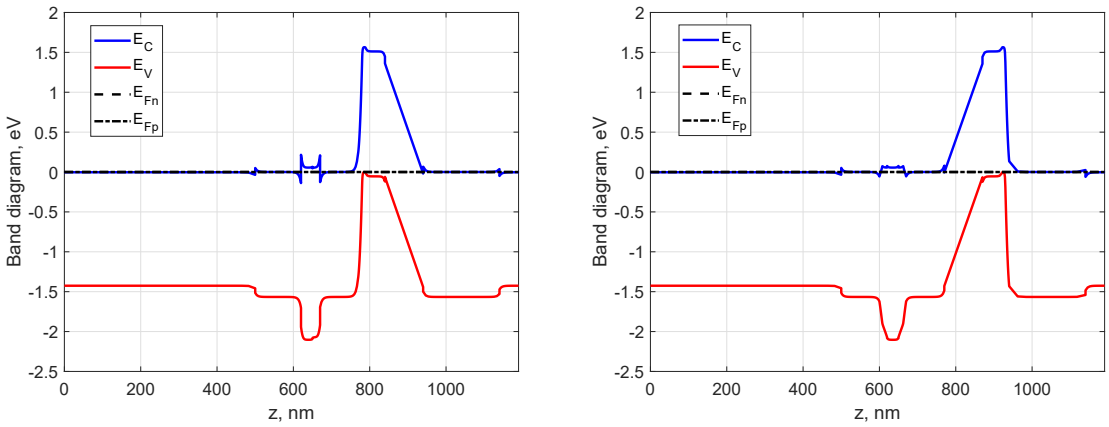
layers), used to better confine the current to a well defined injection area and prevent surface recombination. The details of all the layers are not shown in a table, unlike done in table 4.1 for Lamy's structure, since the devices suggested by Chalmers University are still object of study and experimental results about them have not been published yet. However, they are designed to produce an output light at $\lambda = 850$ nm. Therefore, the ternary alloy $Al_xGa_{1-x}As$ is used, with different values of molar fraction x in each layer. Every alloy parameter is computed by means of a specific linear approximation coming from the AlGaAs section of the NSM archive [24], because now the indirect gap GaAs may lead to major problems if a simple Vegard's approximation similar to (4.1) was used. Anderson's rule is used again here by taking a CBO = 0.65, but now the reference value of E_g and χ is provided by GaAs.

Similarly to the long-wavelength structure analysed in the previous section, the use of a BTJ allows to reduce the amount of long p-doped layers and contacts, so that the overall resistance is decreased and efficiency is improved. Furthermore, the molar fraction in the tunnel junction is tailored to obtain a energy band gap of $E_g = 1.566$ eV, much greater with respect to the InGaAsP-based BTJ. Two configurations are simulated, and in both cases the overall investigated VCSELs are longer than Lamy's one, because a much wider substrate (≈ 500 nm long) is considered in the geometry. Active region is not included in these simulations. For the analysis of AlGaAs BTJ presented afterwards, different doping levels are tested (symmetric and asymmetric), to evaluate the impact of such a parameter on BTBT mechanism in wide band gap materials.

4.2.1 Results

For the two short-wavelength VCSELs, both equilibrium and forward bias conditions are simulated, just like before. The only difference between the two structures is that the BTJ layers have a different order, such that in one case n-side is grown on p-doped layer, and in the other one the opposite occurs. Because of the exchange of position of the pn junction, also some other layers have a different order, but the devices should work approximately in the same way. Moreover, notice that since the tunnel junction properly works in reverse bias, in the two structures the bias is applied at opposite contacts.

At **equilibrium**, in both the devices does not flow any current, and the obtained energy band diagrams are plotted in fig. 4.21. Just like in the long-wavelength VCSEL case, conduction and valence bands in the tunnel junction do not provide any tunneling path to the carriers, such that tunneling current is null. The regions between 600 and 700 nm are where the oxide apertures are actually placed: a graded doping produces this well in the band diagram. Again, since both contacts are both n-doped, at equilibrium their energy levels are exactly the same. The quasi-Fermi levels lie at the same



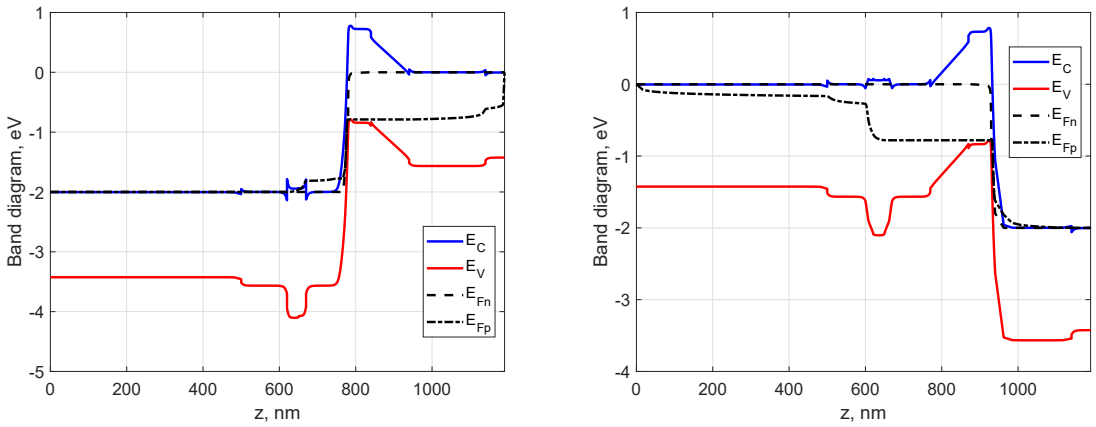
(a) p^+ on n^+ structure (b) n^+ on p^+ structure

Figure 4.21: Energy band diagram of the short-wavelength VCSELs designed by Chalmers University of Technology at **equilibrium** condition.

energy of the conduction band in the neutral regions, i.e., at 0 eV in the whole domain.

However, when a **forward** bias is applied, bringing the BTJ in reverse bias condition, the devices start to work. The corresponding energy band diagrams are reported in fig. 4.22. As already mentioned, the bias is applied on the left contact for the p^+ on n^+ structure in the left-hand side figure, while it is applied at the right contact for the n^+ on p^+ structure. From these, it is clear that the bands corresponding to the tunnel junctions behave aa expected, providing a tunneling path to the holes which now are able to flow across the structure by means of band-to-band tunneling, thus generating a non-zero current. In fact, electron and hole quasi-Fermi levels superimpose in the BTJ region, whereas the E_{fp} separates from E_{fn}, elsewhere: this is actually the reason for which the tunnel junctions are introduced in the structures, that are able to inject holes effectively in absence of p-contacts. Let's remark that the simulation of both structures is not as fast as in the long-wavelength case. This is due to the low current level reached, that makes D1ANA struggle to reach the convergence. A confirmation is given by the fact that by reducing the BTBT rate in the InGaAsP BTJ-VCSEL simulation, a similar issue would be encountered.

The IV curves of VCSELs are of course generated, and superimposed in fig. 4.23. Several observations are needed for such a plot. First of all, as expected, the two curves as almost identical, beside the fact that n^+ on p^+ structure reaches slightly higher values for larger applied bias. Then, a similar overall behaviour with respect to the homologous IV plot for the long-wavelength VCSEL can be appreciated: at low bias the current is almost null; on the other hand, when it overcomes a value of ≈ 0.6 V, a steep increase can be noticed, followed by a more gentle linear increase, due to the overall resistance of the device, just like in the InGaAsP system. Nevertheless, it must be highlighted that the currents magnitude is considerably lower, displaying



(a) p^+ on n^+ structure (b) n^+ on p^+ structure

Figure 4.22: Energy band diagram of the short-wavelength VCSELs designed by University of Chalmers in forward bias condition, with a 2 V bias applied.

values four and even five orders of magnitude smaller. The most clear consequence of this is that the IV plots result to be rather noisy. Therefore, D1ANA code struggles to converge for some values of the voltage sweep, especially near threshold. This is a clear sign that the used local model is not well suited to simulate wide band gap material systems. A comparison with literature [44] and with the experimental results provided by Chalmers University confirms it, since values not so far from InGaAsP device ones are suggested in both of them ($\approx 20 \cdot 10^3$ mA/cm^2).

For this reason, a deeper analysis of the BTJ characteristic is even more needed than the InGaAsP-based one, to understand exactly where Hurkx's model fails. Therefore, in the followings, a version of the code which does not require the already mentioned use of scaling is exploited. It is computationally heavier than the original code, but it possibly allows to obtain slightly more accurate results.

4.2.2 AlGaAs BTJ

Just like in the case of InGaAsP-based BTJ, let's go deeper into the details of the wide gap ($E_g = 1.566$ eV) tunnel junction included in the short-wavelength VCSEL previously described. However, all the analysis about charge density models and recombination mechanisms already performed for the previous BTJ is left out, so that only BTBT is investigated in this case. All the simulations are performed by taking Fermi-Dirac distribution (2.7) and assuming incomplete ionization of dopants model (2.10); reverse bias condition is obtained by applying a 2 V voltage to the n-side of the junction, as usual.

