POLITECNICO DI TORINO

Department of Electronics and Telecommunications

Master's Degree in *Nanotechnologies for ICTs*



Master's Degree Thesis

Accelerating Non-Equilibrium Green's Function Device Simulations with Newton-Raphson and Recursive Green's Function Methods

Supervisor:

Prof. Francesco Bertazzi

Co-supervisors:

Prof. Alberto Tibaldi
Prof. Michele Goano
Dr. Francesco Mercinelli
Dr. Lorenzo Miri

Candidate:

Riccardo Baralis

September 2026

Abstract

Modern nanoscale optoelectronic devices require quantum-transport models capable of describing the interplay of tunneling, confinement, open boundaries, and non-equilibrium carrier distributions. The non-equilibrium Green's function (NEGF) formalism provides such a framework. However, its substantial computational cost limits its application to realistic multilayer structures, as the Dyson and Keldysh equations, discretized over space and energy, must be repeatedly evaluated several times to achieve self-consistency between the carrier Green's functions and the self-energies describing the relevant interactions with photons, phonons, or other carriers.

The scope of this thesis is to develop and assess complementary numerical strategies aimed at reducing the computational cost while preserving consistency with conventional NEGF formulations. First, drift–diffusion (DD) and Schrödinger-corrected DD solutions are used to initialize Poisson–NEGF calculations and, within a frozen-field formulation, to provide a fixed electrostatic potential to the quantum-transport solver. In the latter approach, fixing the electrostatic potential reduces the size of the nonlinear system and enables efficient NEGF analysis of InGaN/GaN single-quantum-well LEDs, including polarization, incomplete dopant ionization, carrier confinement, recombination, and electron-blocking-layer effects. Second, a recursive Green's function (RGF) algorithm is implemented for the evaluation of retarded and lesser Green's functions and combined with Büttiker probes to account phenomenologically for dissipative scattering processes. By exploiting the block-tridiagonal structure of the Hamiltonian, the RGF formulation avoids inversion of the full device matrix, and enables to achieve self-consistency through a Newton-Raphson algorithm, reducing execution time and memory requirements.

The proposed methods are demonstrated on GaAs junctions, a realistic III-nitride visible LED, and a type-II superlattice for infrared photodetectors. For the type-II superlattice, a time saving exceeding 20% is obtained, whereas a substantially larger reduction in per-iteration execution time is achieved for the III-nitride LED. The results further show that the agreement obtained with the frozen-field approach depends on the electrostatic model used to generate the fixed potential.

Overall, the proposed combination of efficient semiclassical initialization, frozen-field transport, and recursive Green's function evaluation extends the practical range of one-dimensional NEGF simulations for nanostructured optoelectronic devices while retaining the quantum-mechanical transport description of the reference formulation.

Contents

1 Introduction . . . 4
2 Drift-Diffusion . . . 6
2.1 Device geometry . . . 6
2.2 Neutrality potential . . . 8
2.3 Drift-Diffusion model . . . 13
2.4 Out-of-equilibrium drift–diffusion . . . 14
2.5 Quantum corrections . . . 14
3 Non-Equilibrium Green's Functions . . . 21
3.1 The contour-ordered non-equilibrium Green's function . . . 21
3.2 Equations of motion . . . 23
3.3 Analytic continuation . . . 25
3.4 Steady state representation . . . 27
3.5 Spatial discretization . . . 28
3.6 Physical Observables . . . 30
3.7 Boundary self-energies . . . 35
3.8 Scattering self-energies . . . 36
3.9 Büttiker probe model . . . 39
4 Finite-difference implementation . . . 42
4.1 System partitioning . . . 43
4.2 Hamiltonian assembling . . . 45
4.3 Self-energy implementation . . . 46
4.4 Dyson equation and Keldysh equation . . . 50
4.5 Numerical Implementation of the Büttiker probe . . . 50
4.6 Physical observable evaluation . . . 52
4.7 Jacobian evaluation . . . 54
4.8 Poisson-NEGF . . . 62
4.9 Recombination currents . . . 64
5 Frozen Field approach . . . 67
5.1 Poisson–drift–diffusion and Poisson–Schrödinger–Drift–Diffusion . . . 67
5.2 Drift–Diffusion as an Initial Guess for Poisson–NEGF . . . 72
5.3 Frozen-Field Non-Equilibrium Green's Function . . . 73
5.4 Drift–Diffusion: Incomplete Ionization . . . 76
5.5 InGaN/GaN Single–Quantum–Well Light–Emitting Diode . . . 77
5.6 Recombination Current Analysis . . . 82
5.7 Internal Quantum Efficiency . . . 86
6 Recursive Green's function . . . 88

6.1 System partitioning . . . 89
6.2 Recursive retarded Green's function algorithm . . . 90
6.3 Recursive lesser Green's function algorithm . . . 94
6.4 Büttiker Probes in the Recursive Green's Function Algorithm . . . 98
6.5 RGF Validation on a GaAs p–n Junction . . . 101
6.6 RGF Validation on a Single–Quantum–Well InGaN LED . . . 104
6.7 RGF Validation on a Type–2 Superlattice . . . 106
7 Conclusion and Outlook . . . 113
A Newton-Raphson algorithm . . . 115

1 Introduction

The continuous evolution of semiconductor optoelectronics has been driven not only by advances in material growth techniques, but also by the ability to engineer materials at the nanometer scale. The introduction of low-dimensional structures has profoundly changed the design of optoelectronic devices, enabling functionalities that cannot be achieved in bulk materials. Nanostructures such as quantum wells, quantum wires, quantum dots, and superlattices allow the electronic band structure to be tailored through quantum confinement, leading to improved performance in conventional devices and to the development of entirely new device concepts.

Among the technologies that have benefited from nanoscale engineering are quantum-well laser diodes, which have become the standard architecture for many semiconductor lasers, quantum cascade lasers, whose operation relies on intersubband transitions in engineered heterostructures, and superlattice-based absorbers employed in infrared photodetectors. More recently, quantum transport phenomena have attracted considerable attention in III-nitride light-emitting diodes, where carrier confinement, tunneling processes, and nonequilibrium transport have been identified as possible contributors to efficiency degradation mechanisms such as efficiency droop. These examples highlight the wide range of applications in which quantum effects play a fundamental role and therefore require modeling approaches capable of describing them accurately.

The increasing importance of quantum confinement poses significant challenges to device simulation. For several decades, the drift-diffusion (DD) model has represented the workhorse of semiconductor device simulation, providing an excellent compromise between physical accuracy and computational efficiency for classical transport. However, its underlying assumptions become inadequate when the characteristic dimensions of the device approach the electron coherence length or when quantum-mechanical effects such as tunneling and confinement dominate carrier transport. Although several quantum corrections have been proposed to extend its applicability, these approaches remain approximate and are often unable to capture the full physics of strongly nanostructured devices.

More generally, the introduction of quantum-mechanical corrections within an inherently classical transport framework remains a challenging and, in some cases, debated issue. In the field of visible III-nitride light-emitting diodes, for example, the long-standing discussion on the microscopic origin of efficiency droop and other carrier loss mechanisms has highlighted the limitations of semiclassical descriptions. These challenges emphasize the need for a unified treatment of electronic transport and optical processes based on genuine quantum-kinetic approaches, such as the density-matrix formalism and the non-equilibrium Green's function (NEGF) theory.

In principle, the most complete description of quantum transport is provided by the many-body Schrödinger equation, which explicitly accounts for all interactions among particles in the system. However, solving the many-body problem for realistic semiconductor devices is computationally intractable due to the exponential growth of the Hilbert space with the number of particles. Consequently, practical quantum transport simulations rely on approximate formalisms that retain the essential quantum-mechanical features while remaining computationally feasible.

Among these approaches, the non-equilibrium Green's function (NEGF) formalism has emerged as one of the most powerful and widely adopted frameworks for nanoscale device simulation. By describing carrier transport through Green's functions and self-energies, NEGF naturally accounts for quantum coherence, tunneling, open boundary conditions, and scattering processes within a unified theoretical framework. As a result, it provides an excellent compromise between physical

accuracy and computational efficiency, making it particularly well suited for the simulation of modern nanostructured optoelectronic devices.

Despite its versatility and predictive capabilities, the NEGF formalism remains computationally demanding when applied to larger-scale devices. The solution of the coupled quantum transport equations requires repeated inversions of large sparse matrices over a dense energy grid, making simulations significantly more expensive than their drift-diffusion counterparts. This burden becomes particularly severe when considering realistic device geometries, scattering mechanisms, or fully self-consistent coupling with the Poisson equation. Consequently, computational cost often limits the range of devices and physical phenomena that can be investigated.

The primary objective of this thesis is to improve the computational efficiency of NEGF simulations while preserving the physical accuracy of the transport model. To this end, several complementary strategies are investigated, each targeting a different computational bottleneck of the NEGF framework.

To illustrate the relevance of these challenges, three representative classes of nanoscale optoelectronic devices are considered in this work.

A first relevant application is that of GaN-based light-emitting diodes (LEDs). These devices typically rely on a multi-quantum-well structure in which carrier injection, confinement, and recombination occur within nanostructured active regions. Despite their technological maturity, GaN LEDs are affected by efficiency degradation mechanisms, commonly referred to as efficiency droop. This effect arises from the interplay between carrier confinement, leakage, and scattering processes, which reduce radiative recombination efficiency at high injection levels. A microscopic understanding of these mechanisms therefore requires a transport description capable of treating quantum confinement and inelastic scattering on equal footing.

A second class of nanostructures is represented by semiconductor superlattices, and in particular type-II superlattices. These systems exhibit a rich interplay between different transport regimes, including miniband transport, Wannier–Stark hopping, and sequential tunneling, depending on structural parameters and applied bias. As a result, their accurate simulation using NEGF is computationally demanding, since multiple transport mechanisms must be resolved self-consistently over extended spatial and energy domains. For this reason, semiclassical approaches such as Poisson–Schrödinger drift-diffusion models are still widely employed due to their significantly lower computational cost, although they cannot fully capture quantum effects such as tunneling-induced interference and coherent transport.

The aim of this work is therefore to develop strategies that reduce the computational cost of NEGF simulations while preserving a quantum-consistent description of transport in superlattice-based devices. This enables a detailed investigation of the underlying carrier transport mechanisms, including recombination processes and their respective contributions, within a simulator that is both quantum-mechanically reliable and computationally efficient.

The methodologies developed in this work are validated on representative nanoscale optoelectronic devices, including III-nitride light-emitting diodes and semiconductor superlattices. While these applications demonstrate the versatility of the simulation framework, the main contribution of this thesis lies in the development and assessment of numerical strategies that significantly improve the efficiency of NEGF-based device simulation.

2 Drift-Diffusion

The drift–diffusion (DD) model is one of the most widely adopted approaches for the numerical simulation of semiconductor devices. Since its formulation by W. Van Roosbroeck in the seminal work of [1], it has become the standard framework for describing charge transport in a broad range of electronic and optoelectronic devices, owing to its favorable balance between physical accuracy and computational efficiency.

The DD formalism provides a semiclassical description of carrier transport under the action of electric fields and concentration gradients. It is based on the assumption that charge carriers undergo frequent scattering events with phonons, impurities, and other carriers, so that their momentum rapidly relaxes and local thermodynamic equilibrium is established. Under these conditions, the carrier dynamics can be described through two transport mechanisms: the *drift* component, driven by the electrostatic field, and the *diffusion* component, originating from spatial gradients of the carrier concentrations due to their thermal motion.

The model is formulated by coupling the Poisson equation, which determines the electrostatic potential inside the device, with the electron and hole continuity equations. Appropriate constitutive relations express the carrier current densities as the sum of drift and diffusion contributions, while generation and recombination mechanisms are incorporated through suitable physical models, such as Shockley–Read–Hall, radiative, and Auger recombination. The resulting nonlinear system is solved self-consistently to obtain the electrostatic potential together with the electron and hole distributions under equilibrium and bias conditions.

Despite its extensive use, the DD model remains a semiclassical approximation. Quantum-mechanical phenomena such as carrier confinement, interference, tunneling, and phase coherence are not explicitly described and therefore become increasingly important as the characteristic dimensions of the device approach the nanometer scale. To partially account for these effects while retaining the computational efficiency of the DD framework, quantum-corrected approaches can be employed. A common strategy consists of coupling the Poisson and continuity equations with the Schrödinger equation, allowing the quantization of the electronic states to modify the carrier densities and, consequently, the electrostatic potential in a self-consistent manner.

This chapter introduces the physical quantities and device parameters required for the DD simulation framework. The device geometry is first presented, followed by the definition of the neutrality potential and the derivation of the drift–diffusion equations under both equilibrium and out-of-equilibrium conditions. Finally, the quantum-correction approach adopted in this work is described as an extension of the semiclassical model.

2.1 Device geometry

To begin the simulations, the first step is to construct the structure of the device under investigation. This requires compiling a consistent set of material parameters for all the semiconductor compounds involved, as reported in the literature. The focus of this thesis is on GaAs- and GaN-based optoelectronics devices, including their associated ternary and quaternary alloys (such as AlGaAs, InGaAs, and AlInGaN). The material parameters depend on the molar fractions and are taken, where possible, from the parameterization provided by Vurgaftman [2].

The following quantities are required as input for the solution of Poisson's equation, which provides the electrostatic potential used to construct the band profile and to initialize subsequent transport simulations:

- molar fractions;
- temperature (simulations are mainly performed at 300 K);
- relative dielectric constant;
- dopant ionization energies;
- electron and hole effective masses;
- band offsets and bandgap energies.

The last group of parameters is relevant since the devices under investigation are heterostructures characterized by abrupt material interfaces, which lead to discontinuities in both the conduction and valence band edges. In order to properly describe these heterojunctions, the valence band edge energy E_V is defined as:

$$E_V = V_{BO} + E_o \tag{2.1}$$

where V_{BO} is the *valence band offset* and E_o is a generic electrostatic potential of reference. Now, the conduction band (E_C) is rewritten by means of the energy gap and, as done for the valence band, in terms of the conduction band offset:

$$E_C = E_V + E_g \;\Rightarrow\; C_{BO} + E_o = V_{BO} + E_o + E_g \;\Rightarrow\; C_{BO} = V_{BO} + E_g \tag{2.2}$$

In this way the *conduction band offset* is also defined.

The final step involves introducing the variations in the conduction and valence bands (ΔE_C and ΔE_V) across the heterojunctions, using the last layer of the structure as a reference. The difference between the i-th layer and the reference layer is expressed as:

$$\Delta E_C(z_i) = E_C(z_i) - E_C(z_{ref}) \tag{2.3}$$

This can also be written as:

$$\boxed{\Delta E_C(z_i) = C_{BO}(z_i) - C_{BO}(z_{ref}) + q\phi_r(z_{ref})} \tag{2.4}$$

where ϕ is the electrostatic potential obtained from the solution of Poisson's equation. The reference potential is fixed by imposing a suitable alignment condition, typically based on the intrinsic Fermi level of the system. This choice determines the absolute energy reference used throughout the simulation. To find the reference term we have that:

$$E_C(z_i) = -q(\phi(z_i) - \phi_r) \tag{2.5}$$

and defining the electrochemical reference energy as the conduction band can be rewritten as

$$E_{F,i}(z_i) = E_C(z_i) - C_{BO}(z_i) \tag{2.6}$$

At thermodynamic equilibrium, Boltzmann statistics gives the electron density as:

$$n_i(z_i) = N_C \, exp\left(\frac{E_{F,i}(z_i) - E_C(z_i)}{k_B T}\right) \tag{2.7}$$

The mass action law provides an alternative expression for the intrinsic carrier concentration:

$$n(z_i)p(z_i) = N_C \, N_V \, exp\left(\frac{-E_g(zi)}{k_B T}\right) \overset{\text{therm. equilib.}}{=} n_i^2(z_i) \tag{2.8}$$

By combining Eqs. (2.7) and (2.8), and assuming thermodynamic equilibrium ($n = p = n_i$), one can relate the intrinsic carrier concentration to the band edges. Assuming a homostructure region with constant bandgap, the reference potential can then be consistently defined after algebraic manipulation as follows:

$$\boxed{\phi_r = \frac{E_g}{2q} + \frac{k_B T}{2q} \ln\left(\frac{N_C}{N_V}\right)} \tag{2.9}$$

where N_C and N_V are the *effective density of states* of the conduction and valence bands of the bulk material[1].

$$N_C = \frac{2(2\pi k_B T m_0)^{3/2}}{h^3}(m_e^*)^{3/2} \tag{2.10a}$$

$$N_V = \frac{2(2\pi k_B T m_0)^{3/2}}{h^3}(m_h^*)^{3/2} \tag{2.10b}$$

The reference potential of the equation 2.9 is evaluated in the last layer of the heterostructures and substituted in 2.4 to quantify the discontinuities in the conduction and valence bands.
The last step is to evaluate the valence band discontinuity:

$$\Delta E_V(z_i) = \Delta E_C(z_i) - E_g(z_i) \tag{2.11}$$

And finally:

$$\boxed{\begin{aligned} E_C(z_i) &= -q\phi(z_i) + \Delta E_C(z_i) \\ E_V(z_i) &= -q\phi(z_i) + \Delta E_V(z_i) \end{aligned}} \tag{2.12}$$

The electrostatic potential ϕ is the last unknown of the band diagram and we will see how to find a suitable guess for the DD simulation in the next section.

Any device that is simulated requires a discretization in space since a real object would be made of infinite points and each physical quantity should be evaluated in any of them. Hence, the term z_i in equations 2.12 refers not only to a single layer of the device, but to a single point of the mesh that we have initially defined.

In this work, the following heterostructures will be of the order of tens of nanometers, and an acceptable value of mesh step is between 0.25 nm and 0.50 nm. A smaller step means adding computational operations and a much longer time of simulation; a longer step means less accuracy in the final results though faster calculations.

2.2 Neutrality potential

The final component needed to fully describe the band diagram of the device is the electrostatic potential in Eq. 2.12, which is calculated using Poisson's equation. Additionally, to characterize the behavior of the device near the metal contacts, it is essential to introduce boundary conditions that allow for a unique solution to Poisson's equation. This is why we impose the assumption of *neutrality* at the contacts, ensuring no charge accumulation at the metal-device interface and, consequently, no electric field distribution near the contacts. The neutrality condition involves

[1]The use of bulk effective density of states is an approximation that neglects the modified density of states in low-dimensional systems, such as quantum wells or superlattices, where confinement leads to discrete or quasi-two-dimensional energy distributions.

solving the neutrality equation for the density charge – a nonlinear equation – defined as well:

$$\rho = q(p - n + N_D^+ - N_A^-) = 0 \tag{2.13}$$

From now on, the doping terms will be compacted in a single term N_{DOP} in such a way to have: $N_{DOP} = N_D^+ - N_A^-$.

The electron and hole densities are calculated exploiting the bulk density of states and the Fermi-Dirac statistics, leading to the following expressions:

$$n = \int_{E_C}^{+\infty} N_C(E) \cdot f_{FD}(E) dE = \int_{E_C}^{+\infty} \left[\frac{\sqrt{2} m_n^{3/2}}{\pi^2 h^3} \sqrt{E - E_C} \right] \left[\frac{1}{1 + exp(\frac{E - E_F}{k_B T})} \right] dE \tag{2.14}$$

$$p = \int_{-\infty}^{E_V} N_V(E) \cdot (1 - f_{FD}(E)) dE = \int_{-\infty}^{E_V} \left[\frac{\sqrt{2} m_p^{3/2}}{\pi^2 h^3} \sqrt{E_V - E} \right] \left[\frac{1}{1 + exp(\frac{E_F - E}{k_B T})} \right] dE \tag{2.15}$$

Using a compact notation and Eq. 2.10a and Eq. 2.10b:

$$\begin{aligned} n &= N_C \, \mathcal{F}_{1/2}(\frac{E_F - E_C}{k_B T}) \\ p &= N_V \, \mathcal{F}_{1/2}(\frac{E_V - E_F}{k_B T}) \end{aligned} \tag{2.16}$$

The Fermi integrals of order 1/2 are numerically estimated using rational approximations, providing an accurate calculation of carrier densities and adjusting the electrostatic potential in the band diagram through a Newton-Raphson algorithm (see Appendix A). In fact, in Eq. 2.16, the conduction and valence bands must be determined through the electrostatic potential, which in turn influences the carrier concentrations, which in turn affect the electrostatic potential, etc. So, the neutrality equation 2.13 is solved self-consistently with the Fermi-Dirac evaluation of the electron and hole densities. The Fermi level is taken as the reference energy throughout the simulations, and is set to zero:

$$E_F = 0. \tag{2.17}$$

This choice fixes the energy reference and simplifies the expressions for the carrier densities and the electrostatic potential, without loss of generality.
The values of the electrostatic potential at the contacts (left and right contact) come from the iterations of the Newton method that, we learned, requires a guess for the simulations. In particular, we need an approximation for the carrier densities and the neutrality potential. Let's start from the intrinsic carrier concentration:

$$n_{int} = \sqrt{N_C N_V} \; exp(\frac{-E_g}{2 k_B T}) \quad \text{from equation 2.8} \tag{2.18}$$

It follows that the carrier densities can be evaluated from Eq. 2.13 exploiting the mass action law (Eq. 2.8) and remembering that the charge density ρ is zero under equilibrium conditions; in this way, a simple quadratic equation is involved, but we have to distinguish two cases based on doping concentration:

- n-type doping layer ($N_{DOP} > 0$), the hole concentration is given by $p = n_i^2/n$, so:

$$n_{neutr} = \frac{1}{2} \left(N_{DOP} + \sqrt{N_{DOP}^2 + 4 n_i^2} \right), \quad p_{neutr} = \frac{n_i^2}{n_{neutr}} \tag{2.19}$$

- p-type doping layer ($N_{DOP} < 0$), the electron concentration is given by $n = n_i^2/p$ so:

$$p_{neutr} = \frac{1}{2}\left(-N_{DOP} + \sqrt{N_{DOP}^2 + 4n_i^2}\right), \quad n_{neutr} = \frac{n_i^2}{p_{neutr}} \tag{2.20}$$

We are finally ready to solve the electrostatic potential. As said before, a Newton-Raphson algorithm has to be implemented, so let's see the application of such a self-consistent solver method. The electrostatic potential guess is evaluated from the Shockley equation of the electrons:

$$n = n_i \, exp\left(\frac{E_F - E_{F,i}}{k_B T}\right) \tag{2.21}$$

where E_F is set to zero and the intrinsic Fermi energy is defined as $E_{F,i} = -q\phi_{neutr}$ to find the following expression:

$$\phi_{neutr} = \frac{k_B T}{q} log\left(\frac{n_{neutr}}{n_i}\right) \tag{2.22}$$

with the previous approximation of the electrostatic potential, we can define the conduction band in Eq. 2.12 and solve the carrier concentration equations (Eq. 2.16) to be substituted in Eq. 2.13. Besides the charge density term, we need the derivative of such expression since the Jacobian matrix is involved in the Newton-Raphson cycles. The only needed terms are the derivatives of the carrier concentrations with respect to the electrostatic potential:

$$\frac{\partial n}{\partial \phi} = q\frac{N_C}{k_B T} \; \mathcal{F}_{-1/2}\left(\frac{-E_C}{k_B T}\right) \tag{2.23}$$

$$\frac{\partial p}{\partial \phi} = -q\frac{N_V}{k_B T} \; \mathcal{F}_{-1/2}\left(\frac{E_V}{k_B T}\right) \tag{2.24}$$

where we have used the standard derivative identity for the Fermi–Dirac integrals,

$$\frac{\mathrm{d}\mathcal{F}_j(\eta)}{\mathrm{d}\eta} = \mathcal{F}_{j-1}(\eta), \tag{2.25}$$

which, for the Fermi–Dirac integral of order $1/2$, gives

$$\frac{\mathrm{d}\mathcal{F}_{1/2}(\eta)}{\mathrm{d}\eta} = \mathcal{F}_{-1/2}(\eta). \tag{2.26}$$

Therefore, the derivative of the carrier concentrations with respect to the electrostatic potential involves the Fermi–Dirac integral of order $-1/2$. The derivative of the charge density is simply:

$$\frac{\partial \rho}{\partial \phi} = q\left(\frac{\partial p}{\partial \phi} - \frac{\partial n}{\partial \phi}\right) = \partial \underline{\rho} \tag{2.27}$$

The electrostatic potential is treated as a spatially discretized quantity, and therefore it is represented as a vector $\underline{\phi} = \phi(z_i)$, defined on the chosen one-dimensional mesh. All quantities depending on the electrostatic potential are accordingly evaluated pointwise on the grid. $\underline{\phi}$ is determined iteratively by implementing a Newton-Raphson algorithm, recalling Eq. A.4:

$$\begin{aligned} \underline{\Delta\phi}_k &= -\frac{\underline{\rho}(\phi_{k-1})}{\partial \underline{\rho}(\phi_{k-1})} \\ \underline{\phi}_k &= \underline{\phi}_{k-1} + \underline{\Delta\phi}_k \end{aligned} \tag{2.28}$$

where $\underline{\phi}_0 = \underline{\phi}_{neutr}$. Once the electrostatic potential has been obtained, the complete band diagram of the device can be reconstructed. This provides a consistent starting point for more advanced transport simulations and, in particular, serves as a physically motivated initial guess for subsequent DD calculations.

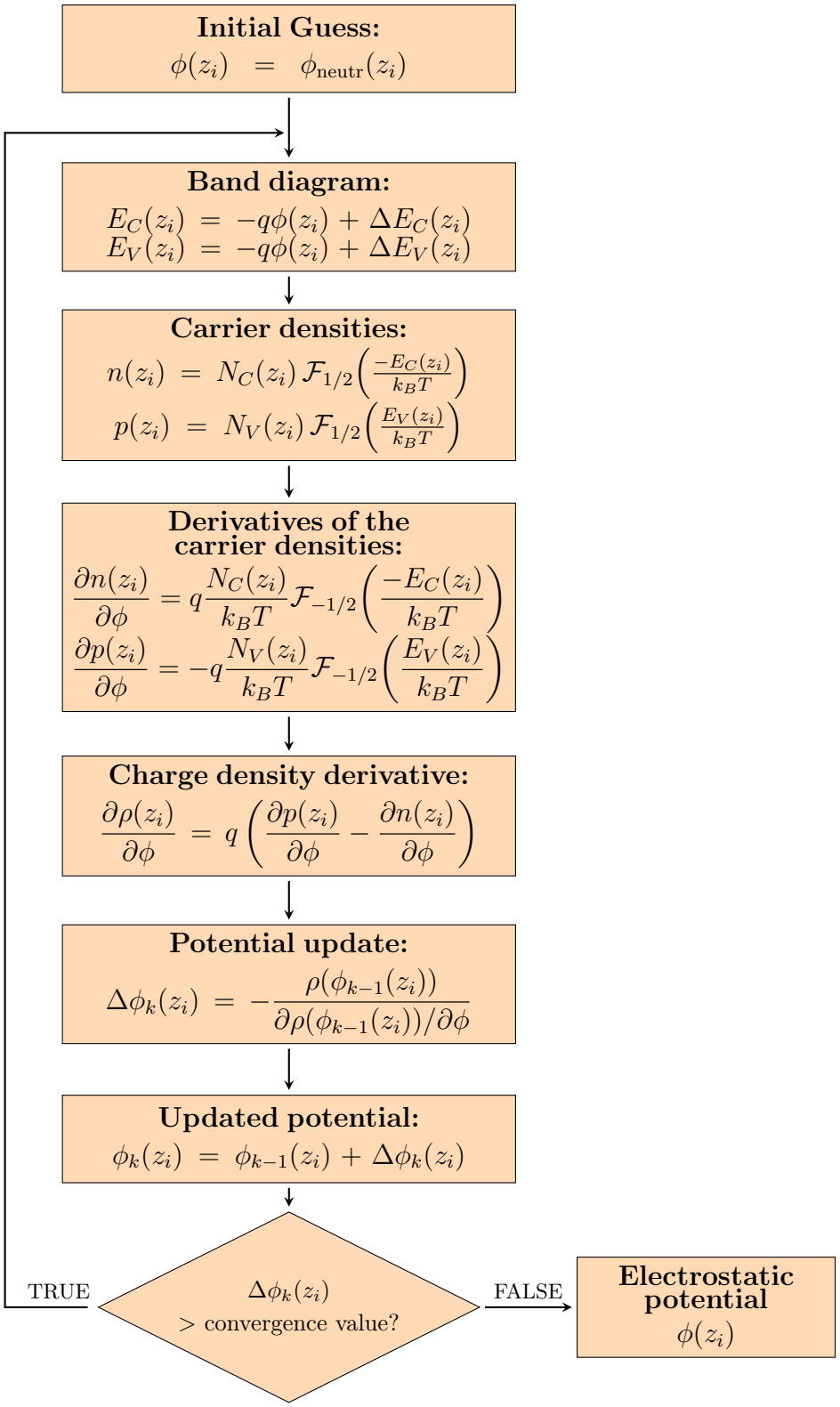


Figure 1: Flowchart of the Newton–Raphson algorithm used to compute the electrostatic potential under charge-neutrality conditions. Starting from an initial potential profile, the conduction and valence band edges, carrier densities, and their derivatives with respect to the electrostatic potential are evaluated. The potential is then iteratively updated until the prescribed convergence criterion is satisfied, providing the initial electrostatic potential used as the starting guess for the drift–diffusion solver.

2.3 Drift-Diffusion model

After constructing the initial geometry of the one-dimensional device of interest—defining doping concentration, dimensions, and materials—the non-equilibrium Green's function (NEGF) simulations require a physically consistent starting point for the evaluation of the key quantities describing the nanostructure. The drift-diffusion (DD) model provides such a starting point due to its straightforward implementation and relatively low computational cost. This model describes electronic transport in semiconductor devices at the macroscopic scale by solving the continuity equations for electrons and holes, coupled self-consistently with Poisson's equation:

- **Poisson's equation**

$$\nabla^2\phi = -\frac{q}{\epsilon}\left(p - n + N_D^+ - N_A^-\right) \tag{2.29}$$

- **Continuity equations**

$$\begin{cases} \frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot J_n - U_n \\ \frac{\partial p}{\partial t} = -\frac{1}{q}\nabla \cdot J_p - U_p \end{cases} \tag{2.30}$$

where n and p are the electron and hole densities, and U_n, U_p are the net recombination-generation rates. The current densities are defined as:

$$J_n = -q\mu_n n\nabla\phi + qD_n\nabla n \tag{2.31a}$$

$$J_p = -q\mu_p p\nabla\phi - qD_p\nabla p \tag{2.31b}$$

The diffusion coefficients D_n and D_p are given by the Einstein relations:

$$D_n = \mu_n\frac{k_BT}{q}, \qquad D_p = \mu_p\frac{k_BT}{q}. \tag{2.32}$$

The recombination-generation terms are modeled using the ABC model, accounting for radiative, Shockley–Read–Hall, and Auger processes. The formulation and implementation of the recombination-generation model will be discussed in detail in a subsequent section.

Thermodynamic equilibrium

The first operating condition considered within the drift-diffusion framework is thermodynamic equilibrium. In this regime, no external bias is applied and no net particle or energy flux exists throughout the device. Consequently, the electrochemical potential is spatially uniform, so that

$$E_{Fn}(x) = E_{Fp}(x) = E_F = 0 \tag{2.33}$$

where the electron and hole quasi-Fermi levels coincide with the equilibrium Fermi level. As a direct consequence, the electron and hole current densities vanish,

$$J_n = 0, \qquad J_p = 0, \tag{2.34}$$

meaning that the drift and diffusion contributions exactly compensate each other.

Under these conditions, the continuity equations are automatically satisfied, and the transport problem reduces to the solution of Poisson's equation alone. The carrier concentrations are evaluated from the Fermi–Dirac statistics introduced in Eq. 2.16, while the electrostatic potential is

obtained by solving the nonlinear Poisson–Fermi problem self-consistently through the Newton–Raphson procedure described in the previous section.

The resulting electrostatic potential defines the equilibrium band diagram of the entire heterostructure, including the built-in electric fields arising from doping and band discontinuities. This equilibrium solution provides the reference state for all subsequent biased drift-diffusion simulations and constitutes the initial condition adopted by the NEGF solver.

2.4 Out-of-equilibrium drift–diffusion

When an external bias is applied, the device is driven out of thermodynamic equilibrium and carrier transport must be explicitly considered. In this regime, electrons and holes are no longer described by a single electrochemical potential, but by two distinct quasi-Fermi levels, E_{Fn} and E_{Fp}, namely

$$E_{Fn}(x) \neq E_{Fp}(x). \tag{2.35}$$

Under steady-state conditions, the continuity equations take the form

$$\nabla \cdot J_n = qU_n, \qquad \nabla \cdot J_p = -qU_p, \tag{2.36}$$

where U_n and U_p represent the net recombination-generation rates. The electron and hole current densities are given by the drift–diffusion constitutive relations

$$J_n = -q\mu_n n\nabla\phi + qD_n\nabla n, \qquad J_p = -q\mu_p p\nabla\phi - qD_p\nabla p, \tag{2.37}$$

with the electric field defined as

$$\mathbf{E} = -\nabla\phi, \tag{2.38}$$

and the diffusion coefficients satisfying the Einstein relation:

$$D_n = \frac{k_B T}{q}\,\mu_n, \qquad D_p = \frac{k_B T}{q}\,\mu_p, \tag{2.39}$$

where μ represents the carrier mobility parameter. The electrostatic potential is obtained from Poisson's equation

$$\nabla^2\phi = -\frac{q}{\varepsilon}(p - n + N_D^+ - N_A^-), \tag{2.40}$$

which closes the electrostatic part of the system.

2.5 Quantum corrections

As discussed in the previous section, the drift-diffusion model relies on a semiclassical description of carrier transport and therefore neglects quantum confinement effects. When the characteristic dimensions of the device become comparable to the carrier de Broglie wavelength, this approximation is no longer adequate. A strategy to overcome this limitation while preserving the computational efficiency of the drift-diffusion framework consists in introducing quantum corrections obtained from the solution of the Schrödinger equation [3] [4]. Within this approach, the electrostatic potential is still obtained from Poisson's equation, while the conduction and valence band profiles are

written as

$$E_C = -q\phi + \Delta E_C, \tag{2.41a}$$
$$E_V = -q\phi + \Delta E_V, \tag{2.41b}$$

where ϕ is the electrostatic potential and ΔE_C and ΔE_V denote the conduction and valence band offsets. In the first iteration, these offsets are determined solely by the material composition of the heterostructure and are evaluated from experimentally determined band parameters [2, 5]. They therefore define the classical band diagram used to initialize the quantum-correction procedure. The resulting band profile is then employed as the potential energy in the effective-mass Schrödinger equation. Once the eigenstates are obtained, the electron density is evaluated according to

$$n(z) = \frac{m_e^*}{2\pi\hbar^2} \int \sum_i |\Psi^{(i)}(E,z)|^2 f_{FD}(E)\, dE, \tag{2.42a}$$
$$f_{FD}(E) = \frac{1}{1+\exp\left(\frac{E-E_F}{k_B T}\right)}.$$

Here, $\Psi(E,z)$ denotes the eigenfunction associated with the eigenenergy E, while f_{FD} is the Fermi–Dirac occupation function. An analogous expression is used to evaluate the hole density. To improve the estimation of the carrier density near the device boundaries, the Schrödinger equation is solved twice by imposing Neumann and Dirichlet boundary conditions, respectively. Since neither choice alone accurately reproduces the carrier distribution, the final carrier density is obtained as the average of the two solutions, following the approach proposed in [6]. The quantum carrier densities are subsequently converted into effective band edges by inverting the Fermi–Dirac statistics. For electrons, the effective conduction band is obtained from

$$E_{C,\text{eff}} = E_{Fn} - k_B T\, \mathcal{F}_{1/2}^{-1}\left(\frac{n}{N_C}\right) \tag{2.43}$$

while an analogous expression is employed for holes. Finally, the quantum-corrected conduction and valence band offsets are computed as

$$\Delta E_C = E_{C,\text{eff}} + q\phi, \tag{2.44}$$

and similarly for ΔE_V.

These corrected band offsets replace the purely material-dependent values in Eq. (2.41), allowing the drift–diffusion model to approximately account for quantum confinement effects. The resulting quantum-corrected electrostatic potential modifies the conduction and valence band profiles, leading to updated carrier densities and an enhanced carrier flux across potential barriers. In addition, the corrected band edges improve the description of carrier confinement within quantum wells through a more accurate redistribution of the carrier population.

Despite these improvements, the transport description remains fundamentally semiclassical. Quantum transport phenomena, such as coherent tunneling, wave interference, and the non-local nature of carrier propagation, are not described explicitly but only approximated through the corrected band profile. Moreover, the electronic states are still represented by bulk local densities of states rather than by the quantized states of an open quantum system. Consequently, while the

quantum-corrected DD approach captures the electrostatic effects of quantum confinement with good accuracy, it cannot fully reproduce intrinsically quantum transport phenomena, which require a wave-mechanical formulation such as the non-equilibrium Green's function (NEGF) formalism.

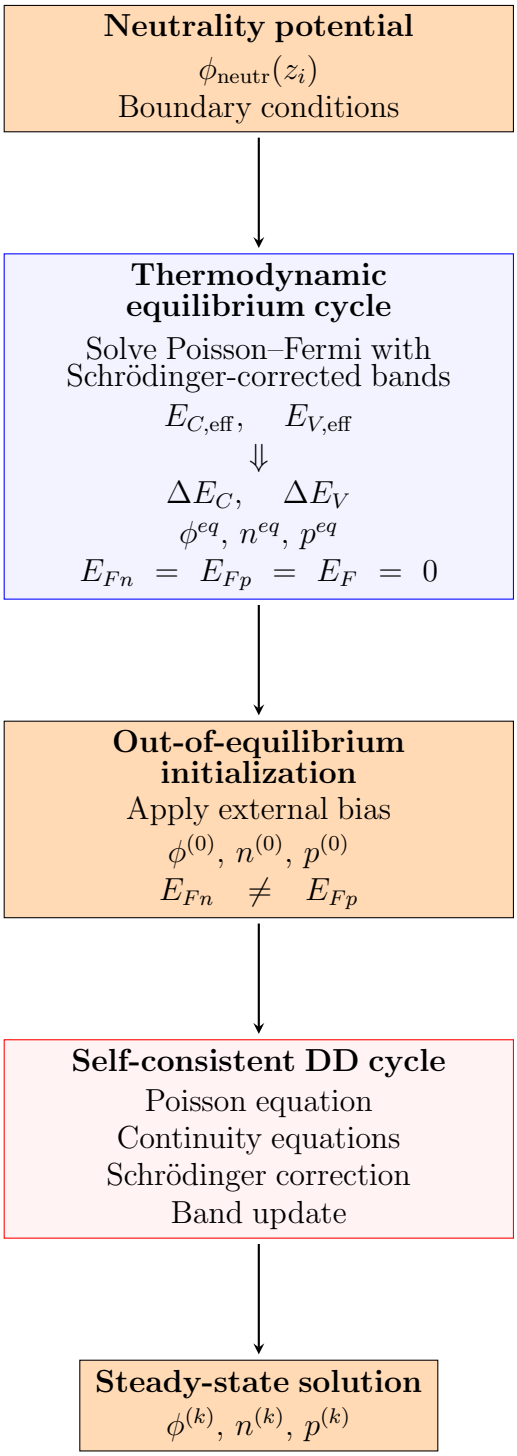


Figure 2: Workflow of the drift–diffusion (DD) simulation procedure. Starting from the neutrality potential and boundary conditions, the equilibrium solution is first obtained through a thermodynamic equilibrium cycle, represented by the blue block scheme below. The resulting electrostatic potential is then used to initialize the out-of-equilibrium problem under the applied bias. A self-consistent DD cycle subsequently solves the coupled Poisson and carrier continuity equations, including the Schrödinger-based band correction, until convergence is reached. The block highlighted in red corresponds to the main self-consistent iterative procedure of the out-of-equilibrium algorithm, which is detailed in the following flowcharts.

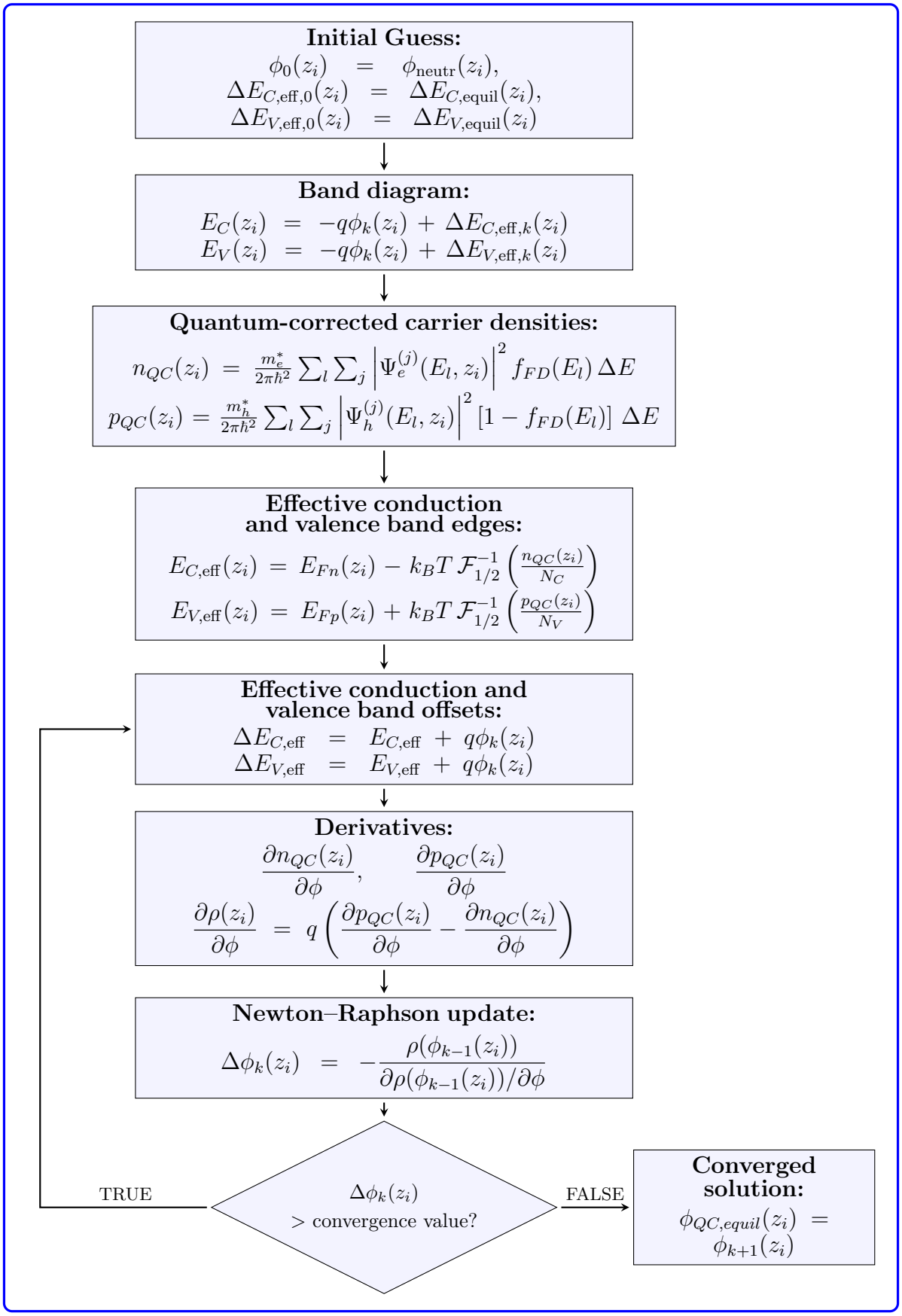


Figure 3: Flowchart of the Newton–Raphson algorithm used to compute the quantum-corrected electrostatic potential under thermodynamic equilibrium conditions. Starting from the neutrality potential, the band edges, quantum-corrected carrier densities, effective band offsets, and their derivatives are evaluated iteratively. The potential is updated until the prescribed convergence criterion is satisfied.

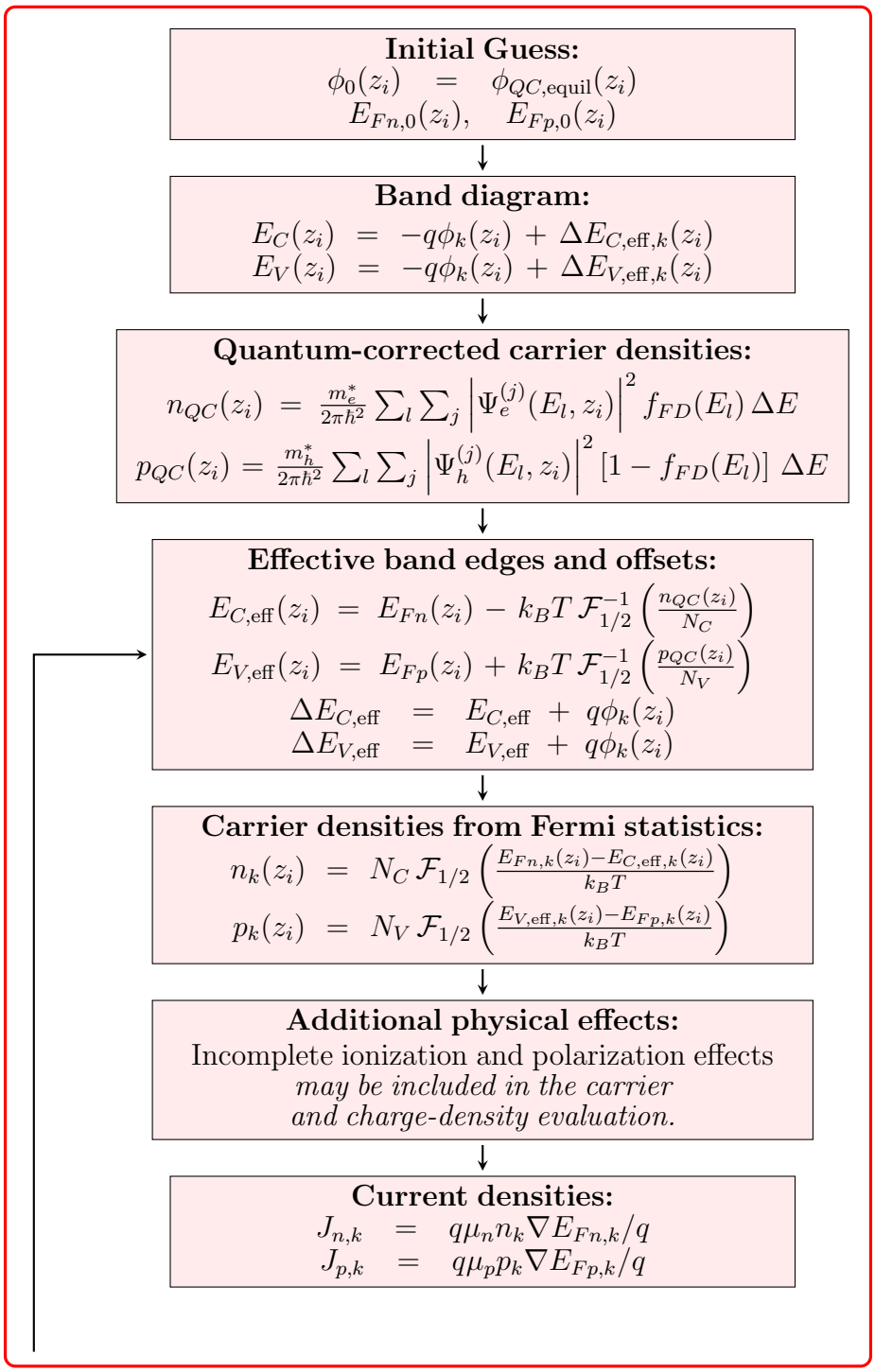


Figure 4: First stage of the self-consistent drift–diffusion algorithm under applied bias. The quantum-corrected band structure and carrier densities are evaluated, followed by the reconstruction of the effective band edges and carrier densities using Fermi statistics. Additional effects such as incomplete ionization and polarization can be included before evaluating the current densities.

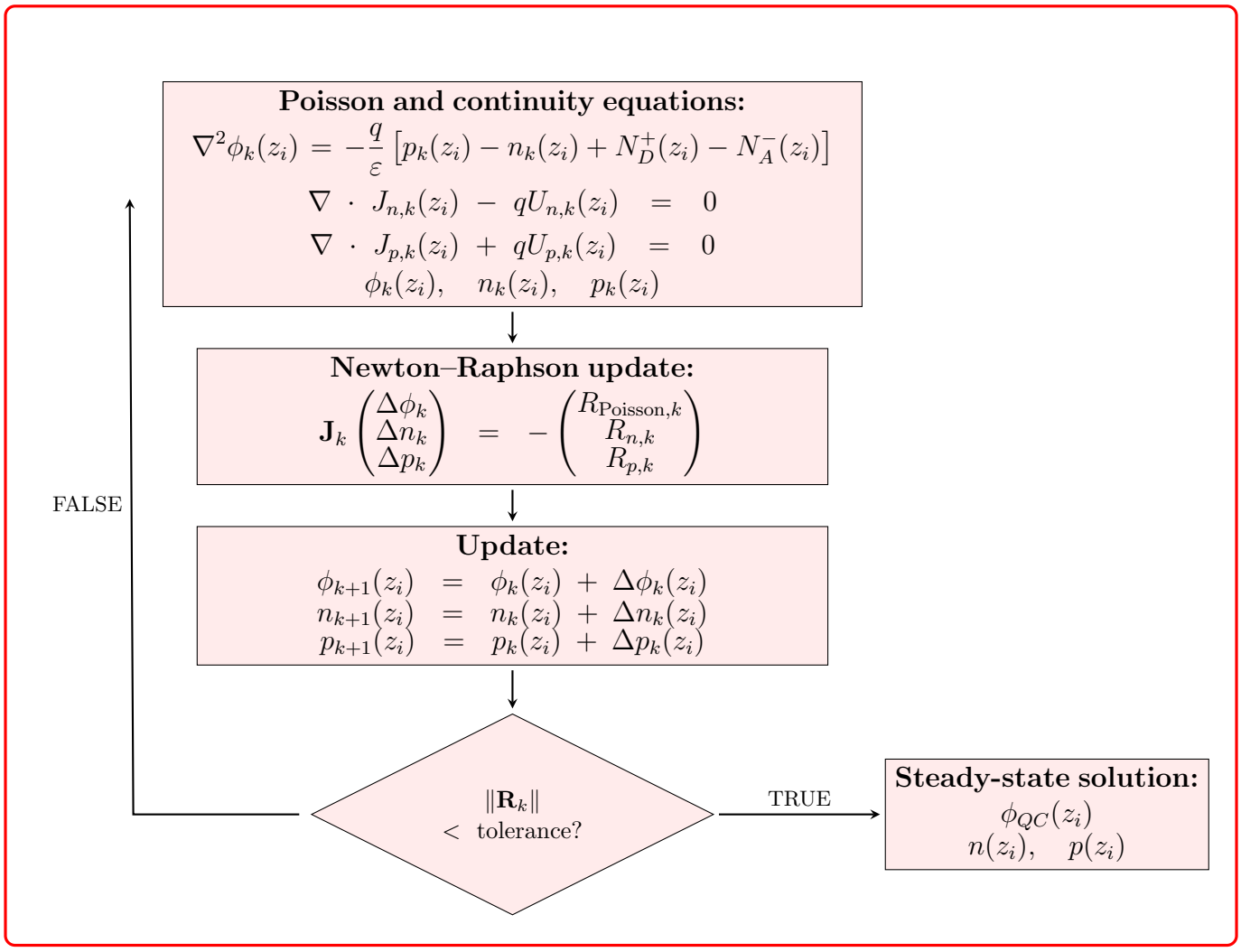


Figure 5: Second stage of the self-consistent drift–diffusion algorithm. The coupled Poisson and carrier continuity equations are solved using a Newton–Raphson procedure. The electrostatic potential and carrier densities are updated and the loop is repeated until the residuals satisfy the prescribed convergence criterion.

3 Non-Equilibrium Green's Functions

In the first part of this thesis, we discussed how the drift–diffusion (DD) model can be employed to obtain an approximate description of nanoscale devices, despite being originally developed for microscale semiconductor structures. We now introduce thenon-equilibrium Green's function (NEGF) formalism, which provides a rigorous quantum-mechanical framework capable of describing transport phenomena that become dominant at the nanoscale, such as quantum confinement and tunneling, and are therefore essential for the analysis of nanostructured devices.

The NEGF formalism describes quantum transport in terms of Green's functions, which are mathematical objects encoding the response of the system to external perturbations. Once the Green's functions have been determined, they provide access to a wide range of physical observables, including carrier densities, current densities, the local density of states, transmission probabilities, and radiative and non-radiative recombination rates.

The non-equilibrium Green's function (NEGF) formalism originates from the nonequilibrium quantum statistical mechanics developed by Schwinger, Kadanoff and Baym, and Keldysh during the 1960s. While Schwinger [7] introduced the closed-time contour formulation, Kadanoff and Baym established the theoretical framework for nonequilibrium Green's functions [8], and Keldysh developed the contour-ordered formalism together with the lesser and greater Green's functions that constitute the foundation of modern NEGF theory [9]. During the following decades, the formalism was adapted to quantum transport in mesoscopic and semiconductor devices, particularly through the work of Datta and co-workers, becoming one of the standard approaches for the simulation of nanoscale electronic devices [10].

Several equivalent Green's function formulations exist, particularly under equilibrium conditions, including the Matsubara formalism based on imaginary-time Green's functions [11]. Since semiconductor devices generally operate under applied bias, this thesis focuses on the nonequilibrium regime, which is described using contour-ordered Green's functions within the Keldysh formalism. The primary objective of the NEGF theory is to determine the Green's functions of the system, from which all relevant physical observables required to characterize the operation of a nanoscale device can be consistently evaluated.

3.1 The contour-ordered non-equilibrium Green's function

The starting point of the NEGF formalism is the description of an interacting many-body quantum system through the Hamiltonian:

$$\mathcal{H}(t) = \hat{H}_0 + \hat{H}^{int} + \hat{H}^{ext}(t) \tag{3.1}$$

where $\hat{H}_0$ represents the single-particle contribution, including the kinetic energy and the external potentials acting on the particles, such as the electrostatic potential profile inside the device. The term $\hat{H}^{int}$ accounts for the interactions among particles, including electron–electron, electron–phonon, and electron–photon interactions. The first two terms describe the intrinsic properties of the isolated system, while $\hat{H}^{ext}(t)$ represents an external perturbation that drives the system away from equilibrium, such as an applied electric field, optical excitation, or the coupling with external reservoirs.

The nanoscale devices considered in this work are open quantum systems composed of interacting particles. Their macroscopic properties are obtained from the evaluation of quantum statistical

averages of suitable observable operators. In this framework, the single-particle dynamics is described through Green's functions, which are two-point correlation functions containing information about the propagation and occupation of particles between two space–time coordinates.

Before introducing the Green's function definition, it is useful to introduce the Keldysh contour shown in Fig. 6. The contour-ordered formalism extends the conventional time-ordering procedure to non-equilibrium systems by defining a closed path in the time domain,

$$C = C_1 \cup C_2,$$

where the forward branch C_1 evolves from the initial time t_0 to the final time, while the backward branch C_2 returns to t_0. The closed-time contour does not represent a physical evolution in which the system returns to equilibrium. Instead, it is a mathematical construction introduced to account for the forward and backward time evolution operators appearing in nonequilibrium quantum statistical averages. This approach allows the definition of contour-ordered Green's functions, providing a unified description of quantum correlations in systems driven away from equilibrium.

The single-particle contour-ordered Green's function is defined as the nonequilibrium ensemble average of contour-ordered field operators:

$$\begin{aligned} G(1,1') &\equiv -\frac{i}{\hbar}\left\langle \hat{T}_C\left\{\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right\}\right\rangle \\ &= -\frac{i}{\hbar}\left(\theta_C(1,1')\left\langle \hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right\rangle \mp \theta_C(1',1)\left\langle \hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\hat{\Psi}_{\mathcal{H}}(1)\right\rangle\right) \end{aligned} \tag{3.2}$$

where the upper sign corresponds to fermions and the lower sign to bosons. The coordinate $1 \equiv (r_1, t_1, \sigma_1)$ identifies the spatial position, time, and spin state of the particle. The operators $\hat{\Psi}$ and $\hat{\Psi}^{\dagger}$ are respectively the annihilation and creation quantum field operators.

The contour-ordering operator $\hat{T}_C$ generalizes the conventional chronological ordering operator by arranging the field operators according to their position along the Keldysh contour:

$$\hat{T}_C\{\hat{A}(t_1)\hat{B}(t_1')\} = \theta_C(t_1,t_1')\hat{A}(t_1)\hat{B}(t_1') \mp \theta_C(t_1',t_1)\hat{B}(t_1')\hat{A}(t_1) \tag{3.3}$$

where θ_C is the contour Heaviside function, defined as:

$$\theta_C(t_1,t_1') = \begin{cases} 1, & \text{if } t_1 \text{ is later than } t_1' \text{ along } C, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

Starting from the definition of the contour-ordered Green's function, the objective of the NEGF formalism is the evaluation of measurable physical quantities, such as carrier densities and current densities, through quantum statistical averages of correlation functions.

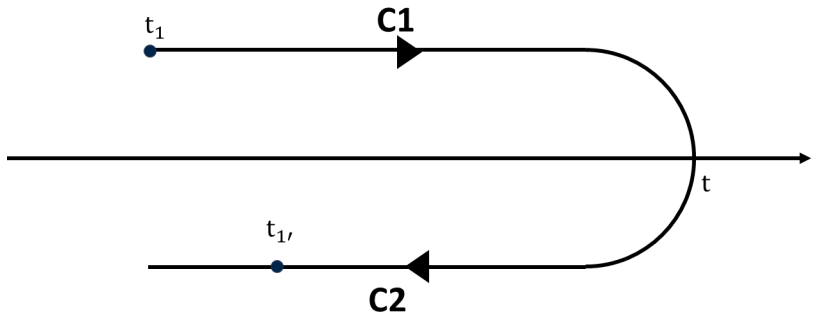


Figure 6: The Keldysh contour $C = C_1 \cup C_2$ used to evaluate nonequilibrium quantum statistical averages. The forward and backward branches of the closed-time contour allow the consistent treatment of time evolution operators in nonequilibrium systems and provide the basis for the definition of contour-ordered Green's functions.