The doping levels initially used for the AlGaAs-based BTJ are supposed to be

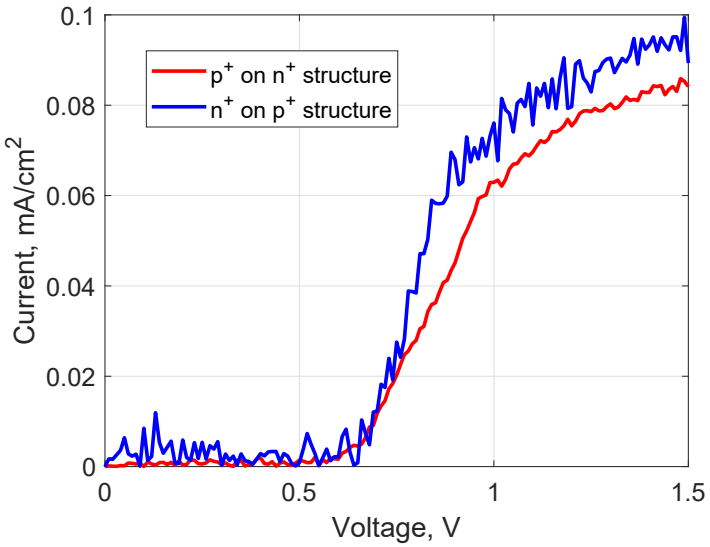


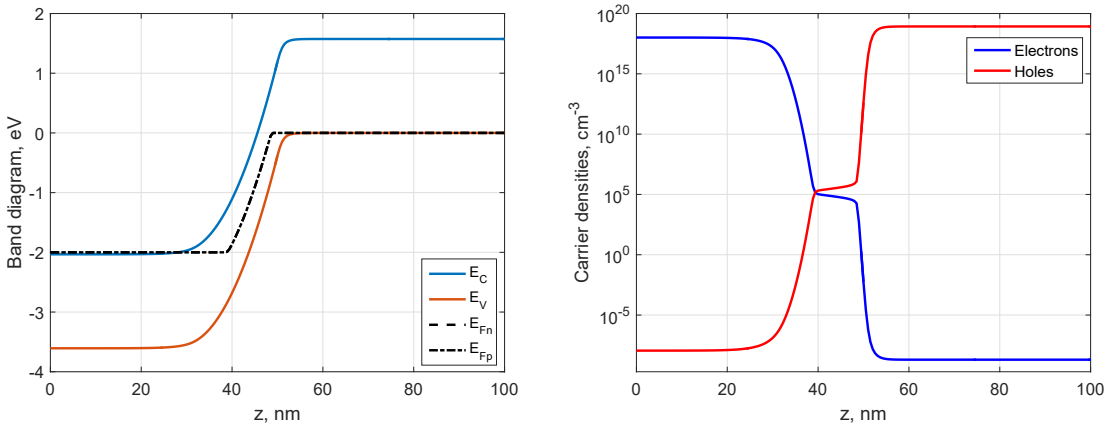
Figure 4.23: Current vs voltage comparison of Chalmers' VCSELs for a forward bias going from 0 to 2 V.

asymmetric, with $N_D = 10^{19}$ cm^{-3} (where Te is the dopant, whose activation energy in AlGaAs is $\Delta E_D = 30$ meV) and $N_A = 10^{20}$ cm^{-3} (with C as a dopant species, $\Delta E_A = 26$ meV). All the recombination mechanisms are now activated in D1ANA, including the BTBT with Kane's parameters obtained by using (3.22) and (3.23):

$$A = 2.03 \cdot 10^{20}\, cm^{-1}s^{-1}V^{-2}, \quad B = 29.1 \cdot 10^{6}\, V/cm$$

The energy band diagram obtained at **reverse** bias condition is reported in fig. 4.24a: since the doping levels are not equal, the depletion region results to be asymmetric as well. Furthermore, by comparing it with the InGaAsP-based BTJ band diagram in fig. 4.16a, it is clear that the bands are much more far away in space, possibly resulting in a lower junction electric field. The charge distribution is similar to the one obtained in the InGaAsP system, and is displayed in fig. 4.24b. This, together with the wider energy gap, surely leads to a significantly lower tunneling rate, as plotted in fig. 4.25. In the latter, one has to notice that the generation peak reaches a value $\approx 5.3 \cdot 10^{21}\, cm^{-3}s^{-1}$, five orders of magnitude lower than in the InGaAsP BTJ, positioned approximately at the center of the depletion region, with just one side peak to the asymmetric doping profile, i.e., where $np > n_i^2$ and the electric field becomes larger than the critical value given by B.

The results just explained have a strong impact on the current-voltage characteristic of the junction, which is plotted in fig. 4.26. In fact, as already mentioned for the IV curve of the whole device, even though the behaviour of the current is basically identical to the InGaAsP one, the values reached are considerably lower (more than five orders of magnitude), such that noise plays an important role (as clear from the result in fig. 4.26b, that comes from the version of the code including scaling). Noise



(a) *Band diagram* (b) *Charge densities*

Figure 4.24: Energy band diagram and charge densities of the **asymmetrically** doped ($N_D = 10^{19}$ cm^{-3}, $N_A = 10^{20}$ cm^{-3}) AlGaAs-based BTJ at reverse bias condition.

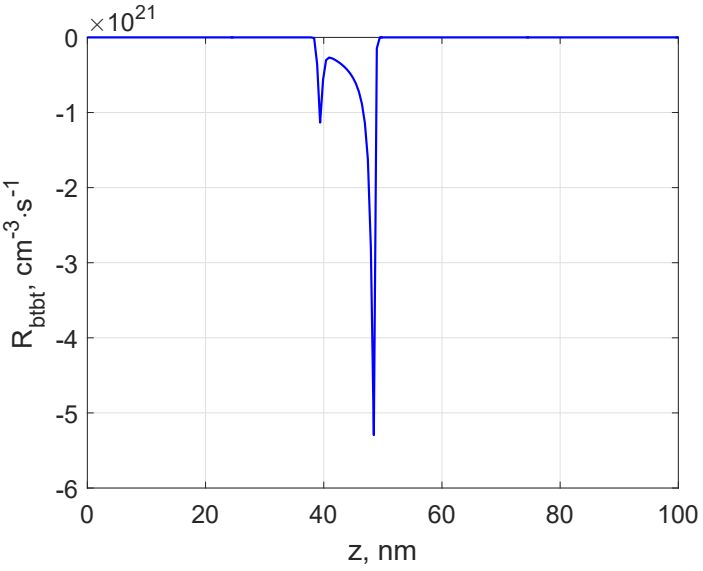
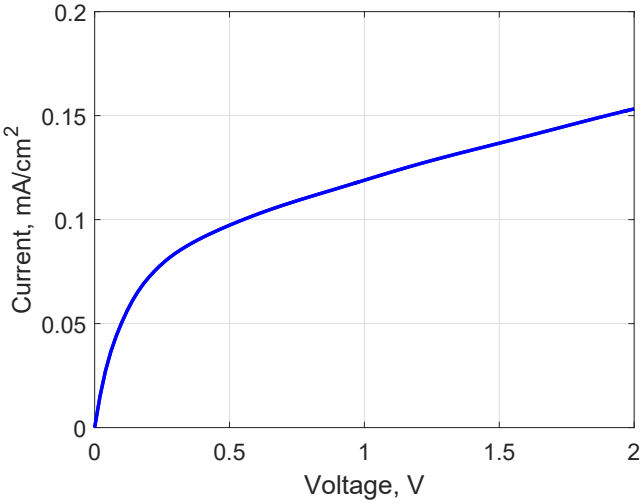
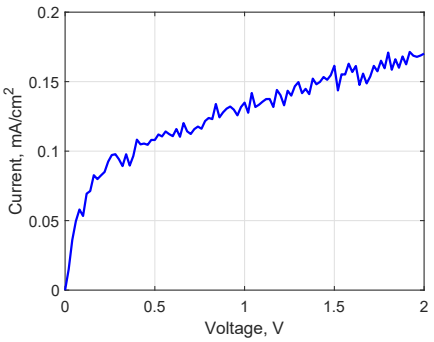


Figure 4.25: BTBT rate across the asymmetrically doped AlGAsAs-based BTJ.



(a) *IV Curve coming from the original version of the code*



(b) *IV Curve coming from the improved version of the code*

Figure 4.26: Current density vs voltage curve of AlGaAs-based BTJ included in Chalmers University of Technology VCSEL for a reverse bias going from 0 to 2 V, computed with different versions D1ANA in case of asymmetric doping profiles.

level is reduced by using the improved version of the code which avoids scaling, thus yielding the curve in fig. 4.26a, at the cost of a slower simulation.

All the analysis showed until this point highlight the fact that the current flowing in AlGaAs-based device in reverse bias due to band-to-band tunneling is much lower with respect to the structure based on lower E_g semiconductor alloys, when modelled with a local model. Therefore, it is important to understand why Hurkx's model predicts such a low level of current, while experimental results coming from Chalmers University of Technology claim for much larger values. To this aim, several material parameters are adjusted one at a time, in order to understand which is the critical one.

The first attempts involves a direct adjustment of the BTBT rate. This is obtained by simply decreasing by two orders of magnitude the **critical electric field** at which tunneling becomes significant:

$$B = 29.1 \cdot 10^4 \; V/cm$$

Due to the exponential dependence of R_{btbt} on it, an increase of the GR rate should be appreciated. Therefore, a simulation (performed with the code including scaling) at reverse bias is performed. Surprisingly, the resulting current has an identical profile with respect to the previous case. From this, it is already clear that Hurkx's model is very limited, and completely unsuited for such a material system. The explanation to such a result can be likely retrieved in the fact that a local R_{btbt} is not able to grasp all the details of such a structure, and the wide E_g imposes a limitation to the maximum reachable current in a DD simulator exploiting this kind of model.

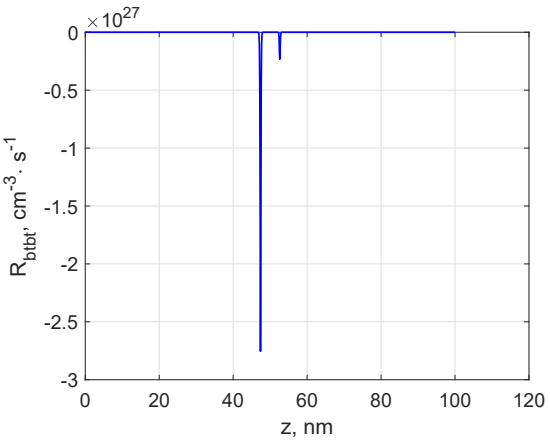
However, let's now move to another approach. In the previous attempt, the **doping levels** of the junction were asymmetric in p and n-sides. Here, they are set to be equal to the values used for the InGaAsP-based BTJ:

$$N_A = N_D = 5 \cdot 10^{19}\ cm^{-3}$$

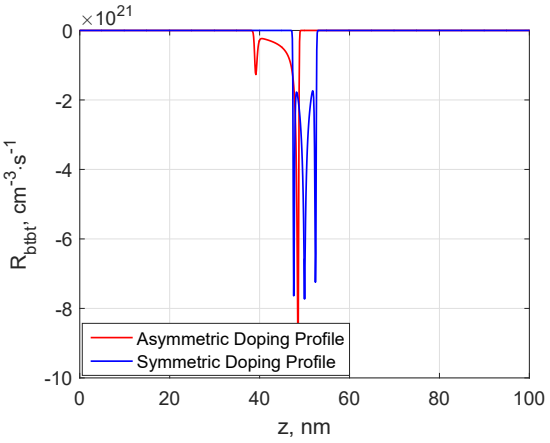
while B returns to its previous value. It is didactic to show the BTBT GR rate computed by the original version of the code, that is plotted in fig. 4.27a. The position of the main peak changes, and reaches a maximum value of $\approx 10^{27}\ cm^{-3}s^{-1}$ which is three orders of magnitude larger than the previous cases. However, such a strong value of R_{btbt} holds only in few mesh points, while it returns to be "small" in all the other ones. As a consequence, it basically does not play a significant role in current transport, and this is the reason why the obtained IV curve (not reported here) does not reach values larger than the previous case.