3.2 Equations of motion

The following step is to determine the time evolution of the non-equilibrium Green's function by simply evaluating the time derivative of the Eq. 3.2:

$$\begin{aligned}
&\frac{\partial}{\partial t}G(1,1') = \frac{\partial}{\partial t}\left(\theta_C(1,1')\langle\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\rangle \mp \theta_C(1',1)\langle\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\hat{\Psi}_{\mathcal{H}}(1)\rangle\right) = \\
&= \delta_C(t,t')\,\langle\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\rangle \mp \delta_C(t',t)\,\langle\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\hat{\Psi}_{\mathcal{H}}(1)\rangle \\
&\quad + \theta_C(1,1')\langle\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(1)\right]\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\rangle \mp \theta_C(1',1)\langle\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(1)\right]\rangle \\
&= \left\langle\hat{T}_C\left\{\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(1)\right]\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right\}\right\rangle + \delta_C(t,t')\left\langle\left[\hat{\Psi}_{\mathcal{H}}(1),\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right]_{\pm}\right\rangle \\
&= \left\langle\hat{T}_C\left\{\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(1)\right]\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right\}\right\rangle + \delta_C(t,t')\,\delta(\mathbf{r}-\mathbf{r'}) \\
&\text{where } \frac{\partial}{\partial t}\theta_C(1',1) = \delta_C(t,t')
\end{aligned} \tag{3.5}$$

The compact notation $[\cdot,\cdot]_{\pm}$ stands for the (anti-)commutator for fermions and bosons. The anti-commutation property of the fermionic operator has been used:

$$\left[\hat{\Psi}_{\mathcal{H}}(1),\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\right] = \delta(\mathbf{r}-\mathbf{r'}) \tag{3.6}$$

Now we have to solve the time derivative of the field operator. Let's start with the Hamiltonian written as a function of three terms: the non-interacting Hamiltonian $\hat{H}_0$, the general particle-particle interaction $\hat{V}$, the single particle potential U, In second quantization, these terms are written in the *normal-ordered* notation:

$$\begin{aligned}
\hat{\mathcal{H}}(\mathbf{r},t) = &\int d\mathbf{r}\,\hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r},t)\left[\hat{H}_0(\mathbf{r}) + U(\mathbf{r},t)\right]\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \\
&+ \int d\mathbf{r}\int d\mathbf{r'}\,\hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r},t)\hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r'},t)\hat{V}(\mathbf{r}-\mathbf{r'})\hat{\Psi}_{\mathcal{H}}(\mathbf{r'},t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t)
\end{aligned} \tag{3.7}$$

The time-derivative of the field operators comes from the Heisenberg equation of motion:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{\Psi}_H(\mathbf{r},t) &= \left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t), \hat{\mathcal{H}}(t)\right] \\ &= \left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t), \hat{H}_0(\mathbf{r}) + U(\mathbf{r},t)\right] + \left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t), \hat{V}(\mathbf{r}-\mathbf{r'})\right] \\ &= \left[\hat{H}_0(\mathbf{r}) + U(\mathbf{r},t)\right] \hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \\ &\quad + \int d\mathbf{r'}\, V(\mathbf{r}-\mathbf{r'})\, \hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r'},t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r'},t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \end{aligned} \tag{3.8}$$

Substituting Eq. 3.8 in Eq. 3.5:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}G(11') &= \left\langle \left[\hat{H}_0(1) + U(1)\right] \hat{T}_C\{\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\}\right\rangle + \delta_C(1,1') \\ &\quad - i\hbar \int d\mathbf{r'}\, V(1-2)\hat{T_C}\{\, \hat{\Psi}^{\dagger}_{\mathcal{H}}(2)\hat{\Psi}_{\mathcal{H}}(2)\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\} \end{aligned} \tag{3.9}$$

where $1'$ and 1 are the endpoints of the Green's function propagator and 2 is a general intermediate interaction point. Recalling the definition of Green's function Eq. 3.2, we obtain:

$$\left(-i\hbar\frac{\partial}{\partial t_1} - H_0(\mathbf{r}_1) - U(1)\right) G(11') = \delta(11') - i\hbar\int_C d2\, V(1-2)\, G^{(2)}(121'2^+) \tag{3.10}$$

$$\left(-i\hbar\frac{\partial}{\partial t_{1'}} - H_0(\mathbf{r}_1) - U(1')\right) G(11') = \delta(11') - i\hbar\int_C d2\, V(1'-2)\, G^{(2)}(12^-1'2) \tag{3.11}$$

Two different expressions arise from the time-derivatives with respect to t and t'. We use a shorthand notation in the previous equations for the *two-particle Green's function*:

$$G^{(2)}(121'2') = (-\frac{i}{\hbar})^2\langle \hat{T_C}\{\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}_{\mathcal{H}}(2)\hat{\Psi}^{\dagger}_{\mathcal{H}}(2')\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\}\rangle \tag{3.12}$$

The notation 2^+ in Eq. 3.10 denotes a space–time point with time coordinate $t_2^+ = t_2 + 0^+$, i.e., infinitesimally later than t_2 along the contour. This prescription enforces the correct ordering of the field operators as required by Eq. 3.8, avoiding ambiguities and equal-time singularities in the time-ordered product.

The evaluation of the two-particle Green's function $G^{(2)}$ is complicated by the emergence of an infinite hierarchy of coupled equations: the equation of motion for $G^{(2)}$ involves the three-particle Green's function $G^{(3)}$, which in turn depends on $G^{(4)}$, and so forth. This generates a chain of interdependent correlations of increasing order, preventing a closed-form solution at the level of $G^{(2)}$.

A practical way to overcome this difficulty is to introduce the self-energy $\Sigma(11')$, which effectively incorporates the interaction effects of all particles into a single quantity. This allows the hierarchy to be recast into a closed equation for the one-particle Green's function through the Dyson equation:

$$\left(-i\hbar\frac{\partial}{\partial t_1} - H_0(\mathbf{r}_1) - U(1)\right) G(11') = \delta(11') - i\hbar\int_C d3\, \Sigma(13)G(31'), \tag{3.13}$$

$$\left(-i\hbar\frac{\partial}{\partial t_{1'}} - H_0(\mathbf{r}_{1'}) - U(1')\right) G(11') = \delta(11') - i\hbar\int_C d3\, G(13)\Sigma(31'), \tag{3.14}$$

A possible solution of the Green's function with respect to the previous equivalence is the following one:

$$\begin{aligned} G(11') &= G_0(11') + \int_C d2 \int_C d3 \;\; G_0(12)\Sigma(23)G(31') \\ G(11') &= G_0(11') + \int_C d2 \int_C d3 \;\; G(12)\Sigma(23)G_0(31') \end{aligned} \tag{3.15}$$

where G_0 is the non-interacting Green's function, solution of the previous equation without considering the two-particle term:

$$G_0(11') = \left[i\hbar \frac{\partial}{\partial t_1} - H_0(\mathbf{r}) - U(1) \right]^{-1} \delta(11') \tag{3.16}$$

To reach these results, it has been necessary to employ the Feynman diagrammatic description.

3.3 Analytic continuation

The question now is: how do we go from the imaginary *Keldysh contour* to the real-time Green's functions? Let us define the following real-time Green's functions based on the contour C (C_1 is the forward branch and C_2 is the backward branch) of Fig. 6:

$$G(11') \equiv \begin{cases} G^C(11') & t_1, t_{1'} \in C_1 \\ G^>(11') & t_1 \in C_2, \;\; t_{1'} \in C_1 \\ G^<(11') & t_1 \in C_1, \;\; t_{1'} \in C_2 \\ G^{\overline{C}}(11') & t_1, t_{1'} \in C_2 \end{cases} \tag{3.17}$$

which are named *chronological*, *greater*, *lesser*, and *anti-chronological* Green's functions, respectively. They are defined as well:

$$\begin{aligned} G^C(11') &= -\frac{i}{\hbar} \langle \hat{T_C} \{ \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^\dagger_{\mathcal{H}}(1') \} \rangle \\ G^{\overline{C}}(11') &= -\frac{i}{\hbar} \langle \hat{T_{\overline{C}}} \{ \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^\dagger_{\mathcal{H}}(1') \} \rangle \\ G^<(11') &= +\frac{i}{\hbar} \langle \hat{\Psi}^\dagger_{\mathcal{H}}(1') \hat{\Psi}_{\mathcal{H}}(1) \rangle \\ G^>(11') &= -\frac{i}{\hbar} \langle \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^\dagger_{\mathcal{H}}(1') \rangle \end{aligned} \tag{3.18}$$

Here, usually, the retarded and advanced Green's functions are introduced replacing the chronological and anti-chronological ones:

$$G^R(11') = -\frac{i}{\hbar} \theta(t_1 - t_{1'}) \langle [\hat{\Psi}_{\mathcal{H}}(1) \, \hat{\Psi}^\dagger_{\mathcal{H}}(1')]_+ \rangle = \theta(t_1 - t_{1'}) [G^>(11') - G^<(11')] \tag{3.19}$$

$$G^A(11') = \frac{i}{\hbar} \theta(t_{1'} - t_1) \langle [\hat{\Psi}_{\mathcal{H}}(1) \, \hat{\Psi}^\dagger_{\mathcal{H}}(1')]_+ \rangle = -\theta(t_{1'} - t_1) [G^>(11') - G^<(11')] \tag{3.20}$$

By subtracting Eq. 3.19 and Eq. 3.20 we obtain:

$$G^R(11') - G^A(11') = G^>(11') - G^<(11') \tag{3.21}$$

The following properties are satisfied on the chosen contour:

$$[G^{\lessgtr}(11')]^{\dagger} = -G^{\lessgtr}(11') \tag{3.22}$$

$$[G^{R}(11')]^{\dagger} = G^{A}(11') \tag{3.23}$$

Now, the contour integrals of Eq. 3.15 can be evaluated by introducing the rules due to Langreth rules [12]. A product of the form

$$D(t,t') = \int_C dt_1 \int_C dt_2 A(t,t_1)B(t_1,t_2)C(t_2,t') \tag{3.24}$$

can be decomposed using the Langreth demonstrations into real-time integrals:

$$\begin{aligned}
D^{R}(t,t') &= \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{+\infty} dt_2\, A^{R}(t,t_1)B^{R}(t_1,t_2)C^{R}(t_2,t') \\
D^{A}(t,t') &= \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{+\infty} dt_2\, A^{A}(t,t_1)B^{A}(t_1,t_2)C^{A}(t_2,t') \\
D^{\lessgtr}(t,t') &= \int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{+\infty} dt_2 \Big[A^{R}(t,t_1)B^{R}(t_1,t_2)C^{\lessgtr}(t_2,t') \\
&\quad + A^{R}(t,t_1)B^{\lessgtr}(t_1,t_2)C^{A}(t_2,t') + A^{\lessgtr}(t,t_1)B^{A}(t_1,t_2)C^{A}(t_2,t')\Big]
\end{aligned} \tag{3.25}$$

Using the results of Eq. 3.25 in Eq. 3.15 and introducing the compact notation without the integration over the states 2 and 3, the lesser and greater Green's functions are written as well:

$$G^{\lessgtr} = G_0^{\lessgtr} + G_0^{R}\Sigma^{R}G^{\lessgtr} + G_0^{R}\Sigma^{\lessgtr}G^{A} + G_0^{\lessgtr}\Sigma^{A}G^{A} \tag{3.26}$$

while G^R and G^A

$$G^{R} = G_0^{R} + G_0^{R}\Sigma^{R}G^{R}, \tag{3.27}$$

$$G^{A} = G_0^{A} + G_0^{A}\Sigma^{A}G^{A}. \tag{3.28}$$

From Eq. 3.26:

$$\begin{aligned}
(1 - G_0^{R}\Sigma^{R})G^{\lessgtr} &= G_0^{\lessgtr} + G_0^{R}\Sigma^{\lessgtr}G^{A} + G_0^{\lessgtr}\Sigma^{A}G^{A} \\
(1 - G_0^{R}\Sigma^{R})^{-1}(1 - G_0^{R}\Sigma^{R})G^{\lessgtr} &= (1 - G_0^{R}\Sigma^{R})^{-1}G_0^{\lessgtr}(1 + \Sigma^{A}G^{A}) + (1 - G_0^{R}\Sigma^{R})^{-1}G_0^{R}\Sigma^{\lessgtr}G^{A}
\end{aligned} \tag{3.29}$$

From Eq. 3.27:

$$(1 - G_0^{R}\Sigma^{R})^{-1}G_0^{R} = G^{R} \tag{3.30}$$

The $G_0^{\lessgtr}$ term of Eq. 3.29 is neglected since it represents the initial conditions of the system before a non-equilibrium state is reached, so that the initial effects are assumed to be damped. $G_0^{\lessgtr}$ is relevant only for transient behavior. The final result is the less/greater Green's function according to the Leonid Keldysh formalism:

$$\boxed{G^{\lessgtr} = G^R \Sigma^{\lessgtr} G^A} \tag{3.31}$$

Using the initial choice of the Keldysh contour and exploiting the integrations over the intermediate states, we finally obtain the *Keldysh equations*:

$$G^{\lessgtr}(11') = \int d2 \int d3 \; G^r(12)\Sigma^{\lessgtr}(23)G^A(31') \tag{3.32}$$

3.4 Steady state representation

In the steady-state regime, the Green's functions depend only on the time difference $\tau = t - t'$. Consequently, both Green's functions and self-energies can be expressed through Fourier transforms, allowing the formulation to be written in the energy domain:

$$\begin{aligned} O^\alpha(r_1, r_2; E) &\equiv \int_{-\infty}^{+\infty} d\tau \; e^{\frac{i}{\hbar}E\tau} O^\alpha(r_1, r_2; \tau) \\ O^\alpha(r_1, r_2; \tau) &\equiv \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dE \; e^{-\frac{i}{\hbar}E\tau} O^\alpha(r_1, r_2; E) \\ O &\in \{G, \Sigma\}, \alpha \in \{R, A, <, >\} \end{aligned} \tag{3.33}$$

Substituting into the Dyson 3.15 and Keldysh equations 3.32, the Green's functions are obtained as well:

$$G^R(r_1, r_{1'}; E) = G_0^R(r_1, r_{1'}; E) + \int dr_2 \int dr_3 \; G_0^R(r_1, r_2; E)\Sigma^R(r_2, r_3; E)G^R(r_3, r_{1'}; E) \tag{3.34}$$

same for G^A, and:

$$G^{\lessgtr}(r_1, r_{1'}; E) = \int dr_2 \int dr_3 \; G^R(r_1, r_2; E)\Sigma^{\lessgtr}(r_2, r_3; E)G^A(r_3, r_{1'}; E) \tag{3.35}$$

The non-interacting Green's function of Eq. 3.34 is given by:

$$G_0^R(r_1, r_{1'}; E) = \delta(r_1 - r_{1'}) \left[E + i\eta - \hat{H}_0(r_1) - u(r_1)\right]^{-1} \tag{3.36}$$

where $\eta \to 0^+$ is a parameter introduced to provide the correct analytical properties. The time reversal symmetry allows us to exploit the following properties

$$\begin{aligned} G^R(11') &= \left[G^A(1,1')\right]^* \to G^R(r_1, r_{1'}, E) = \left[G^A(r_1, r_{1'}, E)\right]^\dagger \\ G^{\lessgtr}(11') &= \left[G^{\lessgtr}(1,1')\right]^* \to G^{\lessgtr}(r_1, r_{1'}, E) = -\left[G^{\lessgtr}(r_1, r_{1'}, E)\right]^\dagger \end{aligned}$$

and reduces the number of independent Green's functions to two. In **equilibrium** conditions, the Green's function expressions can be further simplified through the *fluctuation-dissipation* theorem:

$$\boxed{\begin{aligned} G^<(E) &= if(E)A(E) \\ G^>(E) &= -i\left[1 - f(E)\right]A(E) \end{aligned}} \tag{3.37}$$

where $f(E)$ is the Fermi-Dirac function for fermions and $A(E)$ is defined as the *spectral function*:

$$A(E) \equiv i\left[G^R(E) - G^A(E)\right] = i\left[G^>(E) - G^<(E)\right] \tag{3.38}$$

Following the same road previously done for Green's functions, the self energy terms are written as well:

$$\Sigma^<(r_1, r_{1'}; E) = if(E - E_F)\Gamma(r_1, r_{1'}; E) \tag{3.39}$$

$$\Sigma^>(r_1, r_{1'}; E) = -i\left[1 - f(E - E_F)\right]\Gamma(r_1, r_{1'}; E) \tag{3.40}$$

where we introduced the broadening self-energy $\Gamma(E)$:

$$\boxed{\Gamma(E) = i\left[\Sigma^R(E) - \Sigma^A(E)\right]} \tag{3.41}$$

which is a quantity directly related to the spreading out of the energy levels of the device due to coupling with the environment. So, energy levels are no more described as long and sharp levels, but they acquire a width in energy.

3.5 Spatial discretization

The computational simulations are based on the discretization of the space in order to numerically evaluate the physical observables of a structure. The first step is to choose a finite basis $\phi_m(r, \sigma)$ through which we expand the quantum field operators:

$$\begin{aligned} \hat{\Psi}_{\mathcal{H}}(r, t, \sigma) &\equiv \sum_m \phi_m(r, \sigma)\hat{c}_m(t) \\ \hat{\Psi}^\dagger_{\mathcal{H}}(r, t, \sigma) &\equiv \sum_m \phi^*_m(r, \sigma)\hat{c}^\dagger_m(t) \end{aligned} \tag{3.42}$$

Using the *covariant* notation, and applying the steady state conditions, the Green's function is defined as:

$$G(r_1, r_{1'}; E) = \sum_{mn} \phi_m(r_1) G_{mn}(E) \phi^*_n(r_{1'}) \tag{3.43}$$

where $G_{mn}(E)$ is a matrix-valued function if the basis functions are discrete. Now, recalling Eq. 3.34 and applying the last expression:

$$\begin{aligned} &\sum_{mn} \phi_m(r_1) G^R_{mn}(E) \phi^*_n(r_{1'}) = \\ &\sum_{mn} \phi_m(r_1) \left[G^R_0\right]_{mn}(E) \phi^*_n(r_{1'}) + \int dr_2 \int dr_3 \sum_{m_1 n_1} \phi_{m_1}(r_1) \left[G^R_0\right]_{m_1 n_1}(E) \phi^*_{n_1}(r_2) \\ &\qquad \times \sum_{m_2 n_2} \phi_{m_2}(r_2) \left[\Sigma^R\right]_{m_2 n_2}(E) \phi^*_{n_2}(r_3) \\ &\qquad \times \sum_{m_3 n_3} \phi_{m_3}(r_3) \left[G^R\right]_{m_3 n_3}(E) \phi^*_{n_3}(r_{1'}) = \end{aligned}$$

$$= \sum_{mn} \phi_m(r_1) \left[G_0^R\right]_{mn}(E)\phi_n^*(r_{1'}) + \sum_{m_1n_1}\sum_{m_2n_2}\sum_{m_3n_3} \phi_{m_1}(r_1)\left[G_0^R\right]_{m_1n_1}(E)\Sigma^R_{m_2n_2}(E)G^R_{m_3n_3}(E)\phi^*_{n_3}(r_{1'}) \times \underbrace{\int dr_2 \;\; \phi^*_{n_1}(r_2)\phi_{m_2}(r_2)}_{M_{n_1m_2}} \underbrace{\int dr_3 \;\; \phi^*_{n_2}(r_3)\phi_{m_3}(r_3)}_{M_{n_2m_3}} =$$

$$= \sum_{mn} \phi_m(r_1) \left[G_0^R\right]_{mn}(E)\phi_n^*(r_{1'}) + \sum_{m_1n_1}\sum_{m_3n_3} \phi_{m_1}(r_1)\left[G_0^R\right]_{m_1n_1}(E) \underbrace{\sum_{m_2n_2} M_{n_1m_2}\Sigma^R_{m_2n_2}(E)M_{n_2m_3}}_{\Sigma_{n_1m_3}(E)} G^R_{m_3n_3}(E)\phi^*_{n_3}(r_{1'}) \tag{3.44}$$

Using the matrix notation, the previous equation becomes:

$$G^R(E) = G_0^R + G_0^R(E)\Sigma^R(E)G^R(E) \tag{3.45}$$

equivalently:

$$\boxed{\left[G_0^R(E)^{-1} - \Sigma^R(E)\right] G^R(E) = \mathbb{I}} \tag{3.46}$$

which is recognized as *Dyson's equation*, where the non-interacting term is given by:

$$\begin{aligned}
\left[G_0^R\right]^{-1}_{mn}(E) &= \int dr_1 \int dr_{1'} \;\; \phi^*_m(r_1)\left[G_0^R\right]^{-1}(r_1, r_{1'}; E)\phi_n(r_{1'}) \\
&= \int dr_1 \int dr_{1'} \;\; \phi^*_m(r_1)\left[E + i\eta - H_0(r_1) - U(r_1)\right]\delta(r_1 - r_{1'})(r_1, r_{1'}; E)\;\phi_n(r_{1'}) \\
&= \int dr_1 \;\; \phi^*_m(r_1)\left[E + i\eta - H_0(r_1) - U(r_1)\right]\;\phi_n(r_1) \\
&= (E + i\eta)M_{mn} - H_{mn}
\end{aligned} \tag{3.47}$$

where

$$\begin{aligned}
H_{mn} &= \int dr \;\; \phi *_m (r)\left[H_0(r) + U(r)\right]\phi_n(r) \\
M_{mn} &= \int dr \;\; \phi^*_m(r)\phi_n(r)
\end{aligned}$$

H_{mn} is the Hamiltonian written according to the chosen basis system, while M_{mn} is the overlapping matrix between the basis functions, which means if the basis is orthogonal, this term collapses into δ_{mn}.
In a similar way, the greater and lesser Green's functions can be discretized as well:

$$\sum_{mn} \phi_m(r_1)G^{\lessgtr}_{mn}\phi^*_n(r_{1'}) = \sum_{mn}\phi_m(r_1)\{\sum_{kl} G^R_{mk}(E)\Sigma^{\lessgtr}_{kl}(E)G^A_{ln}(E)\}\phi_n(r_{1'}) \tag{3.48}$$

And by using the linear matrix notation, we recognize the *Keldysh equation*:

$$\boxed{G^{\lessgtr} = G^R(E)\Sigma^{\lessgtr}(E)G^A(E)} \tag{3.49}$$

3.6 Physical Observables

This section introduces the main quantities that characterize an electronic device from a microscopic perspective, with particular focus on physical observables such as carrier density, local density of states, and current density.

The key ingredients are the retarded Green's function G^R (Eq. 3.92), which describes the causal propagation of particle-like excitations, and the advanced Green's function G^A (Eq. 3.93), which is the Hermitian conjugate of G^R and satisfies the boundary conditions associated with time-reversed evolution. The spectral function A encodes the available states of the system and is defined as

$$A(E) = i\left(G^R(E) - G^A(E)\right), \tag{3.50}$$

while the occupation of these states is described by the lesser and greater Green's functions.

The first step is to write a general operator in second quantization as:

$$\hat{O}(t) = \int dr \sum_{\sigma\sigma'} \hat{\Psi}^\dagger_{\mathcal{H}}(r,t,\sigma)\, O(r,t)\, \hat{\Psi}_{\mathcal{H}}(r,t,\sigma'). \tag{3.51}$$

From quantum statistical averaging [13], we obtain:

$$\begin{aligned}
\langle \hat{O}(r,t)\rangle &= \mathrm{Tr}[\rho \hat{O}(r,t)] \\
&= \lim_{r'\to r}\lim_{t'\to t^+} \sum_{\sigma\sigma'} O_{\sigma\sigma'} \mathrm{Tr}\left[\rho \hat{\Psi}^\dagger_{\mathcal{H}}(r',t',\sigma)\hat{\Psi}_{\mathcal{H}}(r,t,\sigma')\right] \\
&= \mp \lim_{r'\to r}\lim_{t'\to t^+} \mathrm{Tr}\left[O(r,t)\, G^<(r,t;r',t')\right].
\end{aligned} \tag{3.52}$$

This expression shows that the expectation value of an observable is directly related to the occupied states through the lesser Green's function $G^<$. The time argument t^+ ensures the correct ordering of the field operators. The upper (lower) sign applies to fermions (bosons). Here, σ and σ' denote the spin indices associated with the two field operators, while $O_{\sigma\sigma'}$ represents the corresponding matrix elements of the observable in spin space.

Carrier density

From the definition of the density operator, the expressions for particle and hole densities are:

a) **Particles:**

$$\begin{aligned}
n(r,t) &= \sum_\sigma \langle \hat{\Psi}^\dagger_{\mathcal{H}}(r,t,\sigma)\hat{\Psi}_{\mathcal{H}}(r,t,\sigma)\rangle \\
&= -i\hbar \lim_{r'\to r}\lim_{t'\to t^+} \sum_\sigma G^<_{\sigma\sigma}(r,r';t,t') \\
&= -i\hbar \sum_\sigma \int \frac{dE}{2\pi\hbar} G^<_{\sigma\sigma}(r,r;E) \\
&= -i \sum_\sigma \int \frac{dE}{2\pi} G^<_{\sigma\sigma}(r,r;E).
\end{aligned} \tag{3.53}$$

b) **Holes:**

$$\begin{aligned} p(r,t) &= \sum_{\sigma} \langle \hat{\Psi}_{\mathcal{H}}(r,t,\sigma) \hat{\Psi}^{\dagger}_{\mathcal{H}}(r,t,\sigma) \rangle \\ &= i\hbar \lim_{r' \to r} \lim_{t' \to t^{+}} \sum_{\sigma} G^{>}_{\sigma\sigma}(r,r';t,t') \\ &= i \sum_{\sigma} \int \frac{dE}{2\pi} G^{>}_{\sigma\sigma}(r,r;E). \end{aligned} \tag{3.54}$$

Local density of states

Recalling Eq. 3.37, the particle density can be written as:

$$n(r) = \sum_{\sigma} \int \frac{dE}{2\pi} f(E)\, A_{\sigma\sigma}(r,r;E). \tag{3.55}$$

The local density of states (LDOS) is defined as:

$$n(r) = \int dE\, f(E)\, \mathrm{LDOS}(r,E), \tag{3.56}$$

so that:

$$\mathrm{LDOS}(r,E) = \frac{1}{2\pi} \sum_{\sigma} A_{\sigma\sigma}(r,r;E). \tag{3.57}$$

where A is defined by Eq. 3.38.

Current

The electronic current is one of the central observables in quantum transport. Within the NEGF formalism, it can be derived from the quantum-mechanical current density operator and subsequently related to the equations of motion of the Green's functions. This connection provides a direct route to the continuity equation and, consequently, to the evaluation of carrier transport and recombination-generation processes.

In second quantization, the charge-current density operator for a parabolic-band system can be written as

$$\hat{\mathbf{J}}(r,t) = \frac{q\hbar}{2mi} \sum_{\sigma} \left[\hat{\Psi}^{\dagger}_{\mathcal{H}}(r,t,\sigma) \nabla_r \hat{\Psi}_{\mathcal{H}}(r,t,\sigma) - \left(\nabla_r \hat{\Psi}^{\dagger}_{\mathcal{H}}(r,t,\sigma) \right) \hat{\Psi}_{\mathcal{H}}(r,t,\sigma) \right], \tag{3.58}$$

where q denotes the elementary charge and the subscript $\mathcal{H}$ indicates the Heisenberg representation.

Using the definition of the lesser Green's function and taking the expectation value of the current operator, the current density can be expressed in terms of $G^{<}$ as

$$\mathbf{J}(r,t) = -\frac{q\hbar^2}{2m} \lim_{\substack{r' \to r \\ t' \to t^{+}}} (\nabla_r - \nabla_{r'}) \sum_{\sigma} G^{<}_{\sigma\sigma}(r,t;r',t'). \tag{3.59}$$

The divergence of the current density is required to establish the continuity equation. To rewrite this quantity in terms of the Hamiltonian, the non-interacting single-particle Hamiltonian is introduced as

$$\hat{H}_0(r) = \frac{\hat{\mathbf{p}}^2}{2m} + V(r), \tag{3.60}$$

where

$$\hat{\mathbf{p}} = -i\hbar\nabla_r \tag{3.61}$$

is the momentum operator and $V(r)$ represents the local potential experienced by the carriers. In the present case, the potential can include the contribution from the crystal lattice, e.g., $V(r) = \sum_i V(r - R_i)$.

Applying the divergence operator to the current density and using the definition of the kinetic-energy operator gives

$$\nabla_r \cdot \mathbf{J}(r,t) = q \lim_{\substack{r'\to r \\ t'\to t^+}} \left[\hat{H}_0(r) - \hat{H}_0(r')\right] \sum_\sigma G^<_{\sigma\sigma}(r,t;r',t'). \tag{3.62}$$

For a local potential, the potential contribution vanishes in the coincidence limit,

$$\lim_{r'\to r} \left[V(r) - V(r')\right] G^<(r,t;r',t') = 0, \tag{3.63}$$

and only the kinetic-energy operator contributes to the divergence of the current.