Therefore, it is better to move to the improved version of D1ANA. In fig. 4.27b, the rates for asymmetric (already computed) and symmetric doping profile cases are displayed. The latter, depicted in blue, has a similar trend to the InGaAsP BTJ, despite having a generation rate five orders of magnitude smaller, obviously meaning that a lower tunneling current level must be expected. Such a curve also confirms that the strong and narrow peaks arising in the code including scaling have no impact in computing tunneling current, but they are probably due to a mathematical noise. For this case, it is also interesting to take a look to the energy band diagram of the tunnel junction in reverse bias condition, which is reported in fig. 4.27c. The bands as well resemble a lot the InGaAsP-based BTJ ones, with the difference that the distance between conduction and valence bands in the depletion region, namely the tunneling path, is slightly larger. Moreover, the depletion region is much shorter, passing from ≈ 20 nm to ≈ 5 nm. This should have an impact on the computed current, because it modifies the mesh points at which function D (3.25) is different from zero. Eventually, the IV curve is plotted in fig. 4.27d, where a comparison with the asymmetric doping case is proposed. It is clear that the current is almost doubled if a symmetric doping is injected; nevertheless this is not enough to bring current to acceptable values, useful to make the BTJ-VCSEL working properly. However, it is interesting to notice that a significant modification of the bands in the depletion region leads to a stronger variation in current with respect to a simple increase of the BTBT generation rate.

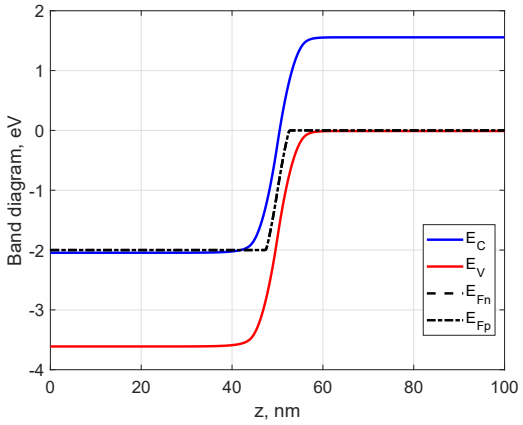
As seen until now, the adjustment of the Hurkx's model parameters and of the doping levels has not yielded the expected results on the current. Therefore, the low values of the latter must be due to one or more material parameters that "kill" the carrier motion inside the tunnel junction. In table 4.2, the main parameters of the two investigated BTJ materials are reported: the biggest differences arise for the energy gap E_g, the electron affinity χ and the electron mobility μ_n. Therefore, simulations on AlGaAs BTJ are performed for each mentioned parameter (keeping the original asymmetric doping levels and B coefficient), by setting it to be equal to its homologous



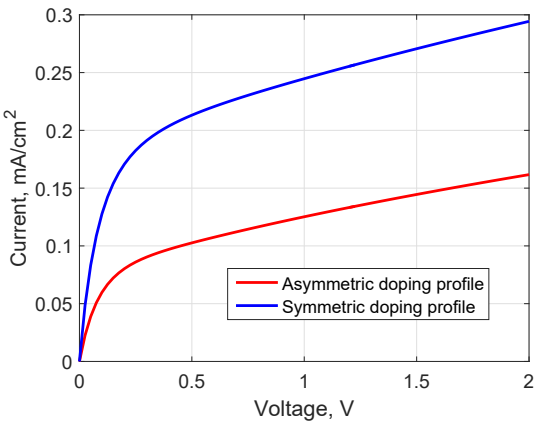
(a) *BTBT generation rate (symmetric case only), with the code including scaling; noise plays a significant role here.*



(b) *BTBT generation rates, obtained with the improved version of D1ANA.*



(c) *Energy band diagram*



(d) *IV curves computed for the symmetric and asymmetric cases.*

Figure 4.27: BTBT recombination rate, energy band diagram and corresponding IV curve for **symmetric** doping levels at both sides of the BTJ.

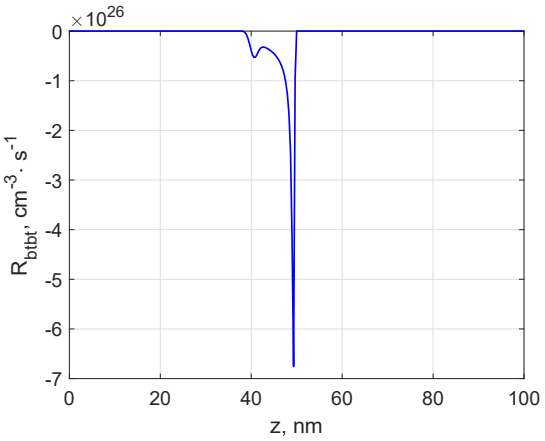
Parameter	Units	$In_{0.66}Ga_{0.34}As_{0.732}P_{0.268}$	$Al_{0.12}Ga_{0.88}As$
E_g	eV	0.97	1.57
χ	eV	4.5	3.98
N_c	cm^{-3}	$5 \cdot 10^{17}$	$5 \cdot 10^{17}$
N_v	cm^{-3}	$8.95 \cdot 10^{18}$	$1.1 \cdot 10^{19}$
m_n^*	m_0	0.0754	0.077
m_p^*	m_0	0.5031	0.58
μ_n	$cm^2 \cdot V^{-1} \cdot s^{-1}$	22400	5500
μ_p	$cm^2 \cdot V^{-1} \cdot s^{-1}$	390	264

Table 4.2: Comparison of the main alloy parameters of the BTJs included in the short and long-wavelength VCSELs.

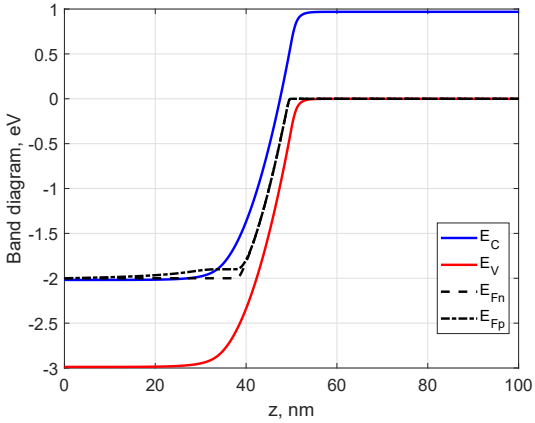
in InGaAsP-based BTJ. While electron affinity modification does not produce any improvement in current levels and electron mobility adjustment has a similar effect of imposing a symmetric doping to the *pn* junction, even a slight decrease of the value of the **energy band gap** has a significant impact on the results.

Let's first try to keep all the parameters untouched, and impose $E_g = E_{g,InGaAsP} =$ 0.97 eV. The same quantities plotted for the previous cases are reported. In particular, generation rate by band-to-band tunneling is plotted in fig. 4.28a: the much smaller E_g leads to a maximum rate five orders of magnitude larger than the original case, again in the whole BTJ region. The results should approach to the small band gap alloy BTJ. In fact, a strong impact on the carrier densities and on the energy band diagram should be expected, mainly because we have deduced that E_g imposes a limit to the maximum reachable current. The latter is displayed in fig. 4.28b. Beside the fact that bands get of course closer in the depletion region (which however remains as wide as before), and thus the tunneling path is shorter, a modification on the holes quasi-Fermi level E_{fp} can be appreciated. Both these factors lead to IV curve (see fig. 4.28c) very similar to the one obtained for the InGaAsP-based tunnel junction (fig. 4.17) analysed in the previous section, at least for what concern the reached values which are on the same order of magnitude ($\approx 10^4$ mA/cm^2). The main difference is that the initial increase is much less steep, and is characterized by a positive curvature, namely a parabolic increase, followed by the usual square root trend.

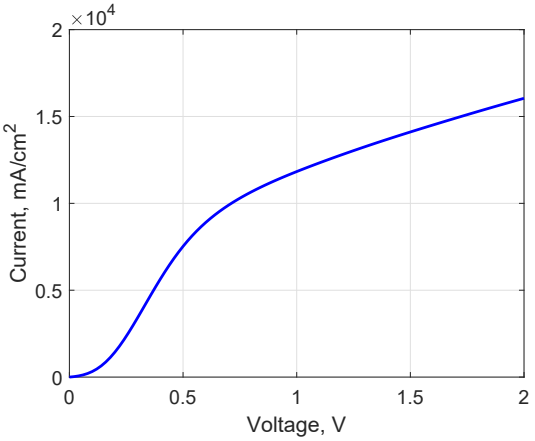
The effect of the energy band gap on tunneling current is worth a further analysis to understand what it actually acts on to allow such levels of carrier generation and tunneling current. To this aim, several simulations are run by reducing each time E_g by 0.1 eV, starting from the original value 1.56 eV. While in the energy band diagram the only clear difference is represented by a decrease of the band gap, it is much more interesting to report the charge densities as a function of E_g. This is done in fig. 4.29a, where gap moves from 1.56 eV to 1.26 eV. Notice that majority carriers are not affected by the gap value; on the other hand, E_g strongly acts on the amount minority



(a) *BTBT generation rate*



(b) *Energy band diagram*



(c) *IV curves of the symmetric and asymmetric cases.*

Figure 4.28: BTBT recombination rate, energy band diagram and corresponding IV curve computed using an **energy gap** equal to the InGaAsP-based BTJ one, with the original asymmetric doping profile.

charges, and more importantly on the population densities inside the depletion region, both surely depending on the enhanced BTBT, depicted in fig. 4.29b. This means that for lower E_g a larger amount of charges is brought from one side to the other of the junction by BTBT mechanism, strongly enhancing the tunnel current, as plotted in fig. 4.29c. Thus, it is clear that even a little decrease of the gap leads to much larger currents, which always follow the usual increasing behaviour.

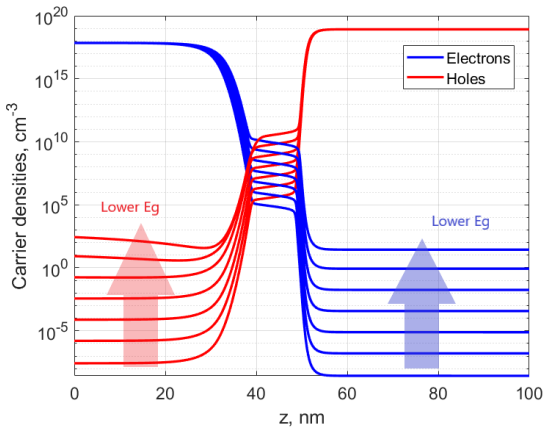
Some final observations on Hurkx's model derive from the whole analysis. First of all, it is straightforward to state that the depletion region length does not play a crucial role in the value reached by the current, even though it can slightly change its value. Furthermore, the generation rate R_{btbt} does not affect the value of the current once it reaches high values (on the other hand, if it is reduced even slightly, almost no current flows any more and the code struggles to converge): what affects the most the tunneling current is the wideness of the region in the junction where it takes those values, such that the strong narrow peaks appreciated in the original version D1ANA actually do not impact on current values at all. As a matter of fact, it is better to have a lower rate in a larger amount of points, rather than obtaining a very strong peak in few nodes of the mesh. However, this is still not enough to reach high values of current: a low gap is also needed, otherwise it remains stuck to noise level. This impacts directly on the charge densities in the depletion region; the same does not happen when one of any other parameter is modified.

Notice that all these remarks concern only the Hurkx's model implemented in D1ANA (and in *Sentaurus Device*) and not the actual physical phenomenon of band-to-band tunneling. In fact, from literature and from Chalmers University of Technology we know that there are several experimental results that display large values of current even for wide gap materials BTJs, like AlGaAs. The conclusion is that for these latter Hurkx's model is no more valid, such that another kind of model is required to reach experimental results. This is exactly what is attempted in the next section, in which is involved the use of the so called non-equilibrium Green's functions formalism to move from a *local* to a *non-local* BTBT model.

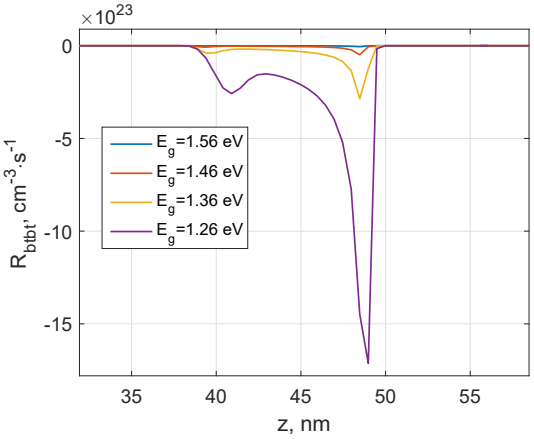
4.3 NEGF simulations

In the previous analysis, it is demonstrated that the simple drift-diffusion simulations on the small band gap (e.g. InGaAsP) BTJ reveal to be accurate when a simple *local* BTBT model is used, whereas completely unsuited for a proper analysis of BTJ current injection in AlGaAs-based VCSELs. These results suggest to move to a radical different way to simulate the tunneling phenomenon, based on non-equilibrium Green's functions (NEGF) formalism.

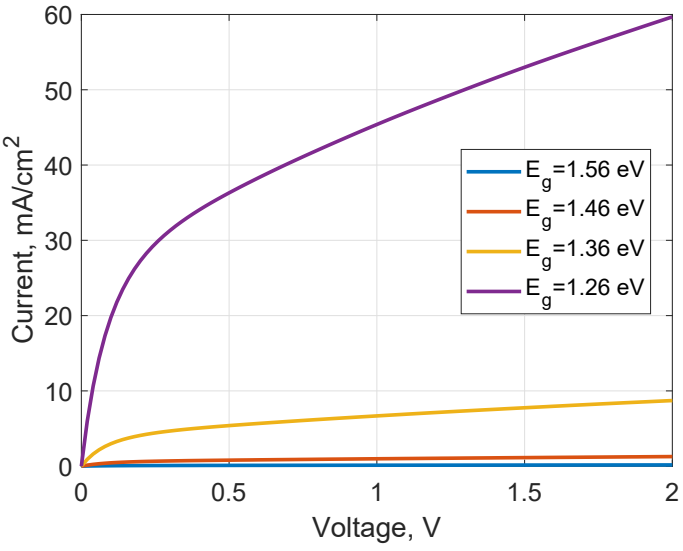
This approach is still based on the modelling of tunneling process as an addition of



(a) *Carrier densities*



(b) *Zoom of BTBT rates in the depletion region*



(c) *IV curves*

Figure 4.29: Charge densities, BTBT and current-voltage curves as a function of the decreasing E_g in the AlGaAs-based BTJ. Doping levels are asymmetric.

a recombination/generation rate R_{btbt}, but NEGF formalism has a more rigorous and solid physical basis, in which tunneling is included naturally, due to the characteristics (material, dimension, doping) of the structure, such that it is capable of provide a *non-local* GR rate. In fact, when non-equilibrium Green's functions are used, it is enough to include the proper system and parameters to the model, so that tunneling is just a consequence of the system band diagram, with a genuine quantum description of BTBT. The novelty of this approach is that NEGF is exploited not only to extract the spectral quantities related to the explored structure, but the resulting tunneling current is used to extract the BTBT non-local GR rate to be included in the DD model in place of Hurkx's rate.

As a matter of fact, NEGF relies on quantum-field theory treating carrier interactions perturbatively by means of self-energies computed within the self-consistent Born approximation (SCBA). Furthermore, the *ballistic* nature of BTBT (tunneling probability is large only in nanoscale dimensions) allows to assume a coherent limit, thus reducing the otherwise staggering computational cost. Basically, the purpose is to retrieve a *non-local* tunneling rate from a 4-band NEGF formalism for the BTJ region only, and then include the obtained results in D1ANA and eventually evaluate the corresponding current-voltage curve. The eigenmodes coming from NEGF approach are evaluated here in a real-space representation, in order to take into account tunneling process from any energy value and not only from band-edges, as done in mode-space representation [45, 46]. Therefore, mode-order reduction [47] is not needed for ballistic BTBT simulations.

Nevertheless, NEGF formalism also has some limitations: in fact, long systems are not simulated because they would require a huge amount of time to be completed. This is the reason why only the BTJ (only few tens of *nm* long) is investigated by means of such an approach. Moreover, the obtained results are often not easy to be interpreted, since a tunneling rate is not returned directly, but only spectral quantities are retrieved, that integrated in space and energy return macroscopic observables, such as current or carrier densities. In particular, this formalism provides a microscopic theory of quantum transport [48] (a simplified treatment is presented in [49]).

However, the nature of the investigated structures allows to restrict the attention to the coherent limit and to the steady-state regime. In this framework, under equilibrium conditions electron density is obtained by filling up the states according to the Fermi function. On the contrary, when a bias is applied to the ballistic device, a flux of carriers naturally builds up between the contacts by tunneling (even though some carriers are reflected back by the potential barrier arisen in reverse bias condition). A device in this condition is very far from equilibrium, given that no scattering mechanism is activated to drive it back to thermodynamic equilibrium. Since tunneling occurs, the different states $\psi_\alpha(r)$ of the Schrödinger equation are occupied in a correlated fashion described

by a density matrix

$$\rho(r, r') = \sum_{\alpha,\beta} \psi_\alpha(r)\psi^*_\beta(r')\tilde{\rho}_{\alpha\beta} \tag{4.2}$$

where α and β indicates different states, and $\tilde{\rho}_{\alpha\beta}$ is a correlation term written in the eigenstate representation. The density matrix actually is the unknown of NEGF formalism, and is needed to compute all the quantities of interest. In particular, at **equilibrium** the sum over its diagonal terms returns the electron density

$$n(r) = \sum_{\alpha} |\psi_\alpha(r)|^2 f_{FD}(E_\alpha - \mu) \tag{4.3}$$

where $f_{FD}(E_\alpha - \mu)$ is the Fermi function, μ is the Fermi level and E_α denotes the eigenenergies of the longitudinal Hamiltonian H: $H\xi = EM\xi$, where M is a matrix coming from the finite-element modelling used to discretize the 1D longitudinal domain, in a very similar fashion already presented for the drift-diffusion model. Such an affinity also reveals to be crucial when the obtained results from NEGF have to be included in D1ANA. Therefore, density matrix in real space can be written performing a unitary transformation:

$$[\rho] = [V]\,[\tilde{\rho}]\,[V]^\dagger \tag{4.4}$$

where $[\rho]$ is a matrix whose columns denote the eigenvectors of $M^{-1}H$, such that density matrix takes the following form:

$$[\rho] = f_{FD}\left(\left[M^{-1}H - \mu I\right]\right) \tag{4.5}$$

$$= \int dE f_{FD}(E-\mu)\delta\left([EM - H]\right) \tag{4.6}$$

where δ is the Kronecker function that can be defined as follows:

$$\delta\left([EM - H]\right) = \frac{i}{2\pi}\left([(E + i\eta)\,M - H]^{-1} - [(E - i\eta)\,M - H]^{-1}\right) \tag{4.7}$$

with η is positive infinitesimal called *contact broadening* parameter that is neglected in the next simulations since it is required only when bound states are involved (while an open system is investigated in this work). From (4.7), *retarded* and *advanced* Green's function are retrieved:

$$G^R(E) = [(E + i\eta)\,M - H]^{-1} \tag{4.8a}$$

$$G^A(E) = (G^R(E))^\dagger = [(E - i\eta)\,M - H]^{-1} \tag{4.8b}$$

which in turn are used to define the *spectral* function:

$$[\mathcal{A}(E)] = i\left(\left[G^R(E)\right] - \left[G^R(E)\right]^\dagger\right) \tag{4.9}$$