The continuity equation is obtained by combining the equations of motion for the Green's functions. These equations follow from the Dyson equation, Eqs. 3.13 and 3.14, which, after writing the equations of motion for the first and second space-time arguments and subtracting them, give

$$\begin{aligned} &\lim_{1'\to 1^+} i\hbar \left(\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_{1'}}\right) G(11') \\ &\quad - \lim_{1'\to 1^+} \left[\hat{H}_0(r_1) - \hat{H}_0(r_{1'})\right] G(11') \\ &= \lim_{1'\to 1^+} \int d2 \, \left[\Sigma(12)G(21') - G(12)\Sigma(21')\right]. \end{aligned} \tag{3.64}$$

Here, the notation

$$\lim_{1'\to 1^+} \equiv \lim_{\substack{r_{1'}\to r_1 \\ t_{1'}\to t_1^+}} \tag{3.65}$$

is adopted. The lesser component of the above equation is then obtained by applying the Langreth rules to the self-energy and Green's function products. Using Eq. 3.62, the resulting continuity equation can be written as

$$\begin{aligned} -\frac{\partial n(r_1,t_1)}{\partial t_1} - \nabla_r \cdot \mathbf{J}(r_1,t_1) = \lim_{1'\to 1^+} \int d2 \Big[&\Sigma^R(12)G^<(21') + \Sigma^<(12)G^A(21') \\ &- G^R(12)\Sigma^<(21') - G^<(12)\Sigma^A(21')\Big]. \end{aligned} \tag{3.66}$$

The right-hand side represents the net scattering contribution to the carrier balance. Depending on the specific self-energy employed, this term accounts for processes such as carrier injection from the contacts, phonon-assisted scattering, and recombination-generation processes. It therefore provides the microscopic NEGF counterpart of the recombination-generation term appearing in the macroscopic continuity equations. The specific generation and recombination processes considered in this work, together with their corresponding self-energies and models, will be discussed in a dedicated chapter.

In steady-state conditions,

$$\frac{\partial n}{\partial t} = 0, \tag{3.67}$$

and the Green's functions depend only on the time difference $\tau = t_1 - t_{1'}$. Fourier transforming

with respect to this time difference yields the energy-resolved form of the continuity equation:

$$\nabla \cdot \mathbf{J}(r,E) = -\frac{1}{2\pi\hbar}\int dr' \Big[\Sigma^R(r,r';E)G^<(r',r;E) + \Sigma^<(r,r';E)G^A(r',r;E) \\ - G^R(r,r';E)\Sigma^<(r',r;E) - G^<(r,r';E)\Sigma^A(r',r;E)\Big]. \tag{3.68}$$

This expression provides the spectral distribution of the current density divergence, which is particularly useful for analysing the energy-dependent transport and scattering processes in the simulated device. The total current density is obtained by integrating the energy-resolved current density over energy. Accordingly, the continuity equation in steady state is obtained by integrating Eq. 3.68 over the entire energy spectrum:

$$\nabla \cdot \mathbf{J}(r) = -\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dE \int dr' \Big[\Sigma^R(r,r';E)G^<(r',r;E) + \Sigma^<(r,r';E)G^A(r',r;E) \\ - G^R(r,r';E)\Sigma^<(r',r;E) - G^<(r,r';E)\Sigma^A(r',r;E)\Big]. \tag{3.69}$$

Terminal current While the current density describes local charge transport within the device, the physically measurable quantity in a transport experiment is the terminal current flowing between the device and the external reservoirs. Within the NEGF formalism, the current flowing from contact α into the device can be expressed through the Meir–Wingreen formula:

$$I_\alpha = \frac{q}{h}\int_{-\infty}^{\infty} dE\, \mathrm{Tr}\left[\Sigma_\alpha^<(E)G^>(E) - \Sigma_\alpha^>(E)G^<(E)\right], \tag{3.70}$$

where α identifies the contact through which the current is evaluated. The lesser and greater self-energies, $\Sigma_\alpha^<$ and $\Sigma_\alpha^>$, describe the injection and extraction of carriers from the corresponding reservoir, while $G^<$ and $G^>$ contain the occupied and unoccupied spectral information within the device, respectively.

For contacts described as non-interacting equilibrium reservoirs, the lesser self-energy can be written as

$$\Sigma_\alpha^<(E) = i f_\alpha(E)\Gamma_\alpha(E) \tag{3.71}$$

where $f_\alpha(E)$ is the Fermi–Dirac distribution of reservoir α, and

$$\Gamma_\alpha(E) = i\left[\Sigma_\alpha^R(E) - \Sigma_\alpha^A(E)\right] \tag{3.72}$$

is the broadening matrix associated with the coupling between the device and contact α. Accordingly, the total lesser self-energy due to all contacts is

$$\Sigma^<(E) = i\sum_\beta f_\beta(E)\Gamma_\beta(E) \tag{3.73}$$

where β is a summation index running over all reservoirs connected to the device. For a two-terminal device, for example, $\beta = L, R$.

In the absence of additional interaction self-energies contributing to the carrier balance, the lesser Green's function can be written as

$$G^<(E) = G^R(E)\Sigma^<(E)G^A(E). \tag{3.74}$$

Under these conditions, the Meir–Wingreen expression can be reduced to the Landauer–Caroli form:

$$I_\alpha = \frac{q}{h} \int_{-\infty}^{\infty} dE \sum_\beta T_{\alpha\beta}(E) \left[f_\alpha(E) - f_\beta(E) \right], \tag{3.75}$$

where

$$T_{\alpha\beta}(E) = \mathrm{Tr} \left[\Gamma_\alpha(E) G^R(E) \Gamma_\beta(E) G^A(E) \right] \tag{3.76}$$

is the energy-dependent transmission probability from reservoir β to reservoir α.

For a two-terminal device, with $\beta = L, R$, the contribution with $\beta = \alpha$ vanishes because $f_\alpha - f_\alpha = 0$. Therefore, the current flowing from the left contact into the device becomes

$$I_L = \frac{q}{h} \int_{-\infty}^{\infty} dE\, T_{LR}(E) \left[f_L(E) - f_R(E) \right], \tag{3.77}$$

and an analogous expression is obtained for the right contact. In the absence of charge accumulation, current conservation requires $I_L = -I_R$.

The Landauer–Caroli expression is particularly useful for analysing the energy-dependent transport properties of the device, since the transmission function directly identifies the energy ranges through which carriers are transmitted between the reservoirs. When additional scattering or recombination self-energies are included, the more general Meir–Wingreen formulation must instead be retained, with the corresponding interaction contributions included in the carrier balance.

Spectral current density It is convenient to define an energy-resolved current contribution (spectral current density):

$$i_\alpha(E) = \frac{q}{h} \sum_\beta \mathrm{Tr} \left[\Gamma_\alpha G^R \Gamma_\beta G^A \right] \left[f_\alpha(E) - f_\beta(E) \right], \tag{3.78}$$

such that the terminal current is obtained by integration:

$$I_\alpha = \int dE\, i_\alpha(E). \tag{3.79}$$

Current conservation In a self-consistent NEGF framework, current conservation is guaranteed by the identity:

$$\sum_\alpha I_\alpha = 0, \tag{3.80}$$

which follows directly from the properties of the self-energies and the Dyson–Keldysh structure.

More generally, the divergence of the current density satisfies:

$$\nabla \cdot \mathbf{J}(r) = - \int \frac{dE}{2\pi\hbar} \left[\Sigma^R G^< + \Sigma^< G^A - G^R \Sigma^< - G^< \Sigma^A \right], \tag{3.81}$$

which expresses local charge conservation in the presence of open boundaries and scattering processes.

3.7 Boundary self-energies

Up to this point, we have shown how the physical observables of a nanodevice can be obtained within the Green's function formalism, which naturally incorporates quantum-mechanical effects into the device description. However, a realistic modeling of electron and hole transport also requires the inclusion of two additional ingredients: carrier exchange between the device and the external contacts, and scattering processes occurring within the device.

To account for boundary effects, the system is partitioned into contacts, leads, and an active region [10] [14] [15] [16]. The contacts are modeled as highly doped, semi-infinite reservoirs characterized by fixed chemical potentials and thermal equilibrium distributions. The leads provide a smooth transition between the contacts and the active region, ensuring a proper description of band bending at the device boundaries.

The total Hamiltonian of the coupled system can be written as

$$H = \begin{bmatrix} H_D & \tau \\ \tau^\dagger & H_R \end{bmatrix}, \tag{3.82}$$

where H_D and H_R denote the Hamiltonians of the device and of the reservoir, respectively, while τ describes the coupling between them. The reservoir is assumed to be in thermal equilibrium and is characterized by a well-defined chemical potential and temperature.

The retarded Green's function is defined as

$$G^R = [(E + i\eta)\mathbb{I} - H]^{-1} = \begin{bmatrix} G^R_D & G^R_{DR} \\ G^R_{RD} & G^R_R \end{bmatrix}. \tag{3.83}$$

The next step consists in deriving the Green's function of the active region in the presence of the contacts. Since the contacts are modeled as semi-infinite equilibrium reservoirs, their effect can be incorporated through the surface Green's function of the isolated reservoirs, defined as

$$g^R(E) = [(E + i\eta)\mathbb{I} - H_R]^{-1}, \tag{3.84}$$

where H_R is the reservoir Hamiltonian and $\eta \to 0^+$ enforces the correct retarded boundary conditions.

By solving the block matrix system, one obtains

$$G^R_{DR} = G^R_D \, \tau \, g^R, \tag{3.85}$$

$$G^R_D = \left[(E + i\eta)\mathbb{I} - H_D - \tau^\dagger g^R \tau\right]^{-1} \equiv \left[(E + i\eta)\mathbb{I} - H_D - \Sigma^{RB}\right]^{-1}. \tag{3.86}$$

From this comparison, the retarded boundary self-energy is identified as

$$\boxed{\Sigma^{RB} = \tau g^R \tau^\dagger}. \tag{3.87}$$

Starting from the Keldysh relation in matrix form,

$$G^< = G^R \Sigma^< G^A, \tag{3.88}$$

the same procedure yields the boundary contributions

$$\boxed{\Sigma^{<B} = \tau g^{<} \tau^{\dagger}}, \qquad \boxed{\Sigma^{>B} = \tau g^{>} \tau^{\dagger}}. \tag{3.89}$$

Since the contacts are assumed to be in thermal equilibrium, their lesser and greater Green's functions satisfy the fluctuation-dissipation theorem:

$$g^{<} = -f_{\mu}(E)\,[g^{R} - g^{A}], \qquad g^{>} = (1 - f_{\mu}(E))\,[g^{R} - g^{A}], \tag{3.90}$$

where $f_{\mu}(E)$ is the Fermi-Dirac distribution associated with the chemical potential μ of the reservoir.

3.8 Scattering self-energies

The treatment of scattering within the NEGF formalism relies on the self-consistent Born approximation (SCBA). This approach assumes a single-particle picture in which carriers move in an effective potential generated by the surrounding particles and scattering mechanisms. Since an exact evaluation of the interaction self-energies is generally not feasible for realistic systems, the Born approximation is employed by retaining the lowest-order scattering contributions to the self-energy. In its self-consistent formulation, the resulting self-energies are evaluated using the interacting Green's functions and subsequently updated iteratively. Starting from an initial estimate, the Green's functions and self-energies are recalculated in successive Born iterations until self-consistency is reached.

For electron–phonon scattering, the lesser and greater self-energies are given by

$$\Sigma^{<,>}(k,E) = \sum_{q} M^{ep}(q)\Big[N_q\, G^{<,>}(k - q_{\parallel}, E \mp \hbar\omega_q) + (N_q + 1)\, G^{<,>}(k - q_{\parallel}, E \pm \hbar\omega_q)\Big] M^{ep}(-q), \tag{3.91}$$

where the upper and lower signs correspond to the lesser and greater components, respectively.

The retarded self-energy can be obtained from its time-domain definition, in analogy with the retarded and advanced Green's functions [17]. In particular, the latter are related to the lesser and greater Green's functions according to

$$G^{R}(11') = \theta(t_1 - t_{1'}) \left[G^{>}(11') - G^{<}(11')\right], \tag{3.92}$$

$$G^{A}(11') = -\theta(t_{1'} - t_1) \left[G^{>}(11') - G^{<}(11')\right]. \tag{3.93}$$

An analogous relation holds for the self-energy,

$$\Sigma^{R}(t) = \theta(t) \left[\Sigma^{>}(t) - \Sigma^{<}(t)\right]. \tag{3.94}$$

Introducing the relative time coordinate $\tau = t_1 - t_{1'}$, the energy representation is obtained through the Fourier transform

$$\Sigma^{R}(E) = \int_{-\infty}^{+\infty} d\tau\, e^{iE\tau/\hbar} \theta(\tau) \left[\Sigma^{>}(\tau) - \Sigma^{<}(\tau)\right]. \tag{3.95}$$

The Fourier transform of the Heaviside function is a distribution containing both a delta-function and a principal-value contribution:

$$\int_{-\infty}^{+\infty} d\tau\, e^{i(E-E')\tau/\hbar} \theta(\tau) = \pi\hbar\,\delta(E - E') + i\hbar\,\mathcal{P}\left(\frac{1}{E - E'}\right), \tag{3.96}$$

where $\mathcal{P}$ denotes the Cauchy principal value. Consequently, after applying the convolution theorem, the retarded self-energy can be written in the general form

$$\begin{aligned}\Sigma^R(E) = \frac{1}{2}\left[\Sigma^>(E) - \Sigma^<(E)\right] \\ + \frac{i}{2\pi}\mathcal{P}\int_{-\infty}^{+\infty} dE' \, \frac{\Sigma^>(E') - \Sigma^<(E')}{E - E'}.\end{aligned} \tag{3.97}$$

The second term represents the dispersive contribution associated with the principal-value integral and contributes to the real part of the retarded self-energy. If this contribution is neglected, the retarded self-energy is approximated by

$$\boxed{\Sigma^R(k,E) \simeq \frac{1}{2}\left[\Sigma^>(k,E) - \Sigma^<(k,E)\right].} \tag{3.98}$$

This approximation is the one adopted in the present model.

Here, $\hbar\omega_q$ is the phonon energy, while $M^{ep}(q)$ denotes the electron–phonon coupling matrix. The quantity N_q is the phonon occupation number, given by the Bose–Einstein distribution,

$$N_q = \frac{1}{\exp\left(\hbar\omega_q/k_BT\right) - 1}, \tag{3.99}$$

which represents the average occupation of the phonon mode with wave vector q.

The electron–phonon matrix elements are defined as

$$M^{ep}_{m,n}(q) = \frac{1}{\sqrt{V}} U_q e^{iq_z(\Delta L + v_{\alpha,z})}, \tag{3.100}$$

where ΔL is the layer spacing and $v_{\alpha,z}$ denotes the z-coordinate of atom α within the unit cell. The quantity V is the normalization volume, while U_q represents the interaction strength associated with the phonon mode q, whose specific form depends on the scattering mechanism considered.

Acoustic phonon scattering self-energy

Assuming a linear dispersion relation for acoustic phonons, the electron–phonon interaction strength can be expressed as

$$|U_q|^2 = \sqrt{\frac{\hbar D_\alpha^2 q}{2V\rho u_l}}, \tag{3.101}$$

where u_l is the longitudinal sound velocity, D_α is the acoustic deformation potential, V is the normalization volume, and ρ is the mass density of the material. In the high temperature limit $N_q \approx N_q + 1 \approx \frac{k_BT}{\hbar\omega_q}$, the self-energy can be written as:

$$\Sigma^{<,>,R}_{ij} = \frac{1}{\Delta_{ij}} \frac{D_\alpha^2 k_B T}{V\rho u_l^2} \sum_{q_{||}} [G^{<,>,R}(q_{||}, E)]_{ij} \tag{3.102}$$

where

$$\frac{1}{\Delta_{ij}} = \sum_{q_z} e^{iq_z(z_i - z_j)} = \frac{\mathcal{L}}{2\pi}\int_{-\pi/2}^{\pi/2} dq_z \;\; e^{iq_z(z_i - z_j)} \approx \delta_{ij}\frac{\mathcal{L}}{a} \tag{3.103}$$

The quantities appearing in the previous expression are defined as follows: $\Sigma^{<,>,R}_{ij}(E)$ denotes

the lesser, greater, and retarded acoustic phonon self-energies between lattice sites i and j. The term k_BT represents the thermal energy, arising from the high-temperature approximation of the phonon occupation. The parameters V and ρ are respectively the normalization volume and the mass density of the material, while u_l is the longitudinal sound velocity. The summation over $q_\parallel$ accounts for all in-plane phonon wavevectors, whereas the dependence on q_z has been absorbed into the factor Δ_{ij}, which enforces (approximately) the locality of the interaction in real space, leading to an almost diagonal self-energy in the site representation. Here, $\mathcal{L}$ denotes the length of the simulated structure along the z-direction, while a is the spatial discretization step. Finally, $G^{<,>,R}(q_\parallel, E)$ denotes the corresponding Green's functions resolved in the in-plane momentum representation.

Polar optical phonon scattering self-energy

According to Fröhlich theory, the electron–phonon coupling strength associated with polar optical phonons is given by

$$U_Q = \sqrt{\frac{e^2\hbar\omega_{LO}}{2V}\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right)}\frac{q}{q^2+q_0^2}, \tag{3.104}$$

where $\hbar\omega_{LO}$ is the longitudinal optical (LO) phonon energy, V is the normalization volume, and ε_∞ and ε_s are the high-frequency and static dielectric constants of the material, respectively. The term

$$q_0 = \sqrt{\frac{e^2 n}{\varepsilon_s k_B T}} \tag{3.105}$$

represents the inverse Debye screening length, where n is the carrier concentration, k_B is the Boltzmann constant, and T is the lattice temperature. The screening wavevector q_0 accounts for the reduction of the long-range Coulomb interaction due to the presence of free carriers.

Under these assumptions, the lesser and greater self-energies associated with polar optical phonon scattering can be expressed as

$$\begin{aligned}\Sigma_{ij}^{<,>}(E) = \frac{e^2\hbar\omega_{LO}}{4\pi^2}\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right)\int dq\, q\, F(q, s_{ij}, k)\Big[&n_{LO}\, G_{ij}^{<,>}(q, E \pm \hbar\omega_{LO}) \\ &+ (n_{LO}+1)\, G_{ij}^{<,>}(q, E \mp \hbar\omega_{LO})\Big],\end{aligned} \tag{3.106}$$

where n_{LO} is the occupation number of longitudinal optical phonons, evaluated through the Bose–Einstein distribution,

$$n_{LO} = \frac{1}{e^{\hbar\omega_{LO}/k_BT} - 1}, \tag{3.107}$$

while $F(q, s_{ij}, k)$ is a form factor accounting for the screened Fröhlich interaction and the spatial dependence of the scattering process. The quantity s_{ij} denotes the separation between the spatial positions associated with the matrix elements i and j, whereas k represents the electronic wavevector.

The first term inside the square brackets describes phonon absorption processes, in which an electron gains an energy $\hbar\omega_{LO}$, while the second term accounts for phonon emission processes, corresponding to an energy loss of $\hbar\omega_{LO}$.

Within the present implementation, the retarded optical phonon self-energy is evaluated from the retarded Green's function at the two energy-shifted arguments. In particular, it is written as

$$\Sigma^R_{\mathrm{opt},ij}(E) = N_{2D} C_{\mathrm{opt}} \int dq\, q\, F(q, s_{ij}, k) \Big[n_{LO}\, i\, \mathrm{Im}\, G^R_{ij}(q, E + \hbar\omega_{LO}) + (n_{LO} + 1)\, i\, \mathrm{Im}\, G^R_{ij}(q, E - \hbar\omega_{LO}) \Big], \tag{3.108}$$

where N_{2D} accounts for the integration over the transverse electronic states and C_{opt} is the coupling prefactor associated with the optical deformation-potential interaction. In the numerical implementation, this prefactor is given by

$$C_{\mathrm{opt}} = \frac{\hbar^2 D_t^2}{2\rho_{\mathrm{mat}} E_{ph} a_{\mathrm{lat}}}, \tag{3.109}$$

with D_t denoting the optical deformation-potential constant, ρ_{mat} the material mass density, $E_{ph} = \hbar\omega_{LO}$ the phonon energy, and a_{lat} the lattice constant.

The retarded self-energy therefore depends on the spectral contribution of the retarded Green's function evaluated at energies shifted by $\pm\hbar\omega_{LO}$. This energy dependence accounts explicitly for the absorption and emission of LO phonons and introduces the corresponding energy broadening into the electronic states.

The inclusion of boundary and scattering self-energies completes the NEGF formulation, enabling a fully open quantum system description of carrier transport in nanoscale devices.

3.9 Büttiker probe model

Although the self-consistent Born approximation provides a rigorous framework for describing carrier scattering within the NEGF formalism, the explicit evaluation of scattering self-energies can become computationally demanding, especially when multiple scattering mechanisms are considered simultaneously. An alternative and computationally efficient approach consists in introducing fictitious reservoirs, known as Büttiker probes, which emulate incoherent scattering processes while preserving the overall efficiency of the simulation [18][19].

In this framework, scattering is modeled as an exchange of carriers between the device and a set of virtual contacts distributed throughout the system. Each probe acts as a local reservoir characterized by its own chemical potential (and, if required, temperature). Carriers entering a probe lose phase coherence and are re-injected into the device, thereby reproducing dephasing and dissipative effects.

The coupling between the device and the probes is introduced through self-energy terms in the NEGF equations. Unlike physical contacts, Büttiker probes are constrained by the condition of zero net current, ensuring charge conservation while still capturing the effect of scattering processes on transport.

At each spatial point z_p of the discretized device ($p = 1, \ldots, N$), a Büttiker probe is attached. The corresponding self-energy is written as the sum of all local contributions:

$$\Sigma^R_{BP}(z, z, E) = \sum_{p=1}^{N} \left[-i \frac{\hbar}{\tau_p} \right] \delta(z - z_p), \tag{3.110}$$

where τ_p is an effective local scattering time and the delta function enforces spatial locality. In this formulation, $B(z, E)$ represents the effective scattering strength, which may depend on both

position and energy.

To ensure current and charge conservation, each probe is associated with a local chemical potential that is determined self-consistently such that the net current flowing into each probe vanishes. This constraint guarantees particle conservation while allowing energy and phase relaxation.

Despite its simplicity, this formulation introduces phenomenological scattering parameters and does not explicitly distinguish between different microscopic scattering mechanisms. More advanced formulations, such as multi-mechanism Büttiker probes, aim to overcome this limitation by deriving scattering strengths from material properties and physical models rather than fitting parameters.

The key idea of these extended approaches is to replace phenomenological relaxation times with quantities computed self-consistently within the NEGF framework, thereby improving predictive capability.

A further simplification of the Büttiker-probe model consists in assuming that the scattering self-energies depend only on the retarded Green's function G^R, without an explicit dependence on the lesser Green's function $G^<$. This approximation follows from the assumption that the probes behave as local quasi-equilibrium reservoirs at finite temperature. Under this assumption, the fluctuation-dissipation relation (see Eq. 3.37) can be used to relate the occupation-dependent Green's functions to the spectral function. Consequently, the calculation of the retarded Green's function can be decoupled from the self-consistent evaluation of the lesser Green's function, significantly reducing the computational cost. Moreover, within the local scattering approximation, only the diagonal elements of G^R are required to evaluate the Büttiker-probe self-energies. These diagonal elements directly determine the local spectral function, and hence the local density of states.

Although this approximation simplifies the numerical scheme, the resulting equations remain self-consistent since the self-energies still depend on G^R, which in turn depends on the self-energies themselves.

Within this framework, the acoustic phonon scattering self-energy in the high-temperature limit ($k_B T \gg \hbar\omega_q$) and assuming linear dispersion can be written as

$$\Sigma^R_{BP,\mathrm{aco}}(z_p, E) = \frac{D^2_{\mathrm{aco}} k_B T}{\rho u_l^2} \frac{1}{2\hbar\omega_D} \int_{E-\hbar\omega_D}^{E+\hbar\omega_D} G^R(z_p, z_p, E')\, dE', \tag{3.111}$$

where ω_D is the Debye frequency and z_p denotes the position of the probe.

For numerical implementation, the weak energy dependence within the Debye window is often neglected, leading to the approximation

$$\Sigma^R_{BP,\mathrm{aco}}(z_p, E) \approx \frac{D^2_{\mathrm{aco}} k_B T}{\rho u_l^2} \frac{1}{a} G^R(z_p, z_p, E), \tag{3.112}$$

where a is the lattice spacing.

For polar optical phonon scattering, the Büttiker probe self-energy can be expressed as

$$\begin{aligned}\Sigma^R_{BP,\mathrm{pol}}(z_p, E) = \frac{\hbar (D_t K)^2}{2\rho\omega_{LO} a} \Big[&(N_{LO} + 1)\, G^R(z_p, z_p, E - \hbar\omega_{LO}) \\ &+ N_{LO}\, G^R(z_p, z_p, E + \hbar\omega_{LO}) \Big],\end{aligned} \tag{3.113}$$

where $D_t K$ is the optical deformation potential and N_{LO} is the Bose–Einstein occupation factor.

The corresponding lesser self-energy is obtained using the fluctuation–dissipation relation:

$$\Sigma^{<}_{BP}(z_p, z_p, E) = i\frac{m^*_n}{\pi\hbar^2}k_B T\mathcal{F}_0\left(\frac{E - E_{BP}}{k_B T}\right)\Gamma_{BP}(z_p, z_p, E), \tag{3.114}$$

where $\mathcal{F}_0$ is the zero-order Fermi integral and

$$\Gamma_{BP}(z_p, z_p, E) = i\left[\Sigma^R_{BP}(z_p, z_p, E) - \Sigma^A_{BP}(z_p, z_p, E)\right]. \tag{3.115}$$

4 Finite-difference implementation

Now that the theoretical framework has been established, the non-equilibrium Green's function (NEGF) simulator can be constructed to evaluate the main physical observables of a nanoelectronic device.

After solving the device using the Poisson–drift–diffusion (DD) model, an approximate band diagram is obtained. However, quantum transport effects are only partially captured within this semiclassical description. For this reason, a NEGF-based implementation is required to achieve a more accurate and physically complete representation of carrier transport at the nanoscale.

Similarly to the DD approach, the numerical solution is obtained through a Newton–Raphson scheme. The first step consists in discretizing the one-dimensional device into n spatial segments. According to Eq. A.9, it is necessary to identify both the system unknowns and the set of governing equations (residuals) that define the NEGF transport problem.

The objective of the simulation is to solve the nonlinear system

$$\underline{r}(\underline{u}) = \underline{0}, \tag{4.1}$$

where the residual vector is defined as

$$\underline{r} = \begin{bmatrix} \underline{r}_p \\ \underline{r}_e \\ \underline{r}_h \\ r_v \\ r_{jl} \\ r_{jr} \\ r_{nl} \\ r_{nr} \\ r_{pl} \\ r_{pr} \end{bmatrix}, \qquad \underline{u} = \begin{bmatrix} \underline{U} \\ \underline{\mu}_n \\ \underline{\mu}_p \\ V \\ J_L \\ J_R \\ \mu_{nl} \\ \mu_{nr} \\ \mu_{pl} \\ \mu_{pr} \end{bmatrix}. \tag{4.2}$$

The meaning of the residuals is the following:

- $\underline{r}_p$: Poisson equation;
- $\underline{r}_e, \underline{r}_h$: electron and hole continuity equations derived from NEGF;
- r_v: circuit equation defining the applied bias;
- r_{jl}, r_{jr}: total current constraints at left and right contacts;
- $r_{nl}, r_{nr}, r_{pl}, r_{pr}$: boundary consistency conditions enforcing continuity between device and reservoir currents.

The unknowns are defined as:

- $\underline{U} = -q\underline{\phi}$, the electrostatic energy vector;
- $\underline{\mu}_n, \underline{\mu}_p$: scattering-related quasi-Fermi levels for electrons and holes;
- V: applied terminal voltage;
- J_L, J_R: terminal currents at left and right contacts;

- $\mu_{nl}, \mu_{nr}, \mu_{pl}, \mu_{pr}$: contact quasi-Fermi levels for electrons and holes.

Recalling Eq. A.10, the Newton update reads

$$\underline{\Delta x}^{(k+1)} = -\underline{\underline{J}}^{-1}\underline{r}^{(k)}. \tag{4.3}$$

In addition to defining the residual equations, the construction of the Jacobian matrix represents one of the main numerical challenges. Each entry requires the evaluation of partial derivatives of all residuals with respect to all unknowns, leading to a highly coupled nonlinear system.

The Jacobian matrix is defined as

$$J_{ij} = \frac{\partial r_i}{\partial u_j}. \tag{4.4}$$

Before explicitly deriving all contributions, several structural properties can be exploited to simplify the implementation:

- the applied voltage V directly affects only the electrostatic and transport-related residuals;
- the terminal currents J_L and J_R act as auxiliary variables and do not enter the internal device equations directly;
- current continuity is enforced through boundary conditions, avoiding the need for explicit coupling between left and right terminal currents inside the Jacobian.

These considerations reveal that the Jacobian matrix has a sparse block structure, determined by the underlying physical couplings. Exploiting this sparsity is essential to reduce computational cost, improve memory efficiency, and accelerate convergence of the Newton iterations.

Having established the general numerical framework, the following sections present the detailed implementation of the simulator, with particular emphasis on the formulation and structure of the residual equations.

4.1 System partitioning

The first step consists of partitioning the device into three main regions: the two contacts and the internal device region. The contacts are modeled as semi-infinite carrier reservoirs, each characterized by its own Hamiltonian, $\hat{H}_L$ for the left and $\hat{H}_R$ for the right side, while $\hat{H}_D$ describes the Hamiltonian of the internal device, which represents the region of primary interest for electronic transport.