that can be seen as the available density of states, such that

$$[\rho] = \int \frac{dE}{2\pi} f_{FD}(E-\mu)\,[\mathcal{A}(E)] \tag{4.10}$$

Far from equilibrium, e.g. when an external tension is applied to the device, μ split into two values because of the action of different bias at the two contacts. However, by neglecting scattering processes (and this is exactly what is done when BTJ BTBT is simulated), left and right eigenstates can be assumed to stay in equilibrium with left and right contacts, respectively. Therefore, the density matrix can be written as [49]:

$$[\rho] = \int \frac{dE}{2\pi} \left[f_{FD}(E - \mu_1) \left[\mathcal{A}_1(E) \right] - f_{FD}(E - \mu_2) \left[\mathcal{A}_2(E) \right] \right] \tag{4.11}$$

where $[\mathcal{A}_{1,2}(E)] = G^R(E)\Gamma_{1,2}(E)G^A(E)$. In the latter,

$$\Gamma_{1,2} = i \left[\Sigma_{1,2}^{RB} - (\Sigma_{1,2}^{RB})^\dagger \right] \tag{4.12}$$

are the so called *broadening* functions. The main advantage of NEGF method is contained in the function Σ^{RB}, namely the *retarded boundary* self-energy, a $nn \times nn$ matrix with all zero entries beside the upper-left and the lower-right corner elements accounting for left and right leads. Retarded boundary self-energy is needed to compute the retarded Green's function of the device, which has a similar form of (4.8a):

$$G_D^R(E) = \left[EM_D - H_D - \Sigma^{RB} \right]^{-1} \tag{4.13}$$

which is known as *Dyson's equation* and shows that the coupling between device and its carrier reservoirs can be described by simply adding the self-energy matrix to the device Hamiltonian. Subscript D stands for "device", and it is dropped in the following expressions.

Dyson's equation must be evaluated numerically by moving from continuous space variables to a discrete basis and by Fourier transforming from time to energy domain (this is allowed in steady-state conditions), allowing to write the needed Green's functions as functions of $\underline{k}$ [50, 51]:

$$G^R(\underline{k}, E) = \left[EM - H(\underline{k}) - \Sigma^R(\underline{k}, E) \right]^{-1} \tag{4.14a}$$

$$G^A(\underline{k}, E) = G^R(\underline{k}, E)^\dagger \tag{4.14b}$$

$$G^{\lessgtr}(\underline{k}, E) = G^R(\underline{k}, E)\Sigma^{\lessgtr}(\underline{k}, E)G^A(\underline{k}, E) \tag{4.14c}$$

where (4.14c) is the *Keldysh's equation*, which defines the lesser and greater Green's functions. The lesser and greater self-energies are computed with the fluctuation-dissipation theorem:

$$\Sigma^<(\underline{k}, E) = i f_{FD}(E - E_F)\Gamma(\underline{k}, E) \tag{4.15a}$$

$$\Sigma^>(\underline{k}, E) = i \left[f_{FD}(E - E_F) - 1 \right] \Gamma(\underline{k}, E) \tag{4.15b}$$

Having defined all the needed Green's functions and the related quantities, let's now express the quantities of interest as functions of the latter. In particular, spectrally-resolved electron density in steady-state is computed by taking the diagonal elements

of the lesser Green's function [52] in the continuum limit, and then assuming axial approximation ($\underline{k} \approx k$), eventually obtaining:

$$n(z_i, E) = -\frac{i}{2\pi} \int dk\, k G^{<}_{i,j}(k, E) \approx i \frac{\Delta k}{2\pi} \sum_{k_j} k_j G^{<}_{i,i}(k_j, E) \tag{4.16}$$

that can be integrated over energy to retrieve the level of electron density along the transport direction z:

$$n(z) = \int \frac{dE}{2\pi} n(z, E) \approx \frac{\Delta E}{2\pi} \sum_{j} n(z, E_j) \tag{4.17}$$

In the NEGF code, Δk and ΔE are two parameters used to define the spatial and energy grid. Current density is computed element-wise, as follows [51]:

$$J_{j+1/2}(E) = \frac{e}{h} \Delta k \sum_{k_j} k_j [(H_{j,j+1}(k_j) - E M_{j,j+1}) G^{<}_{j+1,j}(k_j, E) - G^{<}_{j,j+1}(k_j, E)(H_{j,j+1}(k_j) - E M_{j,j+1})] \tag{4.18}$$

This approach is really powerful when dealing with ultrashort devices, since ballistic assumption permits to divide the contact-device-contact structure into two separated groups whose eigenstates are assumed to remain in equilibrium with the corresponding contacts. This leads to a huge simplification of the calculations, since it actually allows to treat a non-equilibrium problem using equilibrium statistical mechanisms.

4.3.1 AlGaAs BTJ

NEGF formalism is therefore applied to the BTJ included in the short-wavelength VCSEL just presented. The followed procedure is quite cumbersome. First of all, the simple geometry of the structure is produced in D1ANA, one file for each bias point of the voltage sweep. These are the starting point for the NEGF simulations. For each bias point, an output file is generated, and then exported in MATLAB, where spectral quantities can be plotted and investigated. As mentioned, it is not possible to directly extract a BTBT rate as a NEGF output, so the computed spectral current density (4.18) is first integrated in energy domain:

$$J_{NEGF} = \frac{1}{2\pi} \int_{E_{min}}^{E_{max}} J_{spectral}(E) dE \tag{4.19}$$

such that an overall current density (which due to charge conservation is constant across the whole transport direction of the system) is retrieved. The parameters defining the k and energy grid which enter into the expressions (4.16), (4.17) and (4.18) take the values of $\Delta k = 0.0455$ nm^{-1} (with $k_{max} = 2$ nm^{-1}) and $\Delta E = 5$ meV, respectively.

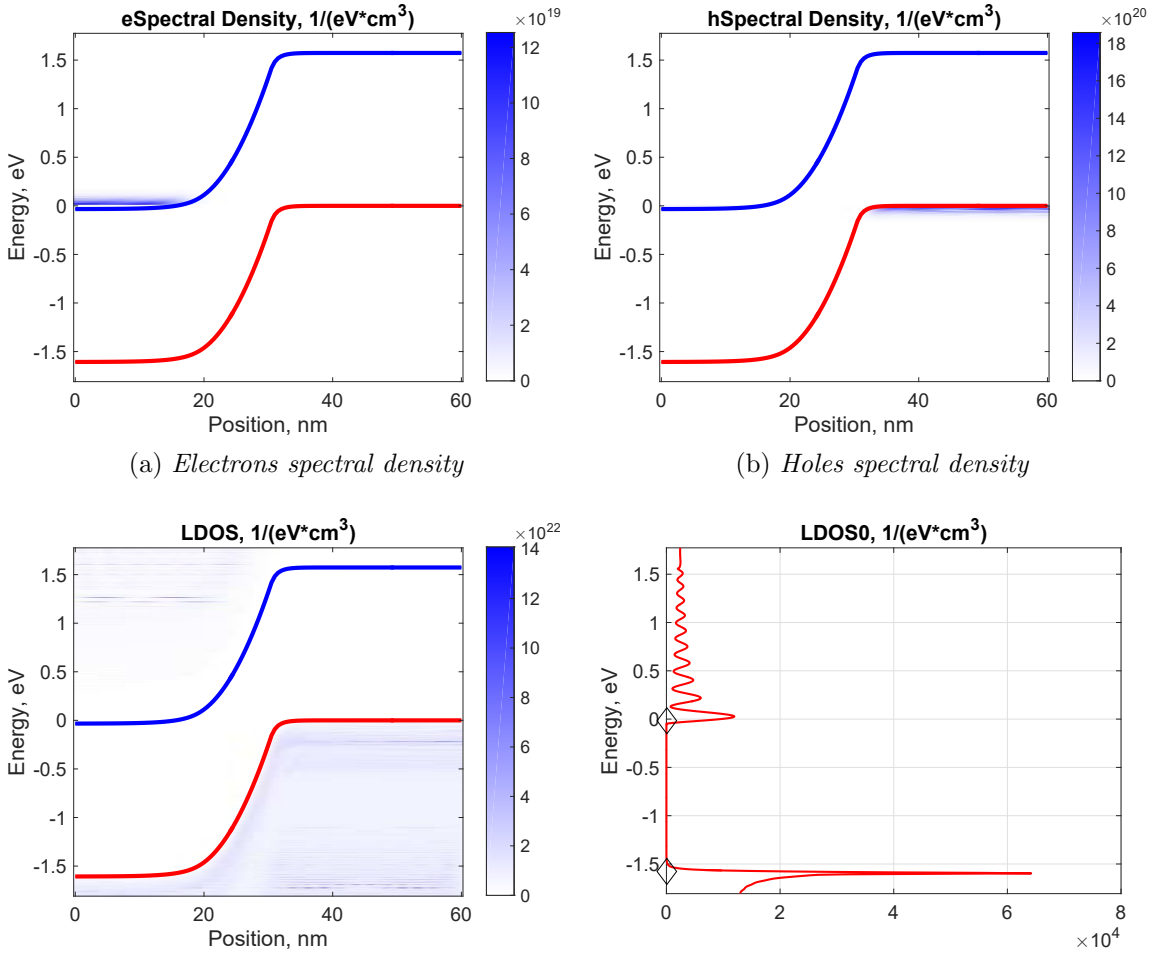
The energy integration spans an energy domain going from below 200 meV from the minimum of the valence band to 200 meV above the maximum of the conduction band.

Then, (4.19) is converted in a GR rate by simply dividing the current value by the elementary charge constant. Such a rate is assigned to those mesh points in the BTJ region where bands are overlooking, while it is equal to zero elsewhere. The drawback of this method is that the retrieved rate does not vary across the points where BTBT can occur. Therefore, even though a *non-local* model is exploited to obtain the current density, tunneling probability does not depends on tunnel path in each point, but comes from an overall current from which a constant rate is computed and plugged in the DD solver, which now reproduces the same value of current coming from NEGF.

Once that BTBT GR rate for each bias point of the voltage sweep is computed from NEGF formalism, a slightly modified version of D1ANA code is launched. In the latter, after the usual thermodynamic equilibrium simulation needed as a starting point, a loop on the bias points begins. For each step, BTBT is initially turned off; then a piece of code recognizes the mesh points where conduction and valence bands are overlooking: these are the ones where the *non-local* GR rate will be plugged in. However, the rate is not immediately set to the value coming from the procedure described before, since this would lead to a convergence failure. Therefore, BTBT is activated and a vector of rates is generated, starting from the value of $10^{11}\ cm^{-3} \cdot s^{-1}$: for each bias point, another loop over the BTBT rate begins until the final value is reached. Even though this results in a slow code, it is the best way to avoid issues related to convergence of the code.