The coupling between the contacts and the device is introduced through the interaction operators $\hat{H}_{LD}$ and $\hat{H}_{RD}$, which account for carrier exchange between the reservoirs and the active region. Under this partitioning scheme, the total quantum system can be represented through the stationary Schrödinger equation written in block matrix form:

$$\begin{pmatrix} \hat{H}_L & \hat{H}_{LD} & 0 \\ \hat{H}_{LD} & \hat{H}_D & \hat{H}_{rD} \\ 0 & \hat{H}_{RD} & \hat{H}_R \end{pmatrix} \begin{bmatrix} |\psi_L\rangle \\ |\psi_D\rangle \\ |\psi_R\rangle \end{bmatrix} = E \begin{bmatrix} \psi_L \\ \psi_D \\ \psi_R \end{bmatrix} \tag{4.5}$$

where ψ_L and ψ_R are the wavefunctions of the contacts, ψ_D is the device wavefunction and E is the carrier energy. Compactly, the full quantum transport problem can be written as:

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{4.6}$$

Recalling the definition of the retarded Green's function, we can write:

$$(E\hat{I} - \hat{H})\hat{G}^R = \hat{I} \tag{4.7}$$

The problem can be reformulated in terms of Green's functions rather than direct wavefunction solutions. This transformation is particularly advantageous because it naturally incorporates open boundary conditions and contact effects.
Expanding the previous expression in block matrix form yields

$$\begin{pmatrix} E\hat{I} - \hat{H}_L & -\hat{H}_{LD} & 0 \\ -\hat{H}_{LD} & E\hat{I} - \hat{H}_D & -\hat{H}_{rD} \\ 0 & -\hat{H}_{RD} & E\hat{I} - \hat{H}_R \end{pmatrix} \begin{pmatrix} \hat{G}^R_L & \hat{G}^R_{LD} & \hat{G}^R_{LR} \\ \hat{G}^R_{LD} & \hat{G}^R_D & \hat{G}^R_{RD} \\ \hat{G}^R_{LR} & \hat{G}^R_{RD} & \hat{G}^R_R \end{pmatrix} = \begin{bmatrix} \hat{I} & 0 & 0 \\ 0 & \hat{I} & 0 \\ 0 & 0 & \hat{I} \end{bmatrix} \tag{4.8}$$

Since the primary objective is to describe carrier transport within the active region of the device, the quantity of main interest is the device retarded Green's function, $\hat{G}^R_D$. It can be shown, by introducing the retarded self-energies of the contacts, that it is equal to:

$$\hat{G}^R_D = [E\hat{I} - \hat{H}_D - \hat{\Sigma}^R_L - \hat{\Sigma}^R_R]^{-1} \tag{4.9}$$

The self-energy terms are energy-dependent operators accounting for:

- carrier injection from the contacts,
- carrier extraction into the contacts,
- level broadening,
- energy renormalization at the device boundaries.

To construct the final expression of the retarded Green's function, it is convenient to introduce the compact notation

$$\hat{h}_D = E\hat{I} - \hat{H}_D \tag{4.10}$$

where $\hat{h}_D$ is the Hamiltonian operator cleaned by contact and scattering corrections. To account for non-ideal transport effects, scattering mechanisms are incorporated through a Büttiker-probe-inspired formalism. In this approach, the scattering contribution is modeled by additional self-energy terms, which are directly added to the boundary self-energies associated with the physical contacts.

$$\hat{G}^R_D = [\hat{h}_D - \hat{\Sigma}_R]^{-1} \tag{4.11}$$

The total retarded self-energy can therefore be written as:

$$\hat{\Sigma}^R = \hat{\Sigma}^R_L + \hat{\Sigma}^R_R + \hat{\Sigma}^R_{BP} \tag{4.12}$$

4.2 Hamiltonian assembling

Several numerical techniques can be employed to solve the quantum carrier transport problem in nanoscale devices under spatial discretization, including finite element methods (FEM), finite box methods (FB), and finite difference (FD) schemes. Among these approaches, the finite difference method represents an effective compromise between physical rigor, mathematical simplicity, and computational efficiency, while also providing a straightforward implementation of the NEGF formalism.

The choice of a FD discretization is particularly advantageous because it transforms the continuous Schrödinger equation into a sparse matrix eigenvalue problem with a clear physical interpretation. By discretizing the device along a one-dimensional spatial grid composed of n nodes with mesh spacing Δz, the device Hamiltonian can be written as

$$\hat{H}_D = -\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + U(z), \tag{4.13}$$

where m^* is the effective mass (assumed constant for each material region in this work), and $U(z)$ denotes the electrostatic potential profile.

The second derivative of the Hamiltonian is approximated using a finite difference scheme:

$$\frac{d^2\psi(z)}{dz^2} \approx \frac{\psi_{i+1} - 2\psi_i + \psi_{i-1}}{(\Delta z)^2}, \tag{4.14}$$

with $\psi_i = \psi(z_i)$.

This leads to the discrete equation at node i:

$$(H_D\psi)_i = \left[\frac{\hbar^2}{m^*(\Delta z)^2} + U(z_i) + \Delta E_{C,i}\right]\psi_i - \frac{\hbar^2}{2m^*(\Delta z)^2}\left(\psi_{i+1} + \psi_{i-1}\right). \tag{4.15}$$

The $n \times n$ matrix form of the Hamiltonian becomes tridiagonal:

$$\underline{\underline{H}}_D = \begin{pmatrix} H_0 & H_1 & 0 & \cdots & 0 \\ H_1 & H_0 & H_1 & \ddots & \vdots \\ 0 & H_1 & H_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & H_1 \\ 0 & \cdots & 0 & H_1 & H_0 \end{pmatrix}, \tag{4.16}$$

where

$$H_{0,i} = \frac{\hbar^2}{m^*(\Delta z)^2} + U(z_i) + \Delta E_{C,i}, \tag{4.17}$$

$$H_{1,i} = -\frac{\hbar^2}{2m^*(\Delta z)^2}. \tag{4.18}$$

The diagonal corner elements are not divided by two; that statement is incorrect for a standard FD discretization.

To improve physical realism, transverse carrier motion is included via separation of variables:

$$E = E_z + E_t, \tag{4.19}$$

with transverse energy:

$$E_t = \frac{\hbar^2 k_t^2}{2m^*}, \tag{4.20}$$

where k_t is the transverse wavevector. The upper limit of the integration over transverse wavevectors is typically of the order of $2 \times 10^9\,\mathrm{m}^{-1}$.

Finally, the operator $\hat{h}_D = E\hat{I} - \hat{H}_D$ in matrix form reads:

$$\underline{\underline{h}}_D = \begin{pmatrix} h_0 & h_1 & 0 & \cdots & 0 \\ h_1 & h_0 & h_1 & \ddots & \vdots \\ 0 & h_1 & h_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & h_1 \\ 0 & \cdots & 0 & h_1 & h_0 \end{pmatrix}, \tag{4.21}$$

where

$$h_{0,i} = E - H_{0,i}, \tag{4.22}$$

$$h_{1,i} = -H_{1,i}. \tag{4.23}$$

4.3 Self-energy implementation

The final ingredient required to determine the retarded Green's function matrix of Eq. 4.11 is the total retarded self-energy operator, $\hat{\Sigma}_R$. Since the boundary self-energies are themselves influenced by carrier scattering processes occurring at and near the contacts, the first step is to properly define the scattering self-energy contributions before constructing the full contact self-energy terms.

$$\underline{\underline{\Sigma}}^R_{BP} = \underline{\underline{\Sigma}}^R_{opt} + \underline{\underline{\Sigma}}^R_{aco} \tag{4.24}$$

Inelastic optical deformation potential scattering.

Recalling the Büttiker-probes of Chapter 3.7, the inelastic optical deformation potential scattering self-energy under matrix notation is:

$$\begin{aligned} \underline{\underline{\Sigma}}^R_{opt}(E) =& C_{opt}[(n_{ph}E_{t,max} + \mathcal{F}_0(E + E_{ph}, \mu_n))\; diag\left\{ i\,\mathbb{I}m\left\{ \underline{\underline{G}}^R(E + E_{ph}) \right\} \right\} \\ &+ (n_{ph} + 1)E_{t,max} + \mathcal{F}_0(E - E_{ph}, \mu_n)\; diag\left\{ i\,\mathbb{I}m\left\{ \underline{\underline{G}}^R(E - E_{ph}) \right\} \right\}] \end{aligned} \tag{4.25}$$

The retarded self-energy is related to the advanced self-energy through Eq. 3.115. Therefore, its anti-Hermitian part can be expressed in terms of its imaginary part. Within the local scattering approximation adopted for the Büttiker probes, only the diagonal elements of the retarded Green's function are required.

The function $\mathcal{F}_0$ results from the numerical evaluation of the Fermi–Dirac integral of order zero and accounts for the occupation of the electronic states involved in the phonon absorption and emission processes:

$$\mathcal{F}_0(E, \mu) = N^{2D} k_B T \left\{ log\left[1 + exp(\frac{\mu - E}{k_B T}) \right] - log\left[1 + exp(\frac{\mu - E - E_{t,max}}{k_B T}) \right] \right\} \tag{4.26}$$

Thus both phonon absorption ($E + E_{ph}$) and emission ($E - E_{ph}$) are incorporated. This provides a physically consistent description of inelastic carrier transitions induced by optical phonon interactions. However, from a numerical standpoint, an additional level of complexity arises during the optical scattering self-energy evaluation. In particular, the optical scattering self-energy at a given carrier energy E explicitly depends on the retarded Green's function evaluated at shifted energy levels: this requires achieving self-consistency between the retarded Green's function and the self-energies, making the problem no longer directly solvable through a single linear inversion. The phonon energy is expressed as:

$$E_{ph} = \hbar\omega_q \tag{4.27}$$

approximately equal to $30\,meV$. The phonon occupation number requires just the temperature of the system besides the phonon energy:

$$n_{ph} = \frac{1}{exp(\frac{E_{ph}}{k_B T}) - 1} \tag{4.28}$$

Finally, all material-dependent and coupling-related parameters are grouped into a single effective constant, which defines the strength of the electron–phonon interaction in the deformation potential approximation:

$$C_{opt} = \frac{hD_{tK}^2}{2\rho_m \omega_q a_{lat}} \tag{4.29}$$

Elastic acoustic deformation potential scattering.

Similar discussion applies to elastic acoustic deformation-potential scattering, but the relative self-energy depends only on the energy level involved, meaning that no Newton-Raphson algorithm is required:

$$\underline{\underline{\Sigma}}^R_{aco}(E) = C_{aco} N^{2D}\ diag\left\{ i\ \mathbb{I}m\left\{ \underline{\underline{G}}^R(E) \right\} \right\} \tag{4.30}$$

where

$$C_{aco} = \frac{D_{aco} k_B T}{\rho_m v_s^2 a_{lat}} \tag{4.31}$$

D_{aco} is the acoustic deformation potential and v_s the sound velocity.

Boundary self-energies

To find an expression for the boundary self-energies, whose role is to implement the carrier injection from the contacts, we recall the hypothesis of a semi-infinitely extended carrier reservoir. The contact is treated as a bulk region where the effective mass m^* still remains the same as that of the internal device. Let's focus on the left contact (the right one follows the same reasoning):

$$\left[\begin{array}{cccc|c} \ddots & \ddots & & & \\ \ddots & h_0 & h_1 & & \\ & h_1 & h_0 & h_1 & \\ & & h_1 & h_0 & h_1 \\ \hline & & & h_1 & h_0 \end{array}\right] = \left[\begin{array}{c|c} \underline{\underline{h}}_L & \underline{h}_{LD} \\ \hline \underline{h}_{LD}^\dagger & h_{DBL} \end{array}\right] \tag{4.32}$$

where $\underline{\underline{h}}_L$ is the semi-infinite left contact Hamiltonian (still no scattering terms are considered), $\underline{h}_{LD}$ is the coupling vector between contact and device, h_{DBL} is the left boundary node of the

device Hamiltonian. By recalling Eq. 4.11, the Green's function associated with the left boundary device node is:

$$G^R_{DBL} = [h_{DBL} - \Sigma^R_L]^{-1} \tag{4.33}$$

with the left boundary self-energy defined as:

$$\underline{\underline{\Sigma}}^R_L = \underline{h}^\dagger_{LD} \underline{\underline{g}}^R_L \underline{h}_{LD} \tag{4.34}$$

where $\underline{\underline{g^R_L}}$ is the retarded Green's function of the isolated semi-infinite left contact. Since the coupling vector $\underline{h}_{LD}$ contains only one nonzero term, corresponding to the nearest-neighbor coupling parameter h_1, only the bottom-right surface element of the contact Green's function contributes. Denoting this boundary term as g^R_{LB}, the matrix product simplifies considerably:

$$\Sigma^R_L = h_1 g^R_{LB} h_1 \tag{4.35}$$

Thus, the entire influence of the semi-infinite left reservoir is condensed into a single complex scalar term acting on the boundary node of the device. Similarly to Eq. 4.33, the left boundary Green's function id defined as:

$$g^R_{LB} = [h_{LB} - \Sigma^R_L]^{-1} \tag{4.36}$$

By applying the definition in Eq. 4.35, the left boundary self energy becomes:

$$\Sigma^R_L = h_1 [h_{LB} - \Sigma^R_L]^{-1} h_1 \tag{4.37}$$

by recognizing the bottom-right element of $\underline{\underline{h}}_L$, h_{LB} has to be equal to h_0, while the equation can be algebraically re-written as:

$$(\Sigma^R_L)^2 - h_0 \Sigma^R_L + h_1^2 = 0 \tag{4.38}$$

The second degree equation can be solved:

$$\Sigma^R_L = \frac{-h_0 \pm \sqrt{h_0^2 - 4h_1^2}}{-2} \tag{4.39}$$

Since the Büttiker-probe self-energies are diagonal within the local scattering approximation, their effect on the semi-infinite contact can be incorporated by renormalizing the on-site term of the contact Hamiltonian according to:

$$h_0 \to \tilde{h_0} = h_0 - \Sigma^R_{BP,L} \tag{4.40}$$

This procedure effectively incorporates the scattering processes occurring inside the contact reservoir while preserving the analytical structure of the surface Green's function formulation. Without this correction, the energy levels in the device are broadened by the scattering self-energies, whereas those in the contact remain sharp. This artificial mismatch in the spectral broadening at the device–contact interface leads to spurious reflections, since carriers propagating from the device encounter contact states with incompatible spectral properties. Including the Büttiker-probe self-energy in the contact Hamiltonian ensures that both the device and the reservoir exhibit a consistent lifetime broadening, thereby providing a physically consistent open boundary condition.

Finally, the surface Green's function becomes:

$$\Sigma_L^R = \frac{-\tilde{h_0} - \sqrt{\tilde{h_0}^{\,2} - 4h_1^2}}{-2} \tag{4.41}$$

The same procedure is applied to the right contact, with the appropriate replacement of boundary indices. Among the mathematical solutions, the physically meaningful one is selected by imposing the retarded condition, $\mathrm{Im}\left(\Sigma_L^R\right) \leq 0$, which guarantees carrier escape into the semi-infinite reservoir. A positive imaginary part would instead imply the unphysical generation of carriers from the reservoir, leading to an artificial injection of infinite energy from the contacts.

The complete boundary self-energy accounts for the contributions of both the left and right contacts and takes the following form:

$$\underline{\underline{\Sigma}}^R = \begin{pmatrix} \Sigma_L^R & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \Sigma_R^R \end{pmatrix}. \tag{4.42}$$

By including the Büttiker-probe self-energies, the total retarded self-energy becomes

$$\boxed{\underline{\underline{\Sigma}}^R_{\mathrm{tot}}(E) = \underline{\underline{\Sigma}}^R(E) + \underline{\underline{\Sigma}}^R_{BP}(E)} \tag{4.43}$$

thus fully defining the retarded self-energy of the system, whose matrix is diagonal under the assumption of locally confined scattering processes.

Recalling Eqs. 3.39 and 3.41, the lesser self-energy is defined as

$$\underline{\underline{\Sigma}}^<(E) = i\, f(E - E_F)\, \underline{\underline{\Gamma}}(E). \tag{4.44}$$

The advanced self-energy is the Hermitian conjugate of the retarded one,

$$\underline{\underline{\Sigma}}^A(E) = \left(\underline{\underline{\Sigma}}^R(E)\right)^\dagger, \tag{4.45}$$

which leads to the standard relation for the broadening function,

$$\underline{\underline{\Gamma}}(E) = i\left(\underline{\underline{\Sigma}}^R(E) - \underline{\underline{\Sigma}}^A(E)\right) = -2\,\mathrm{Im}\left(\underline{\underline{\Sigma}}^R(E)\right). \tag{4.46}$$

Since the contact and Büttiker-probe self-energies are diagonal in the chosen basis, the lesser self-energy reduces to

$$\underline{\underline{\Sigma}}^<(E) = -2i \sum_\alpha f_\alpha(E)\, \mathrm{diag}\left\{\mathrm{Im}\left(\underline{\underline{\Sigma}}^R_\alpha(E)\right)\right\} \tag{4.47}$$

where the Fermi-Dirac distribution of reservoir α is defined as

$$f_\alpha(E) = \frac{1}{1 + \exp\left(\frac{E-\mu_\alpha}{k_B T}\right)}. \tag{4.48}$$

Here, α denotes the index of the reservoir (left, right, or Büttiker probes, which are treated as

fictitious phase-breaking reservoirs).

4.4 Dyson equation and Keldysh equation

The retarded Green's function is obtained by solving the Dyson equation (Eq. 4.11), which requires the inversion of the tridiagonal matrix

$$\underline{\underline{G}}^R(E) = \left[\underline{\underline{h}}_D(E) - \underline{\underline{\Sigma}}^R_{tot}(E)\right]^{-1}. \tag{4.49}$$

The advanced Green's function is given by the Hermitian conjugate of the retarded Green's function,

$$\underline{\underline{G}}^A(E) = \left(\underline{\underline{G}}^R(E)\right)^\dagger. \tag{4.50}$$

Finally, the lesser Green's function, from which the physical observables are computed, is evaluated through the Keldysh equation (Eq. 3.49):

$$\underline{\underline{G}}^<(E) = \underline{\underline{G}}^R(E)\underline{\underline{\Sigma}}^<(E)\underline{\underline{G}}^A(E). \tag{4.51}$$

The implementation of the greater Green's function follows the same computational procedure as that adopted for the lesser Green's function, the only difference being the use of the greater self-energy in the Keldysh equation. This quantity is fundamental as it is directly related to hole transport and plays a key role in the evaluation of physical observables.

4.5 Numerical Implementation of the Büttiker probe

The Büttiker probe self-energies in Eq. 4.25 and Eq. 4.30 require knowledge of the diagonal elements of the retarded Green's function. However, the Green's function of the full system itself depends on the scattering self-energies, which are required for a realistic description of electronic transport within the device.

For this reason, a first NEGF algorithm is implemented under the assumption of negligible scattering processes. In this ballistic approximation, only the contact self-energies are retained, while all Büttiker probe contributions are set to zero, providing an initial approximation of the Green's functions.

At this initial stage, no Newton–Raphson cycle is required, as the objective is to obtain an approximate numerical solution for the Green's functions. In particular, neither the Keldysh equation nor the evaluation of physical observables is performed at this stage, since self-consistency with the scattering self-energies has not yet been established.

A *fixed-point iteration* scheme is then employed to achieve self-consistency between the Dyson equation and the Büttiker probe self-energies. This procedure ensures consistency between the Green's functions and the total self-energies, including both contact and scattering contributions. The resulting Green's functions are then used to construct the Fermi–Dirac distributions and to evaluate the lesser self-energy according to Eq. 3.39.

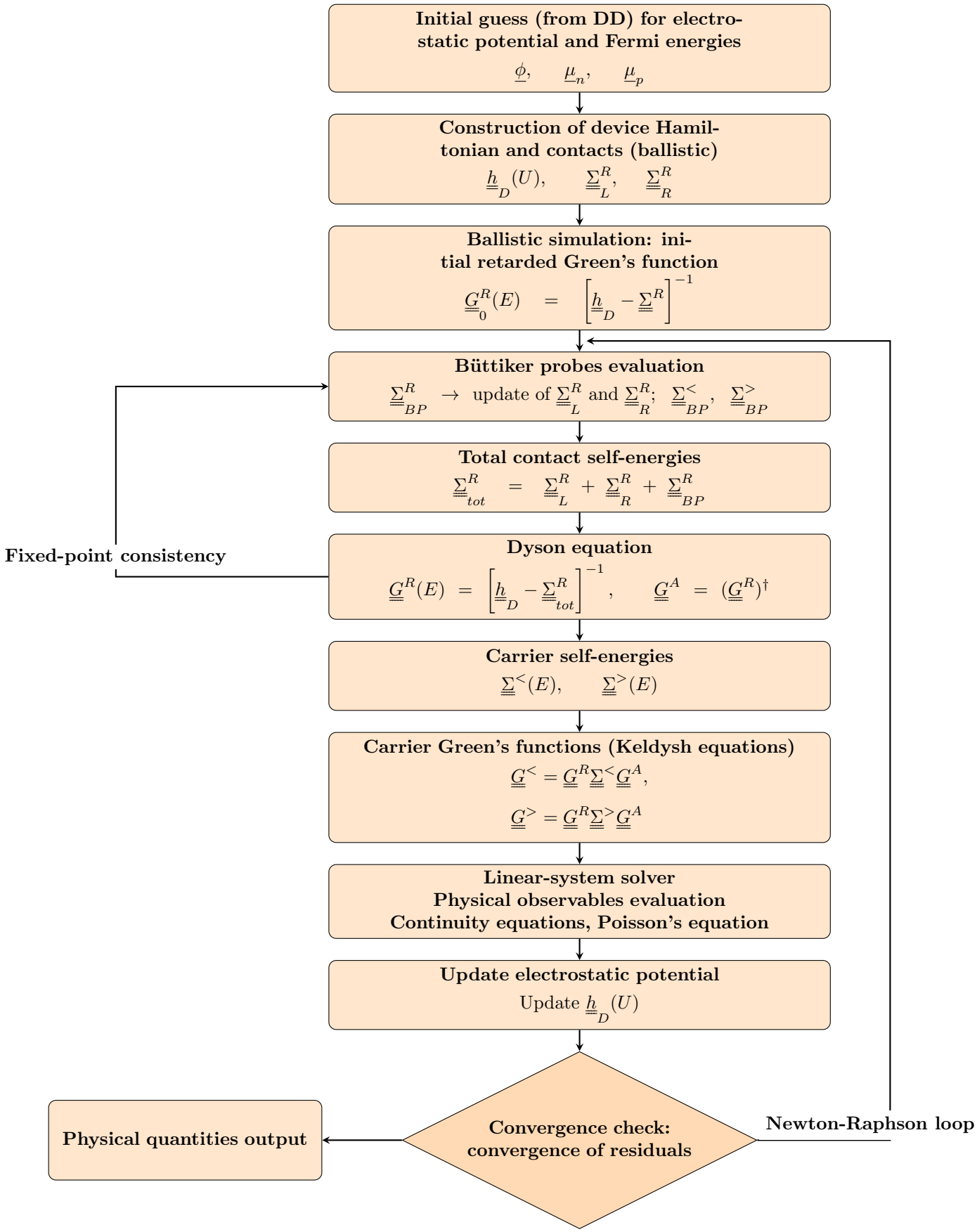


Figure 7: Flowchart of the self-consistent NEGF-Büttiker probe algorithm. Starting from the DD electrostatic potential and carrier densities, the device Hamiltonian and contact self-energies are constructed and the ballistic Green's function is evaluated (*inner iterations*). Büttiker probes are then introduced and iteratively updated through a fixed-point consistency loop coupled to the Dyson equation. Carrier Green's functions are obtained via the Keldysh formalism, from which physical observables are computed by enforcing charge continuity and Poisson's equation. A Newton–Raphson outer loop ensures global self-consistency of the electrostatic potential, while convergence is assessed through the residuals of the coupled equations.

4.6 Physical observable evaluation

Electron current continuity residual

The local charge conservation in the NEGF formalism is verified by evaluating the energy-resolved continuity equation residual after self-consistency of the lesser Green's function is achieved.

In matrix notation, the divergence of the spectral current can be expressed through the commutator between the device Hamiltonian and the lesser Green's function, exploiting the Hermitian nature of the Hamiltonian and the anti-Hermitian property of the lesser Green's function:

$$r(E) = \frac{q}{\hbar}\left[HG^{<}(E) - G^{<}(E)H\right]. \tag{4.52}$$

By exploiting the Hermitian property of the device Hamiltonian, $H_D^{\dagger} = H_D$, together with the anti-Hermitian property of the lesser Green's function, $G^{<\dagger}(E) = -G^{<}(E)$, the two terms in the commutator are complex conjugates of each other. Consequently, the commutator can be rewritten in terms of twice the real part of the corresponding matrix product. Taking the diagonal elements, which correspond to the local divergence of the spectral current at each grid point, yields

$$\boxed{\underline{r}_e(E) = \frac{q}{\hbar}\sigma_{spin}\, 2\,\mathrm{Re}\left\{\underline{\underline{h}}_D\,\underline{\underline{G}}^{<}(E)\right\}} \tag{4.53}$$

In Eq. 4.52, two contributions give origin to the electron net flux in each single node: *in-scattering* term coming from the left node, *out-scattering* term for the right node. By imposing the net current flux to be zero, scattering-related quasi-Fermi levels for electrons ($\underline{\mu_n}$) can be determined at each node of the grid. The spin factor is given by $\sigma_{spin} = 2$.

Hole current continuity residual

The evaluation of hole-related observables follows the same procedure adopted for electrons. The only difference is that the greater Green's function, $G^{>}(E)$, replaces the lesser Green's function, since it describes the occupation of hole states. Accordingly, the greater self-energy, $\Sigma^{>}(E)$, is employed to account for the carrier exchange between the device and the contacts, including the Büttiker probes. The hole continuity residual is therefore obtained from the Hamiltonian–greater Green's function product,

$$\boxed{\underline{r}_h(E) = \frac{q}{\hbar}\sigma_{\mathrm{spin}} 2\,\mathrm{Re}\left\{\underline{\underline{h}}_D(E)\,\underline{\underline{G}}^{>}(E)\right\}} \tag{4.54}$$

Following the same considerations about electrons, the flux of in-coming and out-going holes must be conserved in each node, leading to the numerical evaluation of the quasi-Fermi levels for holes ($\underline{\mu_p}$).

Contact current

Eq. 4.53 describes the energy-resolved divergence of the electron current inside the device region. In a finite open system, special care is required at the device boundaries, since the usual hopping-based expression for the current cannot be evaluated at the first and last lattice nodes without explicitly including the coupling to external reservoirs.

At the contacts, the current is determined through the corresponding contact self-energies, which describe carrier injection into and extraction from the semi-infinite reservoirs. The energy-resolved electron current flowing between the device and the left and right contacts is given by

$$J_L(E) = -\frac{q}{\hbar}\sigma_{\text{spin}}\text{Re}\left\{G^R_{1,1}(E)\Sigma^<_L(E) + G^<_{1,1}(E)\Sigma^A_L(E) - \Sigma^<_L(E)G^A_{1,1}(E) - \Sigma^R_L(E)G^<_{1,1}(E)\right\}, \tag{4.55}$$

$$J_R(E) = +\frac{q}{\hbar}\sigma_{\text{spin}}\text{Re}\left\{G^R_{N,N}(E)\Sigma^<_R(E) + G^<_{N,N}(E)\Sigma^A_R(E) - \Sigma^<_R(E)G^A_{N,N}(E) - \Sigma^R_R(E)G^<_{N,N}(E)\right\} \tag{4.56}$$

These expressions represent the net carrier exchange between the device and the external reservoirs. The contact self-energies depend on the electron and hole quasi-Fermi levels at the contacts, denoted by $\mu_{n,L}$, $\mu_{n,R}$ and $\mu_{p,L}$, $\mu_{p,R}$, respectively, which are determined self-consistently within the Büttiker-probe formalism.

The spectral current satisfies the discrete continuity equation

$$\nabla \cdot \mathbf{J}(E) + \sum_\alpha J_\alpha(E) = 0, \tag{4.57}$$

thereby ensuring global charge conservation even in the presence of incoherent scattering mechanisms.

In conventional drift–diffusion simulations, the applied bias is typically imposed through Dirichlet boundary conditions on the quasi-Fermi levels, namely

$$\mu_n(z_N) - \mu_n(z_1) = -qV, \tag{4.58}$$

where V is the externally applied voltage. Such a boundary condition is appropriate because the current is evaluated throughout the entire computational domain using the same drift–diffusion formulation.

In the NEGF formalism, however, the current flowing inside the device is computed from the Green's functions, whereas the current exchanged with the contacts is obtained from the contact self-energies. Imposing the applied bias directly on the contact quasi-Fermi levels may therefore introduce an artificial discontinuity in the electrochemical potential across the device-contact interface and generally does not guarantee current conservation.

Instead, the contact quasi-Fermi levels are determined by enforcing continuity between the contact current and the current flowing inside the device. In the numerical implementation, this requirement is formulated through the boundary residuals

$$\boxed{r_{n,L}(E) = r_e(z_1, E) - J_L(E)} \tag{4.59}$$

$$\boxed{r_{n,R}(E) = r_e(z_N, E) + J_R(E)} \tag{4.60}$$

where $r_e(z, E)$ denotes the divergence of the electron spectral current evaluated at the first and last lattice sites. During the Newton–Raphson iterations, these residuals are driven to zero, ensuring that the current injected by the contacts exactly matches the current transported through the device.

For hole transport, the same structure applies, with the substitution of lesser Green's functions and self-energies by their greater counterparts. The resulting hole currents at the contacts are therefore:

$$J_\alpha^{(h)}(E) \text{ obtained from } \Sigma^< \to \Sigma^>, \quad G^< \to G^>, \tag{4.61}$$

and the corresponding discrete continuity equation becomes:

$$\nabla \cdot \mathbf{J}_h(E) + \sum_\alpha J_{h,\alpha}(E) = 0, \tag{4.62}$$

which guarantees local charge conservation for the hole population as well.

The corresponding hole continuity residuals at the boundaries are defined in analogy with the electron case as:

$$\boxed{r_{p,L}(E) = r_h(z_1, E) - J_L^{(h)}(E)} \tag{4.63}$$

$$\boxed{r_{p,R}(E) = r_h(z_N, E) + J_R^{(h)}(E)} \tag{4.64}$$

where $r_h(z, E)$ denotes the internal divergence of the hole spectral current, obtained from the greater Green's function.

Local density of states (LDOS), electron and hole density

Recalling Eq. 3.57, the local density of states (LDOS) is obtained from the spectral function, defined as:

$$\underline{LDOS}(E) = \sigma_{spin} Re\{i\left(\underline{\underline{G}}^R(E) - \underline{\underline{G}}^A(E)\right)\} \tag{4.65}$$

Following the structure of Eq. 3.53, the spectral electron density is defined as:

$$\underline{\tilde{n}}(E) = -\sigma_{spin} Re\{i\underline{\underline{G}}^<(E)\} \tag{4.66}$$

The total electron density is obtained by energy integration:

$$\underline{n} = n(z) = \frac{1}{2\pi} \sum_i \underline{\tilde{n}}(E_i) \Delta E_i \tag{4.67}$$

where ΔE is the energy spacing between two consecutive points of the energy discretization.

The hole density, recalling Eq. 3.54, is defined from the greater Green's function:

$$\underline{\tilde{p}}(E) = +\sigma_{spin} Re\{i\underline{\underline{G}}^>(E)\} \tag{4.68}$$

$$\underline{p} = p(z) = \frac{1}{2\pi} \sum_i \underline{\tilde{p}}(E_i) \Delta E_i \tag{4.69}$$

4.7 Jacobian evaluation

The algorithm introduced so far forms the basis for the NEGF evaluation of physical observables through the Green's function matrices. Referring to Eq. 4.2, we have derived the expressions for

the numerical solution of the unknowns of the system; however, the mutual dependence among the different quantities has not yet been explicitly addressed.

As anticipated, a Newton–Raphson scheme is therefore required to compute consistent updates of all unknowns, allowing the system to be solved under realistic operating conditions. The key relation is given in Eq. 4.3, which shows that the correction of the solution vector is obtained by inverting the Jacobian matrix applied to the residual vector.

The Jacobian encodes the full set of interdependencies among all physical quantities describing the system, while the residual vector collects the deviations from the NEGF equations. Together, they define a nonlinear system of equations that must be solved self-consistently.

The structure of the Jacobian matrix is reported below:

$$\underline{\underline{J}} = \begin{pmatrix} \frac{\partial \underline{r}_p}{\partial \underline{U}} & \frac{\partial \underline{r}_p}{\partial \underline{\mu}_n} & \frac{\partial \underline{r}_p}{\partial \underline{\mu}_p} & \frac{\partial r_p}{\partial V} & \frac{\partial r_p}{\partial J_L} & \frac{\partial r_p}{\partial J_R} & \frac{\partial r_p}{\partial \mu_{nl}} & \frac{\partial r_p}{\partial \mu_{nr}} & \frac{\partial r_p}{\partial \mu_{pl}} & \frac{\partial r_p}{\partial \mu_{pr}} \\ \frac{\partial \underline{r}_e}{\partial \underline{U}} & \frac{\partial \underline{r}_e}{\partial \underline{\mu}_n} & \frac{\partial \underline{r}_e}{\partial \underline{\mu}_p} & \frac{\partial r_e}{\partial V} & \frac{\partial r_e}{\partial J_L} & \frac{\partial r_e}{\partial J_R} & \frac{\partial r_e}{\partial \mu_{nl}} & \frac{\partial r_e}{\partial \mu_{nr}} & \frac{\partial r_e}{\partial \mu_{pl}} & \frac{\partial r_e}{\partial \mu_{pr}} \\ \frac{\partial \underline{r}_h}{\partial \underline{U}} & \frac{\partial \underline{r}_h}{\partial \underline{\mu}_n} & \frac{\partial \underline{r}_h}{\partial \underline{\mu}_p} & \frac{\partial r_h}{\partial V} & \frac{\partial r_h}{\partial J_L} & \frac{\partial r_h}{\partial J_R} & \frac{\partial r_h}{\partial \mu_{nl}} & \frac{\partial r_h}{\partial \mu_{nr}} & \frac{\partial r_h}{\partial \mu_{pl}} & \frac{\partial r_h}{\partial \mu_{pr}} \\ \frac{\partial r_v}{\partial \underline{U}} & \frac{\partial r_v}{\partial \underline{\mu}_n} & \frac{\partial r_v}{\partial \underline{\mu}_p} & \frac{\partial r_v}{\partial V} & \frac{\partial r_v}{\partial J_L} & \frac{\partial r_v}{\partial J_R} & \frac{\partial r_v}{\partial \mu_{nl}} & \frac{\partial r_v}{\partial \mu_{nr}} & \frac{\partial r_v}{\partial \mu_{pl}} & \frac{\partial r_v}{\partial \mu_{pr}} \\ \frac{\partial r_{jl}}{\partial \underline{U}} & \frac{\partial r_{jl}}{\partial \underline{\mu}_n} & \frac{\partial r_{jl}}{\partial \underline{\mu}_p} & \frac{\partial r_{jl}}{\partial V} & \frac{\partial r_{jl}}{\partial J_L} & \frac{\partial r_{jl}}{\partial J_R} & \frac{\partial r_{jl}}{\partial \mu_{nl}} & \frac{\partial r_{jl}}{\partial \mu_{nr}} & \frac{\partial r_{jl}}{\partial \mu_{pl}} & \frac{\partial r_{jl}}{\partial \mu_{pr}} \\ \frac{\partial r_{jr}}{\partial \underline{U}} & \frac{\partial r_{jr}}{\partial \underline{\mu}_n} & \frac{\partial r_{jr}}{\partial \underline{\mu}_p} & \frac{\partial r_{jr}}{\partial V} & \frac{\partial r_{jr}}{\partial J_L} & \frac{\partial r_{jr}}{\partial J_R} & \frac{\partial r_{jr}}{\partial \mu_{nl}} & \frac{\partial r_{jr}}{\partial \mu_{nr}} & \frac{\partial r_{jr}}{\partial \mu_{pl}} & \frac{\partial r_{jr}}{\partial \mu_{pr}} \\ \frac{\partial r_{nl}}{\partial \underline{U}} & \frac{\partial r_{nl}}{\partial \underline{\mu}_n} & \frac{\partial r_{nl}}{\partial \underline{\mu}_p} & \frac{\partial r_{nl}}{\partial V} & \frac{\partial r_{nl}}{\partial J_L} & \frac{\partial r_{nl}}{\partial J_R} & \frac{\partial r_{nl}}{\partial \mu_{nl}} & \frac{\partial r_{nl}}{\partial \mu_{nr}} & \frac{\partial r_{nl}}{\partial \mu_{pl}} & \frac{\partial r_{nl}}{\partial \mu_{pr}} \\ \frac{\partial r_{nr}}{\partial \underline{U}} & \frac{\partial r_{nr}}{\partial \underline{\mu}_n} & \frac{\partial r_{nr}}{\partial \underline{\mu}_p} & \frac{\partial r_{nr}}{\partial V} & \frac{\partial r_{nr}}{\partial J_L} & \frac{\partial r_{nr}}{\partial J_R} & \frac{\partial r_{nr}}{\partial \mu_{nl}} & \frac{\partial r_{nr}}{\partial \mu_{nr}} & \frac{\partial r_{nr}}{\partial \mu_{pl}} & \frac{\partial r_{nr}}{\partial \mu_{pr}} \\ \frac{\partial r_{pl}}{\partial \underline{U}} & \frac{\partial r_{pl}}{\partial \underline{\mu}_n} & \frac{\partial r_{pl}}{\partial \underline{\mu}_p} & \frac{\partial r_{pl}}{\partial V} & \frac{\partial r_{pl}}{\partial J_L} & \frac{\partial r_{pl}}{\partial J_R} & \frac{\partial r_{pl}}{\partial \mu_{nl}} & \frac{\partial r_{pl}}{\partial \mu_{nr}} & \frac{\partial r_{pl}}{\partial \mu_{pl}} & \frac{\partial r_{pl}}{\partial \mu_{pr}} \\ \frac{\partial r_{pr}}{\partial \underline{U}} & \frac{\partial r_{pr}}{\partial \underline{\mu}_n} & \frac{\partial r_{pr}}{\partial \underline{\mu}_p} & \frac{\partial r_{pr}}{\partial V} & \frac{\partial r_{pr}}{\partial J_L} & \frac{\partial r_{pr}}{\partial J_R} & \frac{\partial r_{pr}}{\partial \mu_{nl}} & \frac{\partial r_{pr}}{\partial \mu_{nr}} & \frac{\partial r_{pr}}{\partial \mu_{pl}} & \frac{\partial r_{pr}}{\partial \mu_{pr}} \end{pmatrix} \tag{4.70}$$

Based on the structure of the residuals, the Jacobian matrix can be reduced by discarding derivatives with respect to independent variables and rewriting the remaining terms using a compact block-Jacobian notation:

$$\underline{\underline{J}} = \begin{pmatrix} \underline{\underline{J}}^{(p,U)} & \underline{\underline{J}}^{(p,\mu_n)} & \underline{\underline{J}}^{(p,\mu_p)} & \underline{\underline{J}}^{(p,V)} & 0 & 0 & \underline{\underline{J}}^{(p,\mu_{nl})} & \underline{\underline{J}}^{(p,\mu_{nr})} & \underline{\underline{J}}^{(p,\mu_{pl})} & \underline{\underline{J}}^{(p,\mu_{pr})} \\ \underline{\underline{J}}^{(e,U)} & \underline{\underline{J}}^{(e,\mu_n)} & \underline{\underline{J}}^{(e,\mu_p)} & \underline{\underline{J}}^{(e,V)} & 0 & 0 & \underline{\underline{J}}^{(e,\mu_{nl})} & \underline{\underline{J}}^{(e,\mu_{nr})} & 0 & 0 \\ \underline{\underline{J}}^{(h,U)} & \underline{\underline{J}}^{(h,\mu_n)} & \underline{\underline{J}}^{(h,\mu_p)} & \underline{\underline{J}}^{(h,V)} & 0 & 0 & 0 & 0 & \underline{\underline{J}}^{(h,\mu_{pl})} & \underline{\underline{J}}^{(h,\mu_{pr})} \\ 0 & 0 & 0 & J^{(v,V)} & J^{(v,J_L)} & J^{(v,J_R)} & 0 & 0 & 0 & 0 \\ \underline{J}^{(j_L,U)} & \underline{J}^{(j_L,\mu_n)} & \underline{J}^{(j_L,\mu_p)} & 0 & J^{(j_L,J_L)} & 0 & J^{(j_L,\mu_{nl})} & J^{(j_L,\mu_{nr})} & J^{(j_L,\mu_{pl})} & J^{(j_L,\mu_{pr})} \\ \underline{J}^{(j_R,U)} & \underline{J}^{(j_R,\mu_n)} & \underline{J}^{(j_R,\mu_p)} & 0 & 0 & J^{(j_R,J_R)} & J^{(j_R,\mu_{nl})} & J^{(j_R,\mu_{nr})} & J^{(j_R,\mu_{pl})} & J^{(j_R,\mu_{pr})} \\ \underline{J}^{(n_L,U)} & \underline{J}^{(n_L,\mu_n)} & \underline{J}^{(n_L,\mu_p)} & 0 & 0 & 0 & J^{(n_L,\mu_{nl})} & J^{(n_L,\mu_{nr})} & 0 & 0 \\ \underline{J}^{(n_R,U)} & \underline{J}^{(n_R,\mu_n)} & \underline{J}^{(n_R,\mu_p)} & 0 & 0 & 0 & J^{(n_R,\mu_{nl})} & J^{(n_R,\mu_{nr})} & 0 & 0 \\ \underline{J}^{(p_L,U)} & \underline{J}^{(p_L,\mu_n)} & \underline{J}^{(p_L,\mu_p)} & 0 & 0 & 0 & 0 & 0 & J^{(p_L,\mu_{pl})} & J^{(p_L,\mu_{pr})} \\ \underline{J}^{(p_R,U)} & \underline{J}^{(p_R,\mu_n)} & \underline{J}^{(p_R,\mu_p)} & 0 & 0 & 0 & 0 & 0 & J^{(p_R,\mu_{pl})} & J^{(p_R,\mu_{pr})} \end{pmatrix} \tag{4.71}$$

where we introduce the following compact notation for the node-to-node derivatives:

$$\underline{\underline{J}}^{(p,U)} = \begin{pmatrix} \frac{\partial r_{p,1}}{\partial U_1} & \frac{\partial r_{p,1}}{\partial U_2} & \cdots & \frac{\partial r_{p,1}}{\partial U_n} \\ \frac{\partial r_{p,2}}{\partial U_1} & \frac{\partial r_{p,2}}{\partial U_2} & \cdots & \frac{\partial r_{p,2}}{\partial U_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{p,n}}{\partial U_1} & \frac{\partial r_{p,n}}{\partial U_2} & \cdots & \frac{\partial r_{p,n}}{\partial U_n} \end{pmatrix} \tag{4.72}$$

Many Jacobians are scalar quantities, while others involve $n \times n$ matrices associated with the n nodes of the one-dimensional spatial grid. In the following, we focus on the Jacobians related to the Poisson equation and the current-divergence residuals.

The partial derivatives involving the Green's function, in particular its diagonal elements, are generally non-trivial to evaluate. Therefore, the first step consists in computing the derivative of the retarded Green's function with respect to a generic unknown p, by exploiting the identity for the derivative of a matrix inverse.

Generally, the derivative of the inverse of a matrix is given by the key identity:

$$\frac{\partial A^{-1}}{\partial x} = -A^{-1}\frac{\partial A}{\partial x}A^{-1}. \tag{4.73}$$

Starting from the definition

$$G^R = \left(EI - H - \Sigma^R\right)^{-1}, \tag{4.74}$$

its derivative with respect to a generic quantity p is obtained using the identity for matrix inverses:

$$\frac{\partial \underline{\underline{G}}^R}{\partial p} = -\underline{\underline{G}}^R \frac{\partial\left(\underline{\underline{h}}_D - \underline{\underline{\Sigma}}^R\right)}{\partial p}\underline{\underline{G}}^R. \tag{4.75}$$

Electrostatic potential dependence of the residuals

The evaluation of the Jacobian matrix requires determining the dependence of each residual on the unknown variables. Since the scattering self-energies explicitly depend on the retarded Green's function, the first step consists in evaluating the derivatives of the optical and acoustic Büttiker probe self-energies with respect to the electrostatic potential $\underline{U}$. Their dependence is entirely mediated by the derivative of the diagonal elements of the retarded Green's function. In the following, only the **electron contribution** is considered for simplicity, since the same derivation applies to holes after the corresponding substitutions.

$$\begin{aligned} \frac{\partial \underline{\underline{\Sigma}}^R_{opt}(E)}{\partial \underline{U}} = C_{opt}\Bigg[& \left(n_{ph}E_{t,\max} + \mathcal{F}_0(E+E_{ph},\mu_n)\right) i\,\mathrm{Im}\left\{ \left.\frac{\partial\,\mathrm{diag}\{\underline{\underline{G}}^R(E)\}}{\partial \underline{U}}\right|_{E+E_{ph}} \right\} \\ & + \left((n_{ph}+1)E_{t,\max} + \mathcal{F}_0(E-E_{ph},\mu_n)\right) i\,\mathrm{Im}\left\{ \left.\frac{\partial\,\mathrm{diag}\{\underline{\underline{G}}^R(E)\}}{\partial \underline{U}}\right|_{E-E_{ph}} \right\} \Bigg] \end{aligned} \tag{4.76}$$

while the derivative of the acoustic self-energy becomes

$$\frac{\partial \underline{\underline{\Sigma}}^R_{aco}(E)}{\partial \underline{U}} = C_{aco}N^{2D}\, i\,\mathrm{Im}\left\{ \frac{\partial\,\mathrm{diag}\{\underline{\underline{G}}^R(E)\}}{\partial \underline{U}} \right\} \tag{4.77}$$

In the implemented algorithm, the first Newton–Raphson iteration employs the derivative of the ballistic retarded Green's function, i.e., neglecting the dependence of the Büttiker probe self-energies on the Green's function itself. This approximation provides an initial estimate of the scattering self-energy derivatives, which is subsequently refined during the self-consistent NEGF iterations.

The derivative of the total Büttiker probe retarded self-energy is therefore obtained as the sum of the optical and acoustic contributions,

$$\frac{\partial \underline{\underline{\Sigma}}^R_{BP}(E)}{\partial \underline{U}} = \frac{\partial \underline{\underline{\Sigma}}^R_{opt}(E)}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^R_{aco}(E)}{\partial \underline{U}} \tag{4.78}$$

The contact self-energies depend on the electrostatic potential through the effective on-site Hamiltonian entering the contact surface Green's function. Since this effective Hamiltonian includes the Büttiker probe self-energy, an additional implicit dependence on $\underline{U}$ is introduced. Focusing on the left contact,

$$\frac{\partial \Sigma^R_L}{\partial \underline{U}} = \frac{\partial \Sigma^R_L}{\partial \widetilde{h}_0} \frac{\partial \widetilde{h}_0}{\partial \underline{U}} \tag{4.79}$$

where, recalling Eq. 4.41,

$$\frac{\partial \Sigma^R_L}{\partial \widetilde{h}_0} = \frac{1}{2} + \frac{\widetilde{h}_0}{2\sqrt{\widetilde{h}_0^2 - 4h_1^2}}, \tag{4.80}$$

and

$$\frac{\partial \widetilde{h}_0}{\partial \underline{U}} = \frac{\partial \left(h_0 - \Sigma^R_{BP,L}\right)}{\partial \underline{U}} = -\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} - \frac{\partial \Sigma^R_{BP,L}}{\partial \underline{U}} \tag{4.81}$$

The first term accounts for the explicit dependence of the contact on-site energy on the electrostatic potential, whereas the second originates from the implicit dependence through the Büttiker probe self-energy. The same derivation applies to the right contact. Collecting both contributions yields the derivative of the total contact self-energy,

$$\frac{\partial \underline{\underline{\Sigma}}^R_C}{\partial \underline{U}} = \frac{\partial \underline{\underline{\Sigma}}^R_L}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^R_R}{\partial \underline{U}} \tag{4.82}$$

where each contribution is non-zero only at the corresponding contact node.

For the lesser self-energy, the total derivative is given by the sum of the contact and Büttiker probe contributions,

$$\frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{U}} = \frac{\partial \underline{\underline{\Sigma}}^<_C}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^<_{BP}}{\partial \underline{U}} \tag{4.83}$$

The contact contribution is itself the sum of the left and right reservoirs,

$$\frac{\partial \underline{\underline{\Sigma}}^<_C}{\partial \underline{U}} = \frac{\partial \underline{\underline{\Sigma}}^<_L}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^<_R}{\partial \underline{U}} \tag{4.84}$$

Both the contact and the Büttiker probe lesser self-energies depend on the corresponding retarded self-energies through the broadening matrices,

$$\underline{\underline{\Gamma}} = i\left(\underline{\underline{\Sigma}}^{R} - \underline{\underline{\Sigma}}^{A}\right) = -2\,\mathrm{Im}\left\{\underline{\underline{\Sigma}}^{R}\right\} \tag{4.85}$$

Consequently,

$$\frac{\partial \underline{\underline{\Gamma}}_{L}}{\partial \underline{U}} = -2\,\mathrm{Im}\left\{\frac{\partial \underline{\underline{\Sigma}}^{R}_{L}}{\partial \underline{U}}\right\} \tag{4.86}$$

$$\frac{\partial \underline{\underline{\Gamma}}_{R}}{\partial \underline{U}} = -2\,\mathrm{Im}\left\{\frac{\partial \underline{\underline{\Sigma}}^{R}_{R}}{\partial \underline{U}}\right\} \tag{4.87}$$

and

$$\frac{\partial \underline{\underline{\Gamma}}_{BP}}{\partial \underline{U}} = -2\,\mathrm{Im}\left\{\frac{\partial \underline{\underline{\Sigma}}^{R}_{BP}}{\partial \underline{U}}\right\} \tag{4.88}$$

Restricting the discussion to electrons, since the hole counterpart follows the same derivation,

$$\frac{\partial \underline{\underline{\Sigma}}^{<}_{L}}{\partial \underline{U}} = i\,\mathcal{F}_0(\mu_{nl})\frac{\partial \underline{\underline{\Gamma}}_{L}}{\partial \underline{U}} \tag{4.89}$$

$$\frac{\partial \underline{\underline{\Sigma}}^{<}_{R}}{\partial \underline{U}} = i\,\mathcal{F}_0(\mu_{nr})\frac{\partial \underline{\underline{\Gamma}}_{R}}{\partial \underline{U}} \tag{4.90}$$

where the contact quasi-Fermi levels are treated as independent unknowns within the Newton–Raphson scheme. Similarly,

$$\frac{\partial \underline{\underline{\Sigma}}^{<}_{BP}}{\partial \underline{U}} = i\,\mathcal{F}_0(\mu_{n})\frac{\partial \underline{\underline{\Gamma}}_{BP}}{\partial \underline{U}} \tag{4.91}$$

The Fermi–Dirac integral of order zero, $\mathcal{F}_0$, is evaluated numerically at each Newton iteration.

The derivative of the retarded Green's function can be evaluated using the expressions of the retarded self-energies and the identity for the derivative of a matrix inverse (Eq. 4.75). This yields

$$\frac{\partial \underline{\underline{G}}^{R}}{\partial \underline{U}} = \underline{\underline{G}}^{R}\left[q\mathbb{I}_{diag} + \frac{\partial \underline{\underline{\Sigma}}^{R}_{C+BP}}{\partial \underline{U}}\right]\underline{\underline{G}}^{R} \tag{4.92}$$

We now derive the expression for the derivative of the lesser Green's function with respect to the electrostatic potential. Starting from the Keldysh relation

$$\underline{\underline{G}}^{<} = \underline{\underline{G}}^{R}\underline{\underline{\Sigma}}^{<}\underline{\underline{G}}^{A}, \tag{4.93}$$

its derivative with respect to the general unknown $\underline{U}$ is obtained by applying the product rule:

$$\frac{\partial \underline{\underline{G}}^{<}}{\partial \underline{U}} = \frac{\partial \underline{\underline{G}}^{R}}{\partial \underline{U}}\underline{\underline{\Sigma}}^{<}\underline{\underline{G}}^{A} + \underline{\underline{G}}^{R}\frac{\partial \underline{\underline{\Sigma}}^{<}}{\partial \underline{U}}\underline{\underline{G}}^{A} + \underline{\underline{G}}^{R}\underline{\underline{\Sigma}}^{<}\frac{\partial \underline{\underline{G}}^{A}}{\partial \underline{U}}. \tag{4.94}$$

Using the identities

$$\frac{\partial \underline{\underline{G}}^{R}}{\partial \underline{U}} = -\underline{\underline{G}}^{R}\left(\frac{\partial \underline{\underline{H}}}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^{R}}{\partial \underline{U}}\right)\underline{\underline{G}}^{R}, \tag{4.95}$$

and

$$\frac{\partial \underline{\underline{G}}^A}{\partial \underline{U}} = -\underline{\underline{G}}^A \left(\frac{\partial \underline{\underline{H}}}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^A}{\partial \underline{U}} \right) \underline{\underline{G}}^A, \tag{4.96}$$

we obtain:

$$\begin{aligned} \frac{\partial \underline{\underline{G}}^<}{\partial \underline{U}} = &- \underline{\underline{G}}^R \left(\frac{\partial \underline{\underline{H}}}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{U}} \right) \underline{\underline{G}}^R \underline{\underline{\Sigma}}^< \underline{\underline{G}}^A \\ &+ \underline{\underline{G}}^R \frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{U}} \underline{\underline{G}}^A \\ &- \underline{\underline{G}}^R \underline{\underline{\Sigma}}^< \underline{\underline{G}}^A \left(\frac{\partial \underline{\underline{H}}}{\partial \underline{U}} + \frac{\partial \underline{\underline{\Sigma}}^A}{\partial \underline{U}} \right) \underline{\underline{G}}^A. \end{aligned} \tag{4.97}$$

Taking the diagonal part yields the final working expression:

$$\frac{\partial}{\partial \underline{U}} \mathrm{diag}\{\underline{\underline{G}}^<\} = \mathrm{diag} \left\{ \frac{\partial \underline{\underline{G}}^<}{\partial \underline{U}} \right\}. \tag{4.98}$$

The evaluation of the Jacobian matrix requires computing the sensitivity of each residual equation with respect to the electrostatic potential. In particular, the electron and hole current residuals depend explicitly on the Green's functions and on the device Hamiltonian, both of which are functions of the electrostatic potential through the self-consistent electrostatic coupling.

Starting from the definitions of the current residuals

$$\underline{r}_e(E) = \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re} \left\{ \underline{\underline{h}}_D(E) \underline{\underline{G}}^<(E) \right\}, \qquad \underline{r}_h(E) = \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re} \left\{ \underline{\underline{h}}_D(E) \underline{\underline{G}}^>(E) \right\}, \tag{4.99}$$

their derivatives with respect to the electrostatic potential are obtained by applying the product rule to the matrix-valued quantities inside the real part operator.

Since both $\underline{\underline{h}}_D$ and $\underline{\underline{G}}^{</>}$ depend implicitly on $\underline{U}$, the derivative of each residual naturally splits into two contributions: one arising from the explicit dependence of the device Hamiltonian, and one originating from the Green's functions, which encode the full non-equilibrium response of the system.

The real-part operator is linear, therefore the differentiation can be performed directly on the complex quantity inside the brackets. This leads to

$$\frac{\partial}{\partial \underline{U}} \mathrm{Re}\{X\} = \mathrm{Re} \left\{ \frac{\partial X}{\partial \underline{U}} \right\}, \tag{4.100}$$

which allows the derivative to be distributed over the matrix product.

Applying these rules yields the final expressions:

$$\begin{aligned} \frac{\partial \underline{r}_e(E)}{\partial \underline{U}} &= \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re} \left\{ \frac{\partial \underline{\underline{h}}_D(E)}{\partial \underline{U}} \underline{\underline{G}}^<(E) + \underline{\underline{h}}_D(E) \frac{\partial \underline{\underline{G}}^<(E)}{\partial \underline{U}} \right\}, \\ \frac{\partial \underline{r}_h(E)}{\partial \underline{U}} &= \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re} \left\{ \frac{\partial \underline{\underline{h}}_D(E)}{\partial \underline{U}} \underline{\underline{G}}^>(E) + \underline{\underline{h}}_D(E) \frac{\partial \underline{\underline{G}}^>(E)}{\partial \underline{U}} \right\}. \end{aligned} \tag{4.101}$$

These expressions identify two distinct contributions to the Jacobian matrix.

The first contribution (i) originates from the explicit dependence of the device Hamiltonian on the electrostatic potential. This term describes the direct electrostatic response of the system, i.e., the modification of the local energy landscape induced by a variation of $\underline{U}$.

The second contribution (ii) arises from the implicit, non-linear dependence of the Green's functions through the self-energy structure. This term captures the full quantum-transport feedback, since the Green's functions depend on the retarded and lesser self-energies, which are themselves functionals of the electrostatic potential.

The second term is therefore significantly more computationally demanding, as it requires the evaluation of the derivatives of both the retarded and lesser Green's functions with respect to $\underline{U}$, fully embedding the NEGF self-consistency loop.

Fermi energy dependence of the residuals

The following derivation applies equally to electrons and holes, i.e., to the quasi-Fermi energies μ_n and μ_p, after the appropriate substitutions of carrier statistics and contact indices. Unlike the electrostatic potential, the quasi-Fermi energies affect the NEGF equations through the carrier occupation factors entering the Büttiker probe self-energies. This dependence propagates to the contact self-energies through the modified contact Hamiltonian, so that both the retarded and the lesser self-energies contribute to the Jacobian.

For the optical Büttiker probe self-energy, the derivative of the retarded component is

$$\begin{aligned}\frac{\partial \underline{\underline{\Sigma}}^R_{opt}(E)}{\partial \mu} = C_{opt}\Bigg[&\frac{\partial \mathcal{F}_0(E+E_{ph},\mu)}{\partial \mu}\, i\,\mathrm{Im}\left\{\mathrm{diag}\,\underline{\underline{G}}^R(E+E_{ph})\right\} \\ &- \frac{\partial \mathcal{F}_0(E-E_{ph},\mu)}{\partial \mu}\, i\,\mathrm{Im}\left\{\mathrm{diag}\,\underline{\underline{G}}^R(E-E_{ph})\right\}\Bigg].\end{aligned} \tag{4.102}$$

The acoustic contribution does not contain Fermi occupation factors and therefore

$$\frac{\partial \underline{\underline{\Sigma}}^R_{aco}}{\partial \mu} = 0. \tag{4.103}$$

Hence,

$$\frac{\partial \underline{\underline{\Sigma}}^R_{BP}}{\partial \mu} = \frac{\partial \underline{\underline{\Sigma}}^R_{opt}}{\partial \mu}. \tag{4.104}$$

Since the contact self-energies are evaluated from the modified on-site Hamiltonian

$$\widetilde{h}_0 = h_0 - \Sigma^R_{BP},$$

they also inherit an implicit dependence on the quasi-Fermi energies. For the left contact,

$$\frac{\partial \Sigma^R_L}{\partial \mu} = \frac{\partial \Sigma^R_L}{\partial \widetilde{h}_0}\frac{\partial \widetilde{h}_0}{\partial \mu}, \tag{4.105}$$

with

$$\frac{\partial \widetilde{h}_0}{\partial \mu} = -\frac{\partial \Sigma^R_{BP,L}}{\partial \mu}. \tag{4.106}$$

The same expression holds for the right contact.