Before moving to the analysis of the electrical quantities, it is interesting to comment on the results coming from NEGF in different bias conditions, in order to grasp some details about BTBT mechanism across a BTJ. As mentioned, NEGF produces spectral quantities as output. The spatial grid is obtained by using a spacing of 0.5 nm between each point. Let's start the analysis from the **equilibrium** condition, whose related quantities are reported in figures 4.30. Since no bias is applied, no BTBT can occur, therefore the electrons spectral density (fig. 4.30a) is different from zero only at the bottom of the CB of the n-doped side, whereas holes spectral density (fig. 4.30b) is non-zero only at the top of the VB of the p-doped side of the BTJ. Just in a very thin region near the band edges carrier densities take significant values, because of the position of the Fermi level. Local density of states (LDOS) is reported in fig. 4.30c, while in fig. 4.30d a cut of the local density of states is provided, from which it is clear that no states exist below 0 eV level. The ripples in the LDOS are due to the discretization used in the NEGF code.

Then, the system slowly enters the **reverse bias** condition. At V = 0.5 V bands level allows BTBT mechanism to occur. It is interesting to notice that in the cut of the LDOS provided in fig. 4.31a, some states are available below 0 eV, until -0.5 eV, thanks to the action of the applied bias: therefore, the main contribution to the tunneling



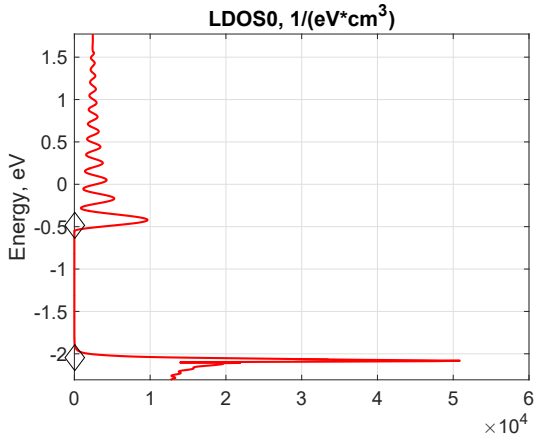
(a) *Electrons spectral density*

(b) *Holes spectral density*

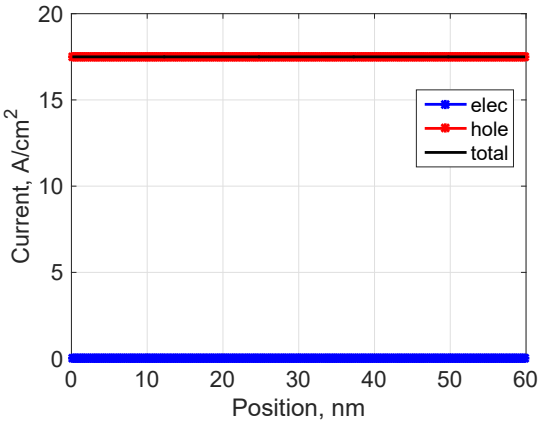
(c) *Local density of states*

(d) *Local density of states*

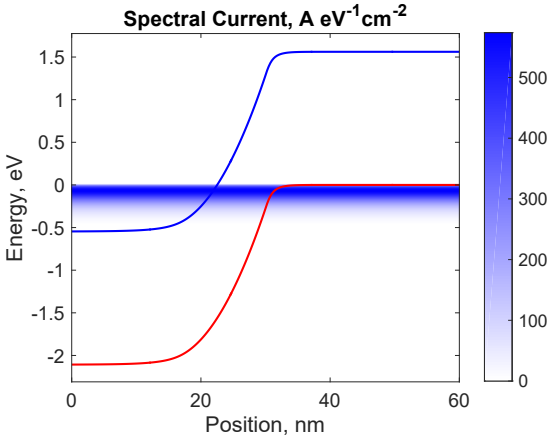
Figure 4.30: Spectral quantities in the AlGaAs BTJ simulated with NEGF at **equilibrium**.



(a) *Local density of states*



(b) *Spectral current density*

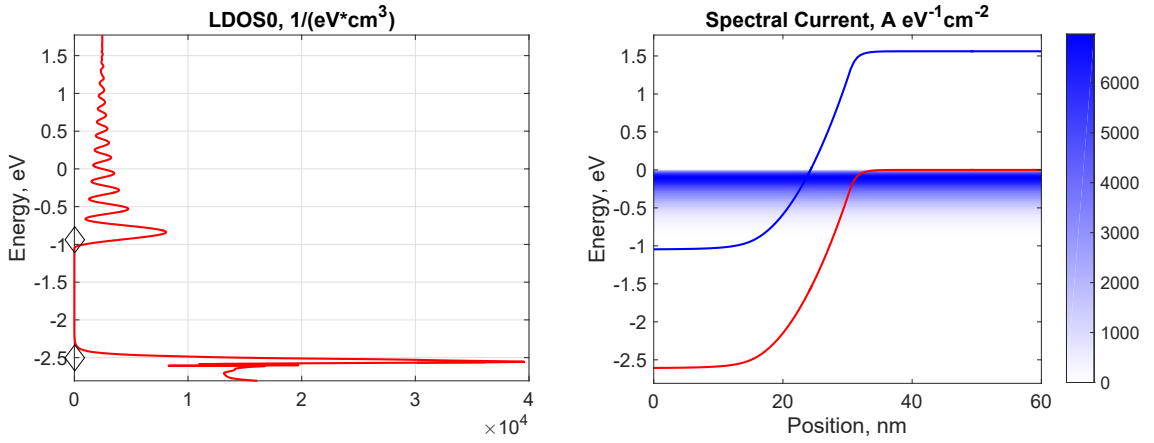


(c) *Spectral current density*

Figure 4.31: Spectral quantities in the AlGaAs BTJ simulated with NEGF at **reverse bias** (0.5 V).

current comes from the holes, as plotted in fig. 4.31b, whereas electron current stays at zero. The spectral current density profile is in reported in fig. 4.31c, from which it should be highlighted that the major contribution to the spectral current density comes from the 0 eV level, corresponding to the VB band edge at p-side, whereas carriers moving deeper in the band participate less to the tunneling process.

In figures 4.32, a bias of 1 V is applied to the junction. The cut of the LDOS is again reported in fig. 4.32a, where a considerably larger density of states is available for tunneling. This result already demonstrates that tunnel junctions can be exploited to effectively inject holes into a VCSEL active region. The value of spectral current, whose corresponding plot is displayed in fig. 4.32b, reaches values of thousands of $A/(eV \cdot cm^2)$, meaning that BTBT is actually capable of building of generating a large current across the junction. Again, (4.19) must be used to retrieve the corresponding current density value.



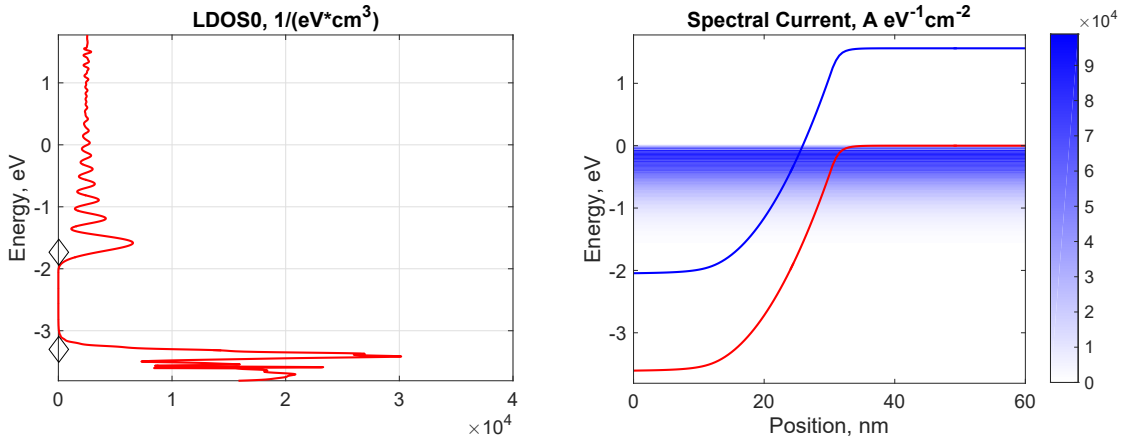
(a) *Local density of states* (b) *Spectral current density*

Figure 4.32: Spectral quantities in the AlGaAs BTJ simulated with NEGF at **reverse bias** (1 V).

Eventually, the results for the last bias point (corresponding to 2 V) of the voltage sweep are reported in fig. 4.33. It is interesting to notice (from fig. 4.33b) how the enlargement of the bias window (and, consequently, of the tunneling window), leads to a larger range of energies for which carriers are able to provide a significant contribution to the tunneling current, which in fact is one order of magnitude greater with respect to the previous case. The effect of the wider tunneling window is even more clear from the LDOS cut reported in fig. 4.33a.