The contact lesser self-energies satisfy the Leibniz rule,

$$\frac{\partial \underline{\underline{\Sigma}}_L^<}{\partial \mu_{nl}} = i\left(\frac{\partial \mathcal{F}_0(\mu_{nl})}{\partial \mu_{nl}} \underline{\underline{\Gamma}}_L + \mathcal{F}_0(\mu_{nl}) \frac{\partial \underline{\underline{\Gamma}}_L}{\partial \mu_{nl}}\right) \tag{4.107}$$

$$\frac{\partial \underline{\underline{\Sigma}}_R^<}{\partial \mu_{nr}} = i\left(\frac{\partial \mathcal{F}_0(\mu_{nr})}{\partial \mu_{nr}} \underline{\underline{\Gamma}}_R + \mathcal{F}_0(\mu_{nr}) \frac{\partial \underline{\underline{\Gamma}}_R}{\partial \mu_{nr}}\right) \tag{4.108}$$

and the Büttiker probe contribution becomes

$$\frac{\partial \underline{\underline{\Sigma}}_{BP}^<}{\partial \mu} = i\left(\frac{\partial \mathcal{F}_0(\mu)}{\partial \mu} \underline{\underline{\Gamma}}_{BP} + \mathcal{F}_0(\mu) \frac{\partial \underline{\underline{\Gamma}}_{BP}}{\partial \mu}\right) \tag{4.109}$$

Remember the broadening factor Γ has no explicit dependence on the Fermi energies, so the partial derivative vanishes:

$$\frac{\partial \underline{\underline{\Gamma}}_L}{\partial \mu_{nl}} = 0, \ \frac{\partial \underline{\underline{\Gamma}}_R}{\partial \mu_{nr}} = 0, \ \frac{\partial \underline{\underline{\Gamma}}_R}{\partial \mu_{nr}} = 0 \tag{4.110}$$

The total lesser self-energy derivative is then

$$\frac{\partial \underline{\underline{\Sigma}}^<}{\partial \mu} = \frac{\partial \underline{\underline{\Sigma}}_C^<}{\partial \mu} + \frac{\partial \underline{\underline{\Sigma}}_{BP}^<}{\partial \mu}, \tag{4.111}$$

where

$$\frac{\partial \underline{\underline{\Sigma}}_C^<}{\partial \mu} = \frac{\partial \underline{\underline{\Sigma}}_L^<}{\partial \mu} + \frac{\partial \underline{\underline{\Sigma}}_R^<}{\partial \mu}. \tag{4.112}$$

The derivative of the diagonal of the lesser Green's function is evaluated using the same optimized form adopted for the electrostatic Jacobian. Applying the product rule to

$$\underline{\underline{G}}^< = \underline{\underline{G}}^R \underline{\underline{\Sigma}}^< \underline{\underline{G}}^A,$$

gives

$$\begin{aligned}\frac{\partial}{\partial \underline{\mu}} \mathrm{diag}\{\underline{\underline{G}}^<\} = \mathrm{diag}\left\{\underline{\underline{G}}^R \frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{\mu}} \underline{\underline{G}}^< + \underline{\underline{G}}^< \frac{\partial (\underline{\underline{\Sigma}}^R)^\dagger}{\partial \underline{\mu}} \underline{\underline{G}}^A + \underline{\underline{G}}^R \frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{\mu}} \underline{\underline{G}}^A\right\} \\ + 2i\,\mathrm{Im}\left\{\mathrm{diag}\left(\underline{\underline{G}}^R \frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{\mu}} \underline{\underline{G}}^<\right) + \mathrm{diag}\left(\underline{\underline{G}}^R \frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{\mu}} \underline{\underline{G}}^A\right)\right\},\end{aligned} \tag{4.113}$$

which is exactly the expression implemented in the numerical algorithm.

Finally, the Jacobian contributions associated with the electron and hole current residuals are

$$\frac{\partial \underline{r}_e}{\partial \mu} = \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re}\left\{\underline{\underline{h}}_D \frac{\partial \underline{\underline{G}}^<}{\partial \mu}\right\}, \tag{4.114}$$

$$\frac{\partial \underline{r}_h}{\partial \mu} = \frac{q}{\hbar} \sigma_{\mathrm{spin}} \mathrm{Re}\left\{\underline{\underline{h}}_D \frac{\partial \underline{\underline{G}}^>}{\partial \mu}\right\}. \tag{4.115}$$

Jacobian of the contact current residuals

The evaluation of the Jacobian contributions associated with the contact current residuals requires the computation of their sensitivity with respect to both the electrostatic potential $\underline{U}$ and the quasi-Fermi energies $\underline{\mu}$. The derivation is presented for the left contact; the right contact follows identically by replacing the corresponding boundary indices.

The left contact current is given by

$$J_L(E) = -\frac{q}{\hbar}\sigma_{\text{spin}}\text{Re}\Big\{G^R_{1,1}\Sigma^<_L + G^<_{1,1}\Sigma^A_L - \Sigma^<_L G^A_{1,1} - \Sigma^R_L G^<\Big\} \tag{4.116}$$

Its Jacobian is obtained by applying the product rule to all matrix-valued quantities, taking into account that both Green's functions and self-energies depend on the unknown variables.

For what concerns the derivative with respect to the electrostatic potential, all quantities exhibit an implicit dependence through the Hamiltonian and the retarded self-energies. The derivative reads

$$\begin{aligned}\frac{\partial J_L(E)}{\partial \underline{U}} = -\frac{q}{\hbar}\sigma_{\text{spin}}\text{Re}\Big\{&\frac{\partial G^R_{1,1}}{\partial \underline{U}}\Sigma^<_L + G^R_{1,1}\frac{\partial \Sigma^<_L}{\partial \underline{U}} + \frac{\partial G^<_{1,1}}{\partial \underline{U}}\Sigma^A_L \\ &+ G^<_{1,1}\frac{\partial \Sigma^A_L}{\partial \underline{U}} - \frac{\partial \Sigma^<_L}{\partial \underline{U}}G^A_{1,1} - \Sigma^<_L\frac{\partial G^A_{1,1}}{\partial \underline{U}} - \frac{\partial \Sigma^R_L}{\partial \underline{U}}G^< - \Sigma^R_L\frac{\partial G^<_{1,1}}{\partial \underline{U}}\Big\}\end{aligned} \tag{4.117}$$

All Green's function derivatives are obtained through the NEGF identity

$$\frac{\partial \underline{\underline{G}}}{\partial \underline{U}} = \underline{\underline{G}}\frac{\partial \underline{\underline{\Sigma}}}{\partial \underline{U}}\underline{\underline{G}} \tag{4.118}$$

In the case of the Fermi energy derivatives, the dependence is restricted to the lesser self-energies through the Fermi occupation functions, while all retarded quantities remain independent of $\underline{\mu}$. Therefore,

$$\frac{\partial \Sigma^R_L}{\partial \underline{\mu}} = 0, \qquad \frac{\partial G^{R/A}}{\partial \underline{\mu}} = 0. \tag{4.119}$$

The only non-vanishing contributions arise from

$$\frac{\partial \Sigma^<_L}{\partial \mu_{nl}} = i\,\frac{\partial \mathcal{F}_0(\mu_{nl})}{\partial \mu_{nl}}\,\Gamma_L. \tag{4.120}$$

Consequently, the Jacobian reduces to

$$\frac{\partial J_L(E)}{\partial \mu_{nl}} = -\frac{q}{\hbar}\sigma_{\text{spin}}\text{Re}\Big\{G^R_{1,1}\frac{\partial \Sigma^<_L}{\partial \mu_{nl}} + \frac{\partial G^<_{1,1}}{\partial \mu_{nl}}\Sigma^A_L - \frac{\partial \Sigma^<_L}{\partial \mu_{nl}}G^A_{1,1} - \Sigma^R_L\frac{\partial G^<_{1,1}}{\partial \mu_{nl}}\Big\}. \tag{4.121}$$

An analogous expression holds for the right contact by replacing $1 \to N$ and $\mu_{nl} \to \mu_{nr}$.

4.8 Poisson-NEGF

After assembling the Jacobian matrix and evaluating the physical observables defining the residual vector, the Newton–Raphson algorithm is iterated until convergence of the unknowns is achieved. These external iterations enforce a fully self-consistent solution of the coupled nonlinear system governing quantum transport in the device.

In particular, the electrostatic potential is updated through the Poisson equation

$$\nabla^2\phi = -\frac{q}{\epsilon}\left(p - n + N_D^+ - N_A^-\right), \tag{4.122}$$

where the charge density is obtained from the NEGF formalism through the electron and hole spectral contributions. Unlike semiclassical drift–diffusion approaches, both densities depend explicitly on the electrostatic potential through the device Hamiltonian and the self-energy operators.

In the discretized formulation, the Poisson residual is written directly in terms of the electrostatic potential vector as

$$\underline{r}_P = -\frac{\partial}{\partial \underline{z}}\epsilon(z)\frac{\partial \underline{\phi}}{\partial \underline{z}} - q\left(\underline{n} - \underline{p} + \underline{N}_D^+ - \underline{N}_A^-\right), \tag{4.123}$$

where the potential vector $\underline{\phi}$ represents the electrostatic potential in the device grid, while the charge contributions are computed from the NEGF formalism.

The corresponding Jacobian contribution is then given by

$$\frac{\partial \underline{r}_P}{\partial \underline{U}} = -\frac{\partial}{\partial \underline{U}}\frac{\partial}{\partial \underline{z}}\epsilon(z)\frac{\partial \underline{\phi}}{\partial \underline{z}} - \frac{\partial \underline{\rho}}{\partial \underline{U}}, \tag{4.124}$$

with

$$\frac{\partial \underline{\rho}}{\partial \underline{U}} = -q\frac{\partial \underline{n}}{\partial \underline{U}} + q\frac{\partial \underline{p}}{\partial \underline{U}}. \tag{4.125}$$

The derivatives of the electron and hole densities are obtained from the functional dependence of the Green's functions on the electrostatic potential. In matrix form, the variation of the lesser Green's function is expressed as

$$\frac{\partial \underline{\underline{G}}^<}{\partial \underline{U}} = \underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{U}}\underline{\underline{G}}^< + \underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{U}}\underline{\underline{G}}^A + \text{h.c.} \tag{4.126}$$

Accordingly, the electron density derivative is computed from the energy-discretized expression

$$\frac{\partial \underline{n}}{\partial \underline{U}} = -\frac{1}{2\pi}\sum_E \Delta E\ \text{diag}\left\{2\,\text{Im}\left[\underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{U}}\underline{\underline{G}}^<\right] + \underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^<}{\partial \underline{U}}\underline{\underline{G}}^A\right\}, \tag{4.127}$$

while the hole contribution follows analogously as

$$\frac{\partial \underline{p}}{\partial \underline{U}} = \frac{1}{2\pi}\sum_E \Delta E\ \text{diag}\left\{2\,\text{Im}\left[\underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^R}{\partial \underline{U}}\underline{\underline{G}}^>\right] + \underline{\underline{G}}^R\frac{\partial \underline{\underline{\Sigma}}^>}{\partial \underline{U}}\underline{\underline{G}}^A\right\}. \tag{4.128}$$

This formulation provides the full quantum correction to the charge response with respect to the electrostatic potential and closes the Poisson block of the NEGF Jacobian within the Newton–Raphson self-consistent loop.

The residual form defined above is then evaluated at each Newton iteration, providing the mismatch between the current electrostatic potential and the charge distribution obtained from the NEGF solution. The condition for self-consistency is therefore expressed as

$$\underline{r}_P = \underline{0}, \tag{4.129}$$

which enforces that the electrostatic potential exactly reproduces the charge density generated by the quantum carrier distribution.

In the Newton–Raphson scheme, this residual is used together with its Jacobian to update the potential according to

$$\Delta \underline{U} = -\left(\frac{\partial \underline{r}_P}{\partial \underline{U}}\right)^{-1} \underline{r}_P, \tag{4.130}$$

ensuring the self-consistent coupling between Poisson and NEGF equations at convergence.

4.9 Recombination currents

The simulation framework developed so far includes several degrees of complexity, making it suitable for the analysis of nanoscale one-dimensional devices, including semiconductor heterostructures. To further improve the predictive capability of the simulator, carrier generation and recombination mechanisms have to be considered, since they play a fundamental role in optoelectronic devices, where optical generation and carrier recombination processes are intrinsically coupled.

Recombination and generation phenomena enter the carrier continuity equations (Eq. 2.30) through the terms U_n and U_p, representing the net electron and hole generation rates, respectively. In a fully microscopic NEGF description, these quantities can be obtained by introducing additional self-energy terms associated with the corresponding scattering processes. For example, starting from the stationary electron continuity equation:

$$\frac{\partial J_n}{\partial z} = qU_n, \tag{4.131}$$

with an analogous expression for holes, the electron generation rate can be expressed in terms of the scattering self-energies as:

$$\begin{aligned} U_n(z) = -\frac{2}{\hbar}\int \frac{dE}{2\pi}\int dz' \Big[&\Sigma^R(z,z',E)G^<(z',z,E) + \Sigma^<(z,z',E)G^A(z',z,E) \\ &- G^R(z,z',E)\Sigma^<(z',z,E) - G^<(z,z',E)\Sigma^A(z',z,E)\Big]. \end{aligned} \tag{4.132}$$

This approach provides a fully quantum-mechanical description of generation and recombination processes. However, the inclusion of additional self-energy contributions significantly increases the computational complexity, since new self-consistent terms must be evaluated together with the Green's functions.

To avoid a substantial increase in the computational cost, a semiclassical description of recombination mechanisms is adopted in this work. The considered processes are radiative recombination and Shockley–Read–Hall recombination, whose effects are introduced through conventional recombination models.

Although the recombination models are treated semiclassically, the approach can be considered as *quantum corrected*, since the carrier densities entering the recombination expressions are not obtained from a classical drift–diffusion model, but are instead evaluated from the NEGF formalism. Consequently, quantum confinement and tunneling effects are implicitly included in the recombination rates through the quantum-mechanical carrier distributions.

Two different approaches can be adopted for the inclusion of recombination phenomena. The first consists in incorporating the recombination rates directly into the NEGF continuity equations, adding their contribution to the electron and hole residuals and, consequently, to the Jacobian matrix during the self-consistent Newton–Raphson iterations. This approach provides a fully self-

consistent treatment of recombination effects, but requires the evaluation of additional derivatives of the recombination rates with respect to the unknown variables, increasing the computational cost.

The second approach consists in evaluating the recombination rates as a post-processing analysis of the converged NEGF solution. In this case, the carrier distributions obtained from the quantum transport calculation are used to estimate the recombination contribution through semi-classical expressions, including radiative, Shockley–Read–Hall (SRH), and Auger recombination mechanisms. Although this approach does not modify the self-consistent NEGF solution, it allows the estimation of recombination effects while avoiding the additional computational burden associated with their iterative inclusion.

Radiative recombination

Radiative recombination describes the spontaneous emission of photons resulting from band-to-band transitions, where electrons in the conduction band recombine with holes in the valence band. The corresponding recombination rate is expressed as:

$$\boxed{U_n^{rad} = B^{rad}\left(np - n_{eq}p_{eq}\right)} \tag{4.133}$$

where B^{rad} is the material-dependent radiative recombination coefficient. The recombination rate is proportional to the product of the electron density in the conduction band and the hole density in the valence band, since both carriers are required for a band-to-band transition. The term $n_{eq}p_{eq}$ ensures that the net recombination rate vanishes under thermodynamic equilibrium, satisfying the detailed balance condition. The photon-emitted energy is of the order of the band gap energy $\hbar\omega \geq E_g$.

Shockley–Read–Hall recombination

Shockley–Read–Hall (SRH) recombination occurs through localized defect or trap states within the semiconductor band gap. Electrons and holes recombine via these intermediate energy levels rather than through a direct band-to-band transition. The recombination rate depends on the trap energy level, whose equilibrium carrier concentrations are given by

$$n_1 = n_i \exp\left(\frac{E_{trap} - E_{F,i}}{k_B T}\right), \tag{4.134}$$

$$p_1 = n_i \exp\left(-\frac{E_{trap} - E_{F,i}}{k_B T}\right), \tag{4.135}$$

where E_{trap} denotes the trap energy level and $E_{F,i}$ is the intrinsic Fermi level. Defining τ_n and τ_p as the electron and hole lifetimes associated with the trap, the SRH recombination rate is expressed as

$$\boxed{U^{SRH} = \frac{np - n_{eq}p_{eq}}{\tau_p(n + n_1) + \tau_n(p + p_1)}} \tag{4.136}$$

where n_i is the intrinsic carrier concentration. The term $n_{eq}p_{eq}$ guarantees that the net recombination rate vanishes under thermodynamic equilibrium.

Auger recombination

In an Auger process, the energy released by the recombination of an electron–hole pair is not emitted as a photon, but is instead transferred to a third carrier, producing a highly energetic (hot) electron or hole. Consequently, Auger recombination is a non-radiative mechanism that becomes particularly significant at high carrier concentrations. The corresponding recombination rate is

$$\boxed{U^{Aug} = (C_n n + C_p p)\,(np - n_{eq}p_{eq})\,,} \tag{4.137}$$

where C_n and C_p are the electron and hole Auger recombination coefficients, respectively.

5 Frozen Field approach

The trade-off between spatial discretization and computational time motivates the development of strategies aimed at reducing the computational cost while preserving the accuracy of the numerical solution. The main challenge arises from the large size of the Jacobian matrix resulting from the coupling between the Poisson equation and the NEGF transport equations, which must be solved self-consistently to determine the electrostatic potential together with the electron and hole transport throughout the device. For complex structures, such as GaN-based light-emitting diodes, the wide bandgap and the corresponding low carrier densities make the evaluation of the spectral charge particularly demanding from a numerical standpoint. Consequently, the Newton–Raphson iterations may become unstable since Poisson equation is not linear, leading to unrealistic electrostatic potential updates and, ultimately, to the divergence of the simulation.

As discussed throughout this work, the main advantage of the NEGF formalism lies in its rigorous quantum-mechanical description of carrier transport, including tunneling and other coherent phenomena. However, for many devices the electrostatic potential profile is already accurately predicted by the drift–diffusion (DD) model, especially when appropriate quantum corrections to the conduction and valence bands are included, as described in Chapter 2.5. Although the DD approach cannot reproduce quantum transport effects, it often provides a reliable description of the device electrostatics.

Based on this observation, a computationally efficient alternative consists of adopting the electrostatic potential obtained from the DD simulation as a fixed input for the NEGF calculations. In this *frozen-field* approach, the Poisson equation is no longer solved self-consistently with the NEGF transport equations. Instead, the NEGF formalism computes the quantum transport properties using the electrostatic potential profile provided by the DD simulation. This decoupling significantly reduces the computational burden and improves numerical robustness. In the following sections, several device structures are investigated to assess the validity of this approximation and to determine whether it can accurately reproduce the transport characteristics while considerably reducing the simulation cost.

To assess the numerical stability and accuracy of the frozen-field approach, the analysis is further extended to include the evaluation of the radiative, Shockley–Read–Hall (SRH), and Auger recombination mechanisms through the implementation of the conventional ABC recombination model. This additional investigation makes it possible to evaluate whether the frozen-field approximation not only reproduces the carrier transport characteristics but also provides an accurate description of the optical recombination processes governing the device performance.

5.1 Poisson–drift–diffusion and Poisson–Schrödinger–Drift–Diffusion

As a first step, drift–diffusion (DD) simulations are investigated to assess the effects introduced by the quantum corrections obtained through the self-consistent solution of the Schrödinger equation. An InGaAs-based light-emitting diode (InGaAs–LED) is adopted as a benchmark structure, allowing the analysis of carrier transport through the conduction and valence bands while maintaining a clear physical interpretation.

InGaAs/AlGaAs multi-quantum-well (MQW) structures have attracted significant interest for infrared optoelectronic applications due to their potential in high-performance light-emitting devices. Despite the continuous improvement in MQW device efficiency, the physical mechanisms governing light emission under different bias conditions are still not completely understood. In

particular, carrier escape from the quantum wells has been observed even under bias conditions, demonstrating that carrier transport phenomena play a relevant role in the emission process. Therefore, an accurate description of InGaAs/AlGaAs MQW LEDs requires the simultaneous consideration of quantum confinement effects and carrier transport mechanisms, including the role of electron blocking layers (EBLs) in controlling carrier injection and confinement.

In this work, the number of quantum wells is reduced to a single quantum well since the main objective is the evaluation and improvement of the simulation framework rather than the reproduction of a fully realistic device architecture. The aim is to compare the carrier densities predicted by the conventional DD model with those obtained from the quantum-corrected Poisson–Schrödinger–drift–diffusion (PSDD) approach and, subsequently, with the NEGF formalism, highlighting the influence of quantum corrections on the electrostatic potential profile and carrier transport.

The device structure is reported in Tab. 1 and the material parameters are reported in Tab. 2. The simulations are performed using a spatial discretization step of 0.5 nm and an energy discretization step of 5 meV. At thermodynamic equilibrium, the Fermi level is fixed at 0 eV, and the carrier distributions are entirely determined by the electrostatic potential profile.

The equilibrium band diagram is shown in Fig. 8. The inclusion of the Schrödinger correction results in smoother conduction and valence band profiles compared with the conventional PDD solution. This correction allows quantum confinement effects to be incorporated into the evaluation of the electrostatic potential and carrier distributions without introducing the computational cost associated with a fully quantum-mechanical transport model. Under forward bias, as shown in Fig. 9, the conventional DD model predicts a steeper potential barrier at the electron blocking layer (EBL), whereas the PSDD approach produces a smoother effective barrier. This correction partially accounts for quantum confinement effects and enhances carrier penetration through the barrier, yielding a transport description that is closer to the NEGF solution, although coherent tunneling is not explicitly modeled.

Table 1: Layer structure of the simulated `InGaAs` LED.

Layer	**Material**	**Thickness (nm)**	**Type**	**Doping (cm^{-3})**
Bias contact	AlAs	20	p	3×10^{18}
EBL	$Al_{0.2}Ga_{0.8}As$	10	n	1×10^{15}
QB	AlAs	3	n	1×10^{15}
QW	$In_{0.25}Ga_{0.75}As$	4	n	1×10^{15}
Ground contact	AlAs	20	n	3×10^{18}

Table 2: Physical and material parameters adopted in the DD and NEGF simulations [2].

Name	Symbol	Value	Unit
Energy of polar optical phonons	$\hbar\omega_{\mathrm{LO}}$	30	meV
Acoustic deformation potential	D_{aco}	10	eV
Lattice constant	a	0.5	nm
Material density	ρ_{mat}	5.32×10^{-3}	kg/cm^3
Sound velocity	u_l	6.56×10^3	m/s
Optical deformation potential	D_{op}	3×10^9	eV/cm
Maximum transverse wavevector	$k_t^{\max}$	2	nm^{-1}
Extinction coefficient	k_0	0.2831	–
Relative dielectric constant	ε_r	12.9	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.067\, m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.200\, m_0$	kg
Well band gap	$E_{g,\mathrm{well}}$	1.4225	eV
Well valence band offset	$\mathrm{VBO}_{\mathrm{well}}$	-0.6762	eV
Barrier band gap	$E_{g,\mathrm{barrier}}$	1.717	eV
Barrier valence band offset	$\mathrm{VBO}_{\mathrm{barrier}}$	-0.906	eV

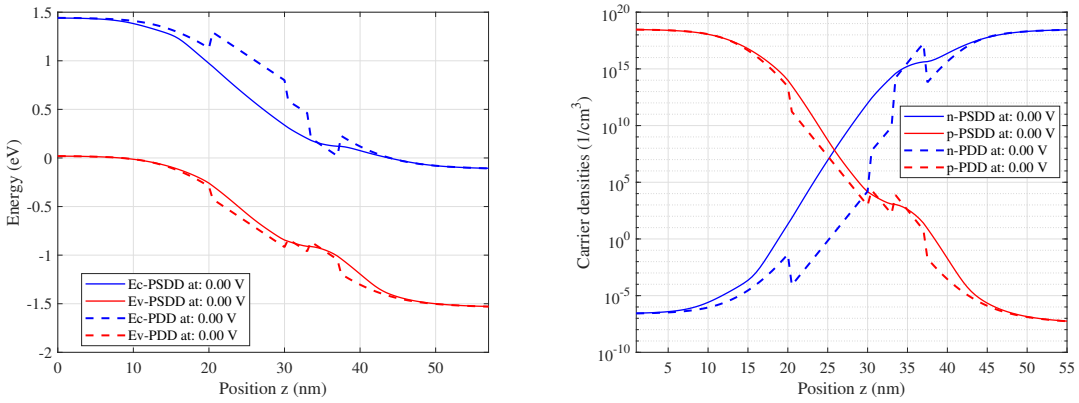


Figure 8: Comparison between the PSDD (solid lines) and conventional PDD (dashed lines) models for the InGaAs-based LED at thermodynamic equilibrium. Left: conduction-band (blue) and valence-band (red) profiles. The figure highlights the band diagrams used by the PSDD (solid lines) and PDD (dashed lines) approaches for the evaluation of the carrier densities. The Schrödinger corrections introduced in the PSDD model result in smoother band profiles, accounting for quantum effects in carrier transport across the quantum well. Right: corresponding electron (blue) and hole (red) carrier-density distributions, highlighting the different carrier redistribution predicted by the two approaches.

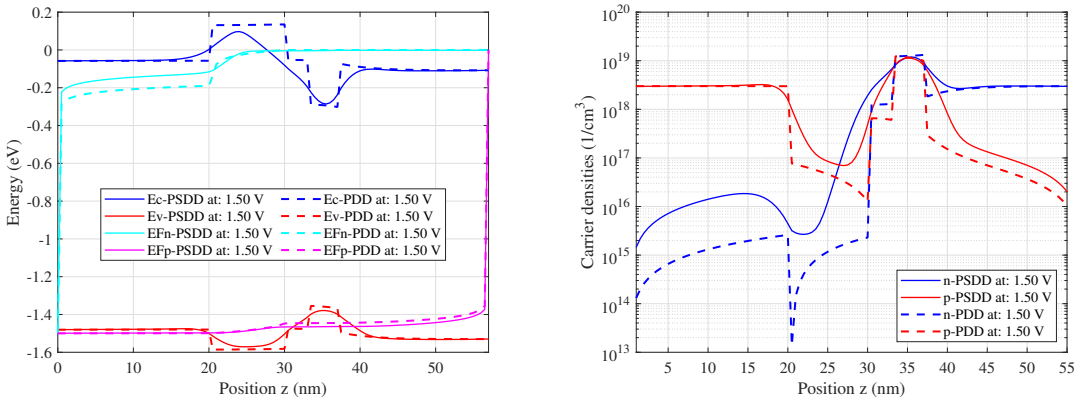


Figure 9: Comparison between the PSDD (solid lines) and the conventional PDD (dashed lines) models for the InGaAs-based LED under a forward bias of 1.5 V. Left: conduction band (blue) and valence band (red) profiles together with the electron (cyan) and hole (magenta) quasi-Fermi levels. The Schrödinger correction produces smoother quasi-Fermi level profiles, reflecting a more realistic description of carrier transport and redistribution. Right: corresponding electron (blue) and hole (red) carrier density distributions. The conventional PDD model underestimates the carrier accumulation in the quantum well and the electron transport across the electron blocking layer compared with the PSDD solution.

The quantum corrections introduced by the PSDD model produce carrier density distributions that are more consistent with quantum transport in nanoscale optoelectronic devices, such as LEDs. Nevertheless, the NEGF formalism provides a more rigorous quantum-mechanical description of carrier transport, explicitly accounting for coherent transport, tunneling, and quantum interference rather than relying on effective band corrections. This improved physical accuracy comes at the expense of a significantly higher computational cost compared with DD-based simulations. Therefore, the PSDD approach is compared with the Poisson–NEGF solution to assess whether the quantum-corrected drift–diffusion model provides an acceptable approximation of the full quantum transport description while maintaining a substantially lower computational burden. Furthermore, the comparison highlights whether the conventional PDD model, which neglects quantum confinement effects, is insufficient for accurately describing carrier transport in nanoscale optoelectronic devices.

Under thermodynamic equilibrium conditions, the PSDD and Poisson–NEGF approaches show a high degree of agreement. The electrostatic potential profile (Fig. 10) is substantially identical for all the considered simulations, while the carrier densities (Fig. 11) highlight the improvement introduced by the quantum-based approaches in reproducing more realistic particle distributions.

When a bias voltage is applied to the device, the different simulation frameworks produce slightly different results. The PDD model shows good agreement with the Poisson–NEGF solution for the electron density profile in the region preceding the electron blocking layer (EBL), as shown in Fig. 13. However, the absence of quantum corrections limits the accuracy of the PDD model in describing carrier transport across the barrier. Conversely, the PSDD approach modifies the effective band profile through Schr"odinger corrections, providing an improved description of carrier confinement. However, it slightly overestimates the tunneling contribution compared with the full quantum-transport solution, resulting in a higher electron accumulation in the region beyond the EBL. The hole density profiles exhibit similar trends for all three simulation approaches under biased conditions. This can be attributed to the lower potential barrier in the valence band, which facilitates hole transport and consequently reduces the impact of quantum corrections on the carrier distribution.