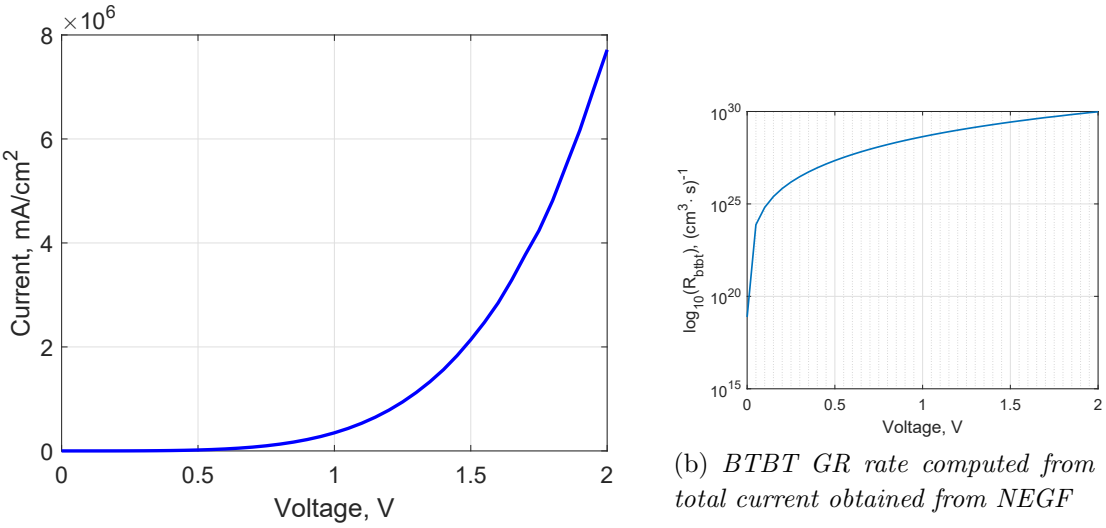
By applying (4.19) to all bias points investigated by means of NEGF formalism, the current density vs voltage curve plotted in fig. 4.34a is obtained. Some remarks are needed for it. First of all, the curve reaches values two times larger with respect to the InGaAsP-based BTJ ones, and several orders of magnitude larger than the current density values coming from Hurkx's model. This means that a short-wavelength emitting device implementing such a tunnel junction may actually work properly. However, one has to take into account that now the curve shows a different trend: while the small band gap system was characterized by an initial steep increase followed by a milder one, here current density stays low for bias smaller than 0.5 V, and then there is a parabolic increase of current. The curve reporting of the rates R_{btbt} (in *semilog* scale) computed from NEGF as a function of the applied bias is displayed in fig. 4.34b. Notice that the maximum value, of course reached at V = 2 V, is four orders of magnitude larger than the peak values coming from Hurkx's model applied to InGaAsP system.

In order to give a satisfactory explanation to the result just presented, it is worth taking a look to energy band diagram and to the charge densities computed in DD when it exploits the *non-local* BTBT GR rate from NEGF formalism, when the applied tension is equal to 2 V. The former is reported in fig. 4.35a, which is very similar to



(a) *Local density of states* (b) *Spectral current density*

Figure 4.33: Spectral quantities in the AlGaAs BTJ simulated with NEGF at **reverse bias** (2 V).

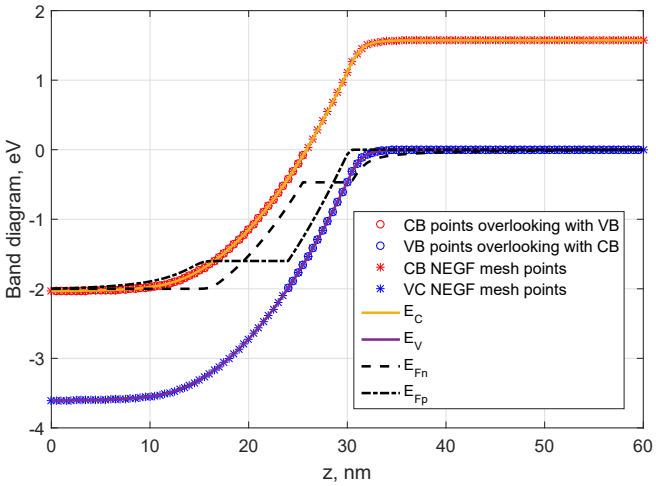


(a) *IV curve*

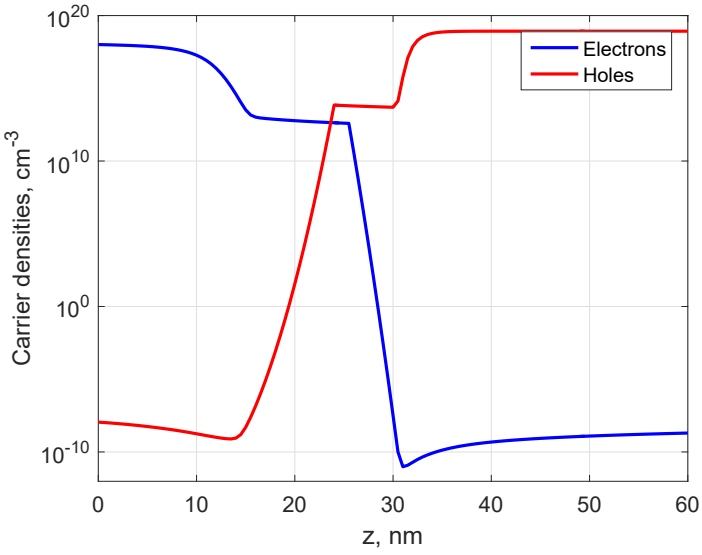
(b) *BTBT GR rate computed from the total current obtained from NEGF*

Figure 4.34: Current density vs applied voltage curve of $Al_{0.12}Ga_{0.88}As$ BTJ, obtained using the *non-local* BTBT GR rate coming from NEGF formalism (the computed rates are displayed in the figure at the right hand side).

the one coming from Hurkx's model. The greatest difference can be appreciated in the behaviour of both the quasi-Fermi levels inside the depletion region. Let's recall that in the *local* BTBT model they were joined in the whole spatial domain (see fig. 4.24a. On the contrary, the use of a *non-local* rate heavily modifies their behaviour, making them splitting and crossing several times. Furthermore, the plot highlights with a circle the points where tunneling of carriers occurs (of course, at equilibrium, there will be no points tagged with a circle, and overall tunneling is equal to zero). The modifications in the energy band diagram obviously lead to huge differences in the carrier densities profile, as clear from fig. 4.35b. The steep changes are due to the fact that tunneling rate is abruptly plugged only at the already discussed tagged points, and it does not reduces mildly close to the neutral sides. However, the most worthwhile difference with respect to the Hurkx's model result (see fig. 4.24b) is that the charge levels in the depletion region are significantly larger, and very close to the level reached by the InGaAsP-based BTJ. This explains the high level of current reached across the junction in this new framework.



(a) *Energy band diagram. Mesh points denoted by a * identify the points where the quantities are interpolated by the code. Those points also marked with a circle are the ones where bands are overlooking, so where tunneling actually takes place.*



(b) *Charge densities*

Figure 4.35: Energy band diagram and carrier densities at reverse bias condition (2 V applied), obtained from D1ANA with *non-local* BTBT GR rate.

CHAPTER 5

Conclusions and outlook

The carrier injection analysis performed about buried tunnel junctions VCSEL throughout the whole thesis work underlines that the goal of reaching a solid modelling of interband band-to-band tunneling is far from being said complete, at least in the drift-diffusion model framework. Nevertheless, it is demonstrated that BTJs efficiently inject carriers in the active region of a VCSEL, thus effectively substituting p-doped contacts. VCSEL performances are consequently improved when a properly designed tunnel junction is included in its layer (as discussed in the first chapter), with the additional advantage of a simplified fabrication process, when compared to the standard structure growth techniques (e.g. oxide aperture realization). However, whereas the simplest models that treat BTBT as a local generation-recombination rate, e.g. Hurkx and Kane's models (which are diffusely described in the third chapter), are suited only for small band gap semiconductors and thus for long-wavelength devices application, as described in section 4.1, on the other hand they completely fail when they are exploited to deal with wide gap materials (see section 4.2) BTJs.

The results coming from Hurkx's model implementation in a drift-diffusion solver suggest the need of more refined and rigorous methods to explore BTBT. The idea of using NEGF formalism comes from the fact that it has already been demonstrated to be a great solution to compute tunneling currents across nanoscaled structure over the years, especially in ballistic approximation and in steady-state conditions. The high degree of accuracy of such an approach, compared to the strongly approximated results coming from drift-diffusion models, is balanced by its huge computational burden, which makes it unsuited to work with structures longer than few hundreds of nanometers. Since it cannot be used to simulate a complete VCSEL, here it is employed for a novel scope. As a matter of fact, in order to go beyond *local* models, it is actually exploited to compute a *non-local* GR rate starting from a genuine quantum mechanical description of the tunneling mechanism. The thesis work only investigates the application of NEGF-based rate in wide-band gap AlGaAs-based BTJ. However, the final goal must involve the simulation of a complete structure, but several convergence

issues must be addressed before being able to obtain results for a BTJ-VCSEL.

A further work should also deal with a more sophisticated method of tunneling points recognition, which should be able to give a larger weight to the tunneling occurring in the points closer in space with respect to the ones more far away. This would lead to a more realistic and accurate simulation of the carrier tunneling phenomenon across a tunnel junction, and it could be also one of the possible way to solve the arising convergence issues. Furthermore, an optoelectronic device needs a multiphysics simulation (such as the academic simulator VENUS) able to solve in a self-consistent fashion optical, thermal and electric problems. Therefore, having demonstrated the improvements provided by the inclusion of a BTJ in VCSELs structures from the electrical standpoint, a deep analysis of the optical and thermal ameliorations is needed to reach a comprehensive understanding of the advantages of buried tunnel junctions.

Bibliography

[1] A. Tibaldi et al. "VENUS: a Vertical-cavity surface-emitting laser Electro-opto-thermal NUmerical Simulator". In: *IEEE Journal of Selected Topics in Quantum Electronics* 25.6 (2019), pp. 1–12.

[2] G. Ghione. *Semiconductor Devices for High-Speed Optoelectronics*. Cambridge University Press, 2009. DOI: 10.1017/CBO9780511635595.

[3] L. Mutter et al. "1.3 μm-wavelength phase-locked VCSEL arrays incorporating patterned tunnel junction". In: *Optics Express* 17.10 (2009). Ed. by Optical Soc. Amer., pp. 8558–8566. ISSN: 1094-4087.

[4] A. Karim et al. "Long-wavelength vertical-cavity lasers and amplifiers". In: *IEEE Journal of Selected Topics in Quantum Electronics* 6.6 (2000). Ed. by IEEE, pp. 1244–1253. ISSN: 1077-260X.

[5] F. Koyama. "Recent Advances of VCSEL Photonics". In: *Journal of Lightwave Technology* 24.12 (2006). Ed. by IEEE, pp. 4502–4513. ISSN: 0733-8724.

[6] R. Michalzik and K. J. Ebeling. *Vertical-Cavity Surface-Emitting Laser Devices*. Ed. by Herbert E. Li and Kenichi Iga. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 53–98. ISBN: 978-3-662-05263-1. DOI: 10.1007/978-3-662-05263-1_3.

[7] E. R. Hegblom et al. "Scattering losses from dielectric apertures in vertical-cavity lasers". In: *IEEE Journal of Selected Topics in Quantum Electronics* 3.2 (1997), pp. 379–389. ISSN: 1077-260X.

[8] D. L. Huffaker and D. G. Deppe. "Improved performance of oxide-confined vertical-cavity surface-emitting lasers using a tunnel injection active region". In: *Applied Physics Letters* 71.11 (1997), pp. 1449–1451. DOI: 10.1063/1.119933.

[9] M. C. Amann et al. "Low-resistivity p-side contacts for InP-based devices using buried InGaAs tunnel junction". In: *Electronics Letters* 36.1 (2000). Ed. by IEE-INST Elec. Eng., pp. 87–88. ISSN: 0013-5194.

[10] P. Legay et al. "Oxide confining layer on an InP substrate". In: *Journal of Applied Physics* 85.4 (1999), pp. 2428–2430. ISSN: 0021-8979.

[11] A. Mircea et al. "Cavity Mode-Gain Peak Trade-off for 1320-nm Wafer-Fused VCSELs With 3-mW Single-Mode Emission Power and 10-Gb/s Modulation Speed Up to 70° C". In: *IEEE Photonics Technology Letters* 19.2 (2007). Ed. by IEEE, pp. 121–123. ISSN: 1041-1135.

[12] M. X. Feng et al. "High efficient GaN-based laser diodes with tunnel junction". In: *Applied Physics Letters* 103.4 (2013), p. 043508.

[13] P. Y. Benjamin et al. "Demonstration of a III-nitride edge-emitting laser diode utilizing a GaN tunnel junction contact". In: *Opt. Express* 24.7 (Apr. 2016), pp. 7816–7822. DOI: 10.1364/OE.24.007816.

[14] V. Fan Arcara et al. "(Ga,In)N/GaN light emitting diodes with a tunnel junction and a rough n-contact layer grown by metalorganic chemical vapor deposition". In: *AIP Advances* 9.5 (2019), p. 055101. DOI: 10.1063/1.5092693.

[15] L. Esaki. "New phenomenon in narrow germanium p-n junctions [3]". In: *Physical Review* 109.2 (1958), pp. 603–604. ISSN: 0031899X.

[16] N. Nishiyama et al. "Long-wavelength vertical-cavity surface-emitting lasers on InP with lattice matched AlGaInAs-InP DBR grown by MOCVD". In: *Selected Topics in Quantum Electronics, IEEE Journal of* 11-5 (Oct. 2005), pp. 990–998. DOI: 10.1109/JSTQE.2005.853841.

[17] J. Piprek. *Handbook of Optoelectronic Device Modeling and Simulation Volume 1: Fundamentals, Materials, Nanostructures, LEDs, and Amplifiers*. CRC Press, 2017.

[18] R. Sacco et al. "Quantum-corrected drift-diffusion models for transport in semiconductor devices". In: *Journal of Computational Physics* 204.2 (2005), pp. 533–561.

[19] M. Goano. "Series expansion of the Fermi-Dirac integral $\mathcal{F}_j(x)$ over the entire domain of real j and x". In: *Solid State Electronics* 36.2 (1993), pp. 217–221. ISSN: 0038-1101.

[20] J.S. Blakemore. "Approximations for Fermi-Dirac integrals, especially the function $\mathcal{F}_{1/2}(\eta)$ used to describe electron density in a semiconductor". In: *Solid-State Electronics* 25.11 (1982), pp. 1067–1076. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(82)90143-5.

[21] A. Quarteroni and S. Quarteroni. *Numerical models for differential problems*. Vol. 2. Springer-Verlag Italia, 2009.

[22] W. Hänsch. *The Drift Diffusion Equation and Its Applications in MOSFET Modeling*. Computational Microelectronics. Springer Vienna, 2012. ISBN: 9783709190951.

[23] R. E. Bank, W.M. Coughran Jr., and L. C. Cowsar. "The Finite Volume Scharfetter-Gummel method for steady convection diffusion equations". In: *Computing and Visualization in Science* 1.3 (1998), pp. 123–136. ISSN: 1432-9360.

[24] *NSM Archive - Physical Properties of Semiconductors*. URL: http://www.ioffe.ru/SVA/NSM/Semicond/.

[25] X. Gao, B. Kerr, and A. Huang. "Analytic band-to-trap tunneling model including band offset for heterojunction devices". In: *Journal of Applied Physics* 125.5 (2019), p. 054503. DOI: 10.1063/1.5078685.

[26] S. M. Myers, W. R. Wampler, and N. A. Modine. "Recombination by band-to-defect tunneling near semiconductor heterojunctions: A theoretical model". In: *Journal of Applied Physics* 120.13 (2016), p. 134502. DOI: 10.1063/1.4963873.

[27] R. Adar. "Spatial integration of direct band-to-band tunneling currents in general device structures". In: *IEEE Transactions on Electron Devices* 39.4 (1992), pp. 976–981. ISSN: 0018-9383.

[28] R. H. Fowler and L. Nordheim. "Electron emission in intense electric fields". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 119.781 (1928), pp. 173–181.

[29] H. Ilatikhameneh et al. "From Fowler–Nordheim to Nonequilibrium Green's Function Modeling of Tunneling". In: *IEEE Transactions on Electron Devices* 63.7 (July 2016), pp. 2871–2878. ISSN: 1557-9646. DOI: 10.1109/ted.2016.2565582.

[30] E. O. Kane. "Theory of Tunneling". In: *Journal of Applied Physics* 32.1 (1961), pp. 83–91. DOI: 10.1063/1.1735965.

[31] E. O. Kane. "Zener tunneling in semiconductors". In: *Journal of Physics and Chemistry of Solids* 12.2 (1960), pp. 181–188. ISSN: 0022-3697.

[32] M. Mandurrino. "Advances in quantum tunneling models for semiconductor optoelectronic device simulation". Ph.D. Thesis. Politecnico di Torino, 2017.

[33] M. Hermle et al. "Numerical simulation of tunnel diodes for multi-junction solar cells". In: *Progress in Photovoltaics: Research and Applications* 16.5 (2008), pp. 409–418. ISSN: 1062-7995.

[34] G. A. M. Hurkx, F. G. O'Hara, and M. P. G. Knuvers. "Modelling Forward-Biased Tunneling". In: *ESSDERC '89: 19th European Solid State Device Research Conference* (1989). Ed. by IEEE, pp. 793–796.

[35] Synopsys, ed. *Sentaurus^{TM} Device User Guide*. Version K-2015.06. 2015.

[36] M. K. Ieong et al. "Comparison of raised and Schottky source/drain MOSFETs using a novel tunneling contact model". In: IEEE, 1998, pp. 733–736.

[37] A. Mereuta et al. "1.5 μm VCSEL structure optimization for high-power and high-temperature operation". In: *Journal of Crystal Growth* 272.1-4 (2004), pp. 520–525. ISSN: 0022-0248.

[38] S. Islam, A. Islam, and S. Islam. "Integrated Duo Wavelength VCSEL Using an Electrically Pumped GaInAs/AlGaAs 980 nm Cavity at the Bottom and an Optically Pumped GaInAs/AlGaInAs 1550 nm Cavity on the Top". In: *International Scholarly Research Notices* 2014 (Oct. 2014), pp. 1–10. DOI: 10.1155/2014/627165.

[39] A. Islam and S. Islam. "Designing a high speed 1310nm AlGaInAs/AlGaInAs VCSEL using MgO/Si top DBR and GaInAsP/InP bottom DBR". In: *American Journal of Optics and Photonics* 2 (Jan. 2014), pp. 37–44. DOI: 10.11648/j.ajop.20140203.14.

[40] J.-M. Lamy et al. "Design of an InGaAs/InP 1.55 μm electrically pumped VCSEL". In: *Optical and Quantum Electronics* 40.14 (Nov. 2008), pp. 1193–1198. ISSN: 1572-817X. DOI: 10.1007/s11082-009-9306-1.

[41] R. E. Nahory et al. "Band gap versus composition and demonstration of Vegard's law for $In_{1-x}Ga_xAs_yP_{1-y}$ lattice matched to InP". In: *Applied Physics Letters* 33.7 (1978), pp. 659–661. DOI: 10.1063/1.90455.

[42] R. L. Anderson. "Germanium-gallium arsenide heterojunctions (letter to the editor)". In: *IBM Journal of Research and Development* 4.3 (1960), pp. 283–287.

[43] R. Faez, A. Marjani, and S. Marjani. "Design and simulation of a high power single mode 1550nm InGaAsP VCSELs". In: *IEICE Electronics Express* 8.13 (2011), pp. 1096–1101.

[44] N. Moulin et al. "Tunnel junction I(V) characteristics: Review and a new model for p-n homojunctions". In: *Journal of Applied Physics* 126.3 (2019), p. 033105. DOI: 10.1063/1.5104314.

[45] S. Li et al. "Computing entries of the inverse of a sparse matrix using the FIND algorithm". In: *Journal of Computational Physics* 227.22 (2008), pp. 9408–9427.

[46] A. Afzalian et al. "An efficient tight-binding mode-space NEGF model enabling up to million atoms III-V nanowire MOSFETs and TFETs simulations". In: *arXiv preprint arXiv:1705.00909* (2017).

[47] J. Z. Huang et al. "Quantum Transport Simulation of III-V TFETs with Reduced-Order $k \cdot p$ Method". In: *Tunneling Field Effect Transistor Technology*. Ed. by Lining Zhang and Mansun Chan. Cham: Springer International Publishing, 2016, pp. 151–180. ISBN: 978-3-319-31653-6. DOI: 10.1007/978-3-319-31653-6_6.

[48] P. Vogl and T. Kubis. "The non-equilibrium Green's function method: an introduction". In: *Journal of Computational Electronics* 9.3 (Dec. 2010), pp. 237–242. ISSN: 1572-8137. DOI: 10.1007/s10825-010-0313-z.

[49] S. Datta. *Quantum Transport: Atom to Transistor*. Cambridge University Press, 2005. DOI: 10.1017/CBO9781139164313.

[50] U. Aeberhard. "A microscopic theory of quantum well photovoltaics". PhD thesis. ETH Zurich, 2008.

[51] S. Steiger. "Modelling nano-LEDs". PhD thesis. ETH Zurich, 2009. DOI: 10.3929/ethz-a-005834599.

[52] H. Haug and A.-P. Jauho. *Quantum kinetics in transport and optics of semiconductors*. Vol. 2. Springer, 2008.