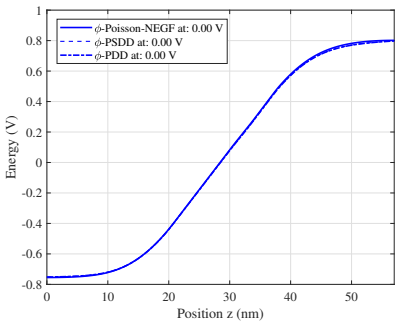


Figure 10: Electrostatic potential profile of the InGaAs-based junction at thermodynamic equilibrium obtained using the different simulation approaches. The high degree of agreement among the employed methods confirms the consistency of the electrostatic solution, while the main differences arise in the quantum description of carrier distributions.

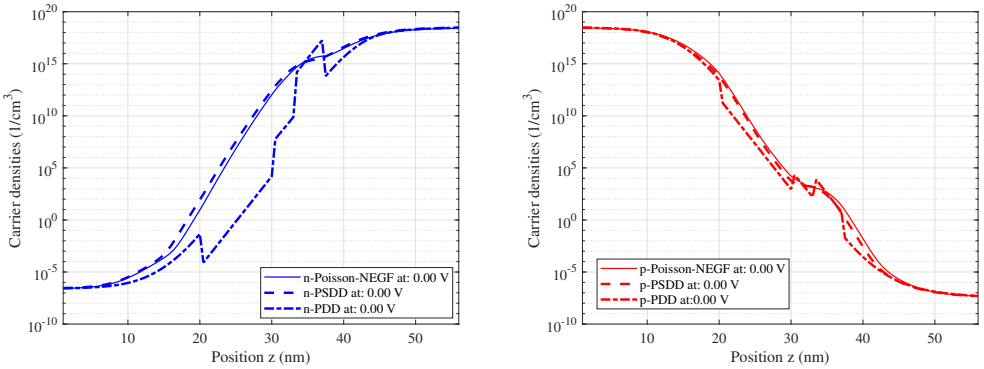


Figure 11: Comparison between the Poisson–NEGF (solid lines), PSDD (dashed lines), and PDD (dash-dotted lines) models for the InGaAs-based LED at thermodynamic equilibrium. Left: electron density profiles. Right: hole density profiles. The inclusion of Schrödinger corrections in the DD framework leads to a significantly better agreement between the PSDD results and the quantum description provided by the Poisson–NEGF approach, demonstrating the capability of effective band corrections to reproduce the main quantum confinement effects.

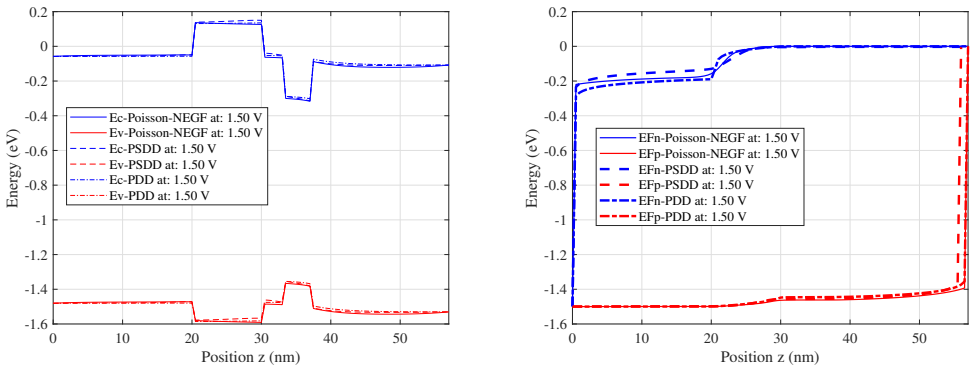


Figure 12: Comparison between the Poisson–NEGF (solid lines), PSDD (dashed lines), and PDD (dash-dotted lines) models for the InGaAs-based LED under a forward bias of 1.5 V. Left: band-diagram profiles (blue: conduction band; red: valence band). Right: quasi-Fermi energy profiles. The band diagrams predicted by the three simulation approaches are nearly equivalent, whereas the quasi-Fermi energy profiles exhibit different behavior depending on whether quantum corrections are included in the model, resulting in different carrier distributions.

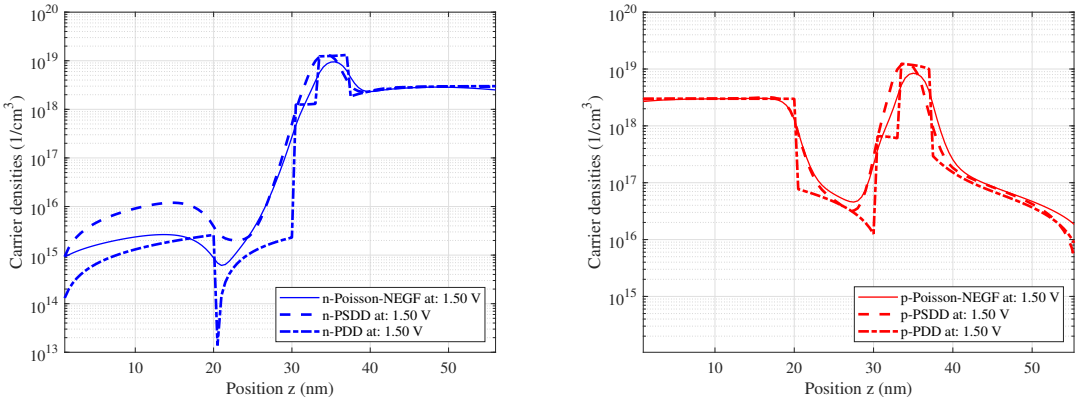


Figure 13: Comparison between the Poisson–NEGF (solid lines), PSDD (dashed lines), and PDD (dash-dotted lines) models for the InGaAs-based LED under a forward bias of 1.5 V. Left: electron density profiles. Right: hole density profiles. The inclusion of Schrödinger corrections in the DD framework provides a good description of carrier confinement within the quantum well and improves the agreement with the Poisson–NEGF results. However, the effective band correction approach slightly overestimates the tunneling contribution under biased conditions compared with the full quantum transport description.

5.2 Drift–Diffusion as an Initial Guess for Poisson–NEGF

The Poisson–NEGF simulations can be initialized using the results of a drift–diffusion simulation. In particular, the electrostatic potential obtained from the DD simulation is used to construct an accurate initial band diagram for the ballistic transport calculation. The ballistic simulation then provides the quasi-Fermi levels, which are used to initialize the Büttiker probe self-energies and thus provide a suitable initial approximation of the scattering processes.

The Poisson–NEGF framework relies on a Newton–Raphson iterative scheme, in which the choice of the initial guess plays a crucial role in the convergence of the physical observables and, consequently, in the overall simulation time. Since the converged Poisson–NEGF solution is independent of the initial guess, provided that convergence is achieved, the main difference between alternative initializations lies in the number of iterations required to reach self-consistency. As shown in Fig. 14, employing the PSDD solution as the initial guess significantly accelerates the convergence of both the electrostatic potential and the carrier-density residuals, which are strongly coupled to the quasi-Fermi level updates. The conventional PDD solution also represents a suitable initial guess for the NEGF solver, although it requires a slightly larger number of iterations to achieve convergence.

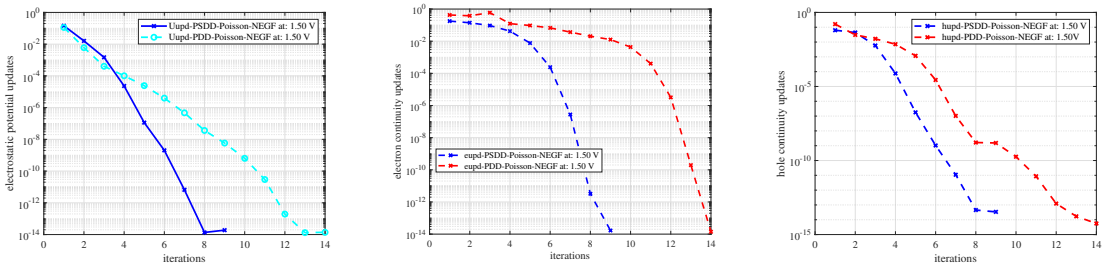


Figure 14: Convergence comparison between PSDD and PDD initial guesses for the Poisson–NEGF simulation of the InGaAs-based LED under forward bias of 1.5 V. Left: electrostatic potential update. Center: electron continuity residual update. Right: hole continuity residual update. The PSDD initialization consistently accelerates convergence, reducing the number of Newton–Raphson iterations required to reach the self-consistent solution.

5.3 Frozen-Field Non-Equilibrium Green's Function

In this section, the frozen-field approach is investigated by removing the Poisson equation from the Newton–Raphson self-consistent solution of the NEGF framework. In this formulation, the electrostatic potential is kept fixed throughout the simulation and is no longer treated as an unknown, allowing the self-consistent procedure to involve only the transport equations.

The objective is to determine which electrostatic potential profile provides the most suitable compromise between computational efficiency and transport accuracy. In particular, electrostatic potentials obtained from the conventional DD model and from the quantum-corrected PSDD approach are considered as frozen-field inputs to the NEGF solver.

The analysis is carried out on the InGaAs-based LED introduced in the previous section. The frozen-field implementation is obtained by removing the Jacobian blocks associated with the Poisson equation of Eq. 4.71, so that the electrostatic potential is excluded from the set of unknowns. As a consequence, the size of the Jacobian matrix is reduced, leading to lower computational cost and shorter simulation times.

$$\underline{\underline{J}} = \begin{pmatrix}
\underline{\underline{J}}^{(p,U)} & \underline{\underline{J}}^{(p,\mu_n)} & \underline{\underline{J}}^{(p,\mu_p)} & \underline{J}^{(p,V)} & 0 & 0 & \underline{J}^{(p,\mu_{nl})} & \underline{J}^{(p,\mu_{nr})} & \underline{J}^{(p,\mu_{pl})} & \underline{J}^{(p,\mu_{pr})} \\
\underline{\underline{J}}^{(e,U)} & \underline{\underline{J}}^{(e,\mu_n)} & \underline{\underline{J}}^{(e,\mu_p)} & \underline{J}^{(e,V)} & 0 & 0 & \underline{J}^{(e,\mu_{nl})} & \underline{J}^{(e,\mu_{nr})} & 0 & 0 \\
\underline{\underline{J}}^{(h,U)} & \underline{\underline{J}}^{(h,\mu_n)} & \underline{\underline{J}}^{(h,\mu_p)} & \underline{J}^{(h,V)} & 0 & 0 & 0 & 0 & \underline{J}^{(h,\mu_{pl})} & \underline{J}^{(h,\mu_{pr})} \\
0 & 0 & 0 & J^{(v,V)} & J^{(v,J_L)} & J^{(v,J_R)} & 0 & 0 & 0 & 0 \\
\underline{\underline{J}}^{(j_L,U)} & \underline{\underline{J}}^{(j_L,\mu_n)} & \underline{\underline{J}}^{(j_L,\mu_p)} & 0 & J^{(j_L,J_L)} & 0 & \underline{J}^{(j_L,\mu_{nl})} & \underline{J}^{(j_L,\mu_{nr})} & \underline{J}^{(j_L,\mu_{pl})} & \underline{J}^{(j_L,\mu_{pr})} \\
\underline{\underline{J}}^{(j_R,U)} & \underline{\underline{J}}^{(j_R,\mu_n)} & \underline{\underline{J}}^{(j_R,\mu_p)} & 0 & 0 & J^{(j_R,J_R)} & \underline{J}^{(j_R,\mu_{nl})} & \underline{J}^{(j_R,\mu_{nr})} & \underline{J}^{(j_R,\mu_{pl})} & \underline{J}^{(j_R,\mu_{pr})} \\
\underline{\underline{J}}^{(n_L,U)} & \underline{\underline{J}}^{(n_L,\mu_n)} & \underline{\underline{J}}^{(n_L,\mu_p)} & 0 & 0 & 0 & \underline{J}^{(n_L,\mu_{nl})} & \underline{J}^{(n_L,\mu_{nr})} & 0 & 0 \\
\underline{\underline{J}}^{(n_R,U)} & \underline{\underline{J}}^{(n_R,\mu_n)} & \underline{\underline{J}}^{(n_R,\mu_p)} & 0 & 0 & 0 & \underline{J}^{(n_R,\mu_{nl})} & \underline{J}^{(n_R,\mu_{nr})} & 0 & 0 \\
\underline{\underline{J}}^{(p_L,U)} & \underline{\underline{J}}^{(p_L,\mu_n)} & \underline{\underline{J}}^{(p_L,\mu_p)} & 0 & 0 & 0 & 0 & 0 & \underline{J}^{(p_L,\mu_{pl})} & \underline{J}^{(p_L,\mu_{pr})} \\
\underline{\underline{J}}^{(p_R,U)} & \underline{\underline{J}}^{(p_R,\mu_n)} & \underline{\underline{J}}^{(p_R,\mu_p)} & 0 & 0 & 0 & 0 & 0 & \underline{J}^{(p_R,\mu_{pl})} & \underline{J}^{(p_R,\mu_{pr})}
\end{pmatrix} \tag{5.1}$$

The DD simulations are consistently faster than the corresponding NEGF simulations, independently of the device dimensions and material system considered. To reach a high forward-bias voltage, both approaches generally require a step-by-step bias ramp, in which the solution obtained at the previous bias point is used as the initial guess for the next one. However, the computational time required to reach a given forward bias is significantly lower when using the DD model. This difference can be exploited to reduce the computational cost of NEGF simulations. Instead of progressively increasing the bias within the NEGF framework, the DD model can first be used to ramp the device voltage up to the desired operating point. The converged DD solution at this voltage is then employed as the initial guess for a NEGF simulation. In this way, the quantum transport description is introduced only at the target operating bias, avoiding the computationally expensive step-by-step bias ramp required by the NEGF approach. This DD-to-NEGF initialization strategy therefore allows the device to be simulated directly within the NEGF framework at the desired operating bias, while retaining the computational advantage of the faster DD bias ramp.

For clarity, the three simulation approaches considered in the following sections are summarized below:

1. **PDD–NEGF**: frozen-field NEGF simulation initialized using the electrostatic potential obtained from a PDD simulation;

2. **PSDD–NEGF**: frozen-field NEGF simulation initialized using the quantum-corrected elec-

trostatic potential obtained from a PSDD simulation;

3. **PSDD–Poisson–NEGF**: self-consistent Poisson–NEGF simulation initialized using the converged solution of a PSDD simulation. The choice of using either a PDD or PSDD solution as the initial guess for the Poisson–NEGF calculation does not affect the final converged solution, but only the convergence time, as discussed in Section 5.2.

The first configuration considered employs the electrostatic potential obtained from the conventional DD simulation as the frozen-field input to the NEGF solver.

Simulation guess: Poisson–Drift–Diffusion

The first approach consists of evaluating the physical observables using the electrostatic potential obtained from a PDD simulation, where no quantum corrections are applied to the conduction and valence band profiles. Under thermodynamic equilibrium conditions, the PDD–NEGF approach shows excellent agreement with the conventional Poisson–NEGF simulation, both in terms of the electrostatic potential profile and the electron and hole density distributions (Fig. 15).

Under forward bias conditions, as shown in Fig. 16, the carrier density distributions remain in good agreement with the reference Poisson–NEGF simulation. A slight deviation can be observed in the hole density profile, which results in a small modification of the hole transport contribution and consequently affects the total current density along the device.

Using a DD potential profile does not compromise the NEGF description of quantum confinement, even under non-equilibrium conditions. Consequently, PDD simulations can be used to reach high bias voltages without significantly compromising the accuracy of the quantum transport description.

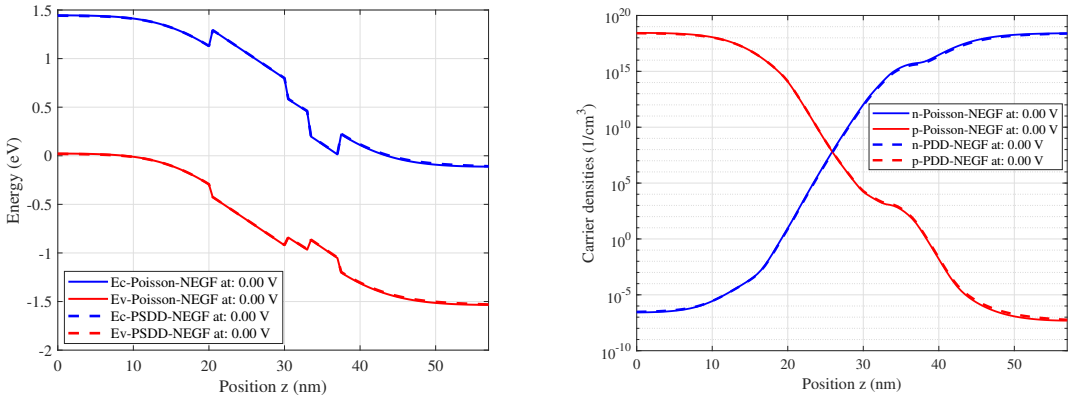


Figure 15: Comparison between the conventional Poisson–NEGF (solid lines) and the PDD–NEGF (dashed lines) models for the InGaAs-based LED at thermodynamic equilibrium. Left: band diagram profiles. Right: electron (blue) and hole (red) density distributions. Both quantities show excellent agreement between the two approaches.

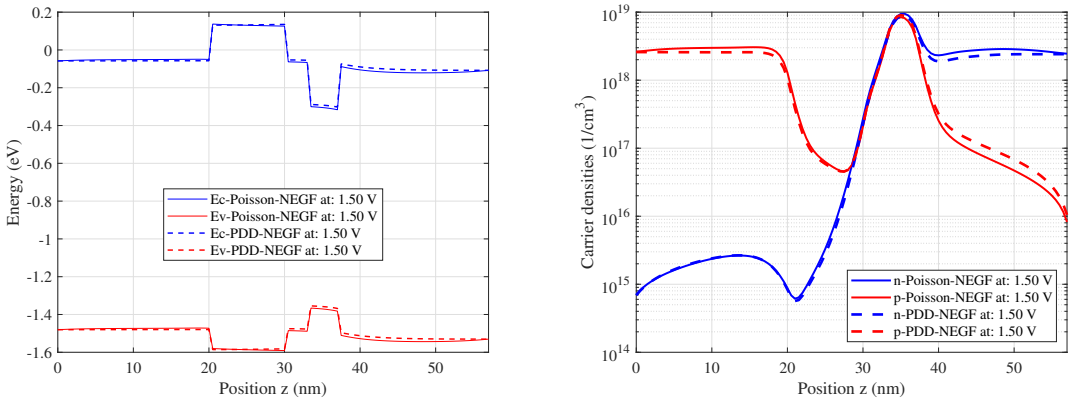


Figure 16: Comparison between the conventional Poisson–NEGF (solid lines) and the PDD–NEGF (dashed lines) models for the InGaAs-based LED under a forward bias of 1.5V. Left: band diagram profiles. Right: electron (blue) and hole (red) density distributions. Both quantities show excellent agreement between the two approaches.

Simulation guess: Poisson–Schrödinger–Drift–Diffusion

The inclusion of quantum corrections in the PSDD simulation produces only minor differences compared with the PDD-based frozen-field approach. Since the equilibrium simulations yield essentially identical results, only the band diagram and carrier density distributions under forward bias are reported. As shown in Fig. 17, the PSDD–NEGF results exhibit excellent agreement with those obtained from the fully self-consistent Poisson–NEGF simulation. These results demonstrate that both the PDD and PSDD electrostatic potential profiles provide reliable initial guesses for the frozen-field NEGF approach, enabling an accurate description of carrier transport and the associated physical observables while significantly reducing the computational cost.

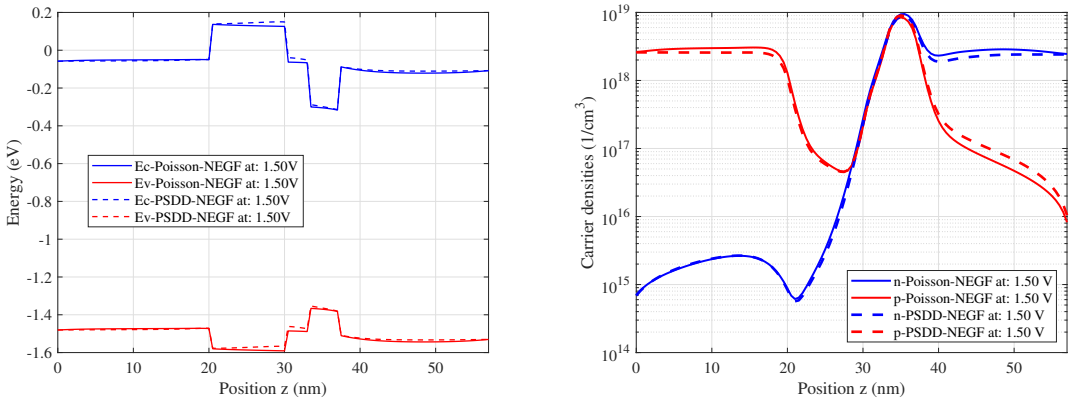


Figure 17: Comparison between the conventional Poisson–NEGF (solid lines) and the PSDD–NEGF (dashed lines) models for the InGaAs-based LED under a forward bias of 1.5 V. Left: band diagram profiles. Right: electron and hole carrier density profiles, showing excellent agreement between the two simulations, with only a slight underestimation of the carrier densities inside the quantum well in the PSDD–NEGF approach.

5.4 Drift–Diffusion: Incomplete Ionization

To further improve the accuracy of the electrostatic potential employed in the frozen-field NEGF simulations without increasing the computational cost of the quantum transport calculation, incomplete thermal ionization of the dopants is included at the drift–diffusion level. The ionized donor and acceptor concentrations are corrected according to

$$N_D^+ = \frac{N_D}{1 + g_D \frac{n}{n_1}}, \tag{5.2}$$

$$N_A^- = \frac{N_A}{1 + g_A \frac{p}{p_1}}, \tag{5.3}$$

where g_D and g_A are the donor and acceptor degeneracy factors (equal to 2 in the present simulations), while n_1 and p_1 are the characteristic carrier densities associated with the donor and acceptor ionization energies ΔE_D and ΔE_A,

$$n_1 = N_C \exp\left(-\frac{\Delta E_D}{k_B T}\right), \tag{5.4}$$

$$p_1 = N_V \exp\left(-\frac{\Delta E_A}{k_B T}\right). \tag{5.5}$$

These corrections modify the ionized dopant distribution in the Poisson equation, resulting in a different electrostatic potential profile used as the frozen field in the NEGF calculation. As shown in Fig. 18, incomplete ionization slightly reduces both electron and hole densities throughout the device, with a stronger effect on the majority carriers. This reduction may decrease radiative recombination and, consequently, the internal quantum efficiency of the LED. However, the impact of incomplete ionization depends strongly on the dopant ionization energy and the semiconductor material, and may be negligible at room temperature for sufficiently shallow dopants. The results show that incomplete ionization can be implemented within the simpler DD formalism without requiring additional corrections to the NEGF simulations, thus avoiding any additional computational cost.

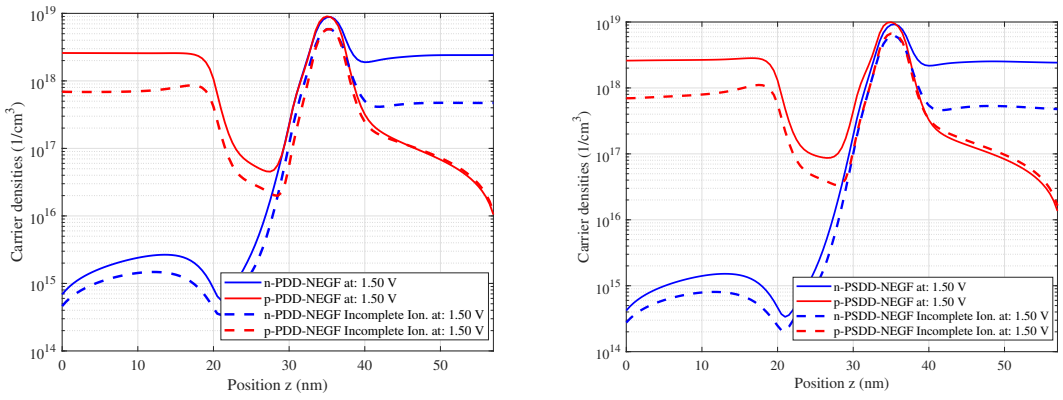


Figure 18: Comparison of NEGF simulations for the InGaAs-based junction under a forward bias of 1.5 V, using electrostatic potentials obtained from PDD (left) and PSDD (right). Solid lines correspond to complete ionization, while dashed lines include incomplete ionization in the drift–diffusion stage. Electron and hole density profiles show the impact of incomplete ionization on the carrier distributions.

5.5 InGaN/GaN Single–Quantum–Well Light–Emitting Diode

The device investigated in this section is LED based on an InGaN alloy with a single quantum well (QW), first demonstrated in 1993 [20]. III–nitride LEDs are affected by several efficiency limitations, including the efficiency droop phenomenon [21] [22] and the so-called "green gap" [23], namely the reduction of the emission efficiency with increasing wavelength. These effects are mainly related to material defects, carrier leakage mechanisms, and polarization-induced fields in the heterostructure. For short-wavelength emission, InGaN/GaN LEDs can achieve internal quantum efficiencies of approximately 80–90%, making them a relevant platform for high-efficiency optoelectronic applications [24] [25] [26] [27] [28].

A relevant phenomenon in these devices is trap-assisted tunneling (TAT) [29] [30] [31], in which electrons are transferred to a trap level within the band gap and subsequently tunnel toward the valence band by means of multiphonon emission or absorption. These defect-mediated transitions introduce additional leakage pathways, increasing the non-radiative current component and consequently reducing device efficiency [30]. A quantum-corrected Shockley–Read–Hall model provides an effective approach for incorporating TAT-related effects, as demonstrated by Gonzalez et al. [32]. Although a fully quantum-mechanical description based on NEGF self-energies would provide a more accurate treatment, it requires the implementation of non-local self-energies, resulting in a significant increase of the computational cost. Besides defect-assisted transport, polarization effects play a crucial role in III–nitride heterostructures, especially in the presence of quantum wells. The lack of inversion symmetry in the wurtzite crystal structure generates spontaneous polarization, while strain induced by lattice mismatch introduces a further piezoelectric contribution. The resulting polarization charges modify the electrostatic potential inside the device, producing strong internal electric fields that affect carrier confinement and spatial separation of electrons and holes. This redistribution reduces the electron– hole wavefunction overlap and consequently decreases the radiative recombination efficiency.

The polarization parameters used in this work are taken from [33] and [34]. The total polarization vector is obtained as the sum of the spontaneous and piezoelectric contributions:

$$\mathbf{P}_{tot} = \mathbf{P}_{sp} + \mathbf{P}_{pz}. \tag{5.6}$$

The associated bound charge density is obtained from the divergence of the polarization field:

$$\rho_{pol} = -\nabla \cdot \mathbf{P}_{tot}. \tag{5.7}$$

For a one-dimensional structure, assuming the polarization direction aligned with the growth axis z, this expression becomes:

$$\rho_{pol}(z) = -\frac{\partial P_{tot}(z)}{\partial z}. \tag{5.8}$$

Therefore, abrupt variations of the polarization at heterointerfaces generate localized sheet charges. For an interface between two materials with polarization values P_1 and P_2, the interface charge density is:

$$\sigma_{pol} = P_2 - P_1. \tag{5.9}$$

The polarization contribution is included in the electrostatic calculation as an additional charge term in the Poisson equation:

$$\nabla \cdot (\epsilon \nabla \phi) = -(\rho_{free} + \rho_{pol} + qN_{DOP}), \tag{5.10}$$

where the free-carrier contribution is given by:

$$\rho_{free} = q(p - n). \tag{5.11}$$

Hence, the total charge density employed in the electrostatic solution of Eq. 2.29 is:

$$\rho_{tot} = q(p - n + N_{DOP}) + \rho_{pol}. \tag{5.12}$$

Finally, the larger band gap of III–nitride materials imposes more demanding numerical conditions compared with the previously investigated InGaAs-based p–i–n junction. To improve numerical stability, the simulations are performed using quadruple precision, rather than the double precision adopted for the InGaAs devices. Since significant current levels are obtained only at relatively high applied voltages, the analysis is restricted to the voltage range between 2 V and 4 V.

Polarization charges and incomplete dopant ionization are included in the DD simulations, with their impact being particularly relevant near the contacts, as shown in Fig. 20, where the majority-carrier densities are significantly reduced. In contrast, the carrier distributions within the quantum well are only marginally affected by these effects, despite slight modifications of the electrostatic potential profile across the quantum well.

Table 3: Layer structure of the simulated `pin_AlInGaN` single–quantum–well LED.

Layer	**Material**	**Thickness (nm)**	**Type**	**Doping (cm^{-3})**
Bias contact	GaN	8.0	p	5×10^{19}
EBL	$Al_{0.13}Ga_{0.87}N$	17.25	p	5×10^{19}
QB	GaN	4.0	n	1×10^{16}
QW	$In_{0.195}Ga_{0.805}N$	4.0	n	1×10^{16}
Spacer	GaN	4.0	n	1×10^{20}
Ground contact	GaN	8.0	n	1×10^{20}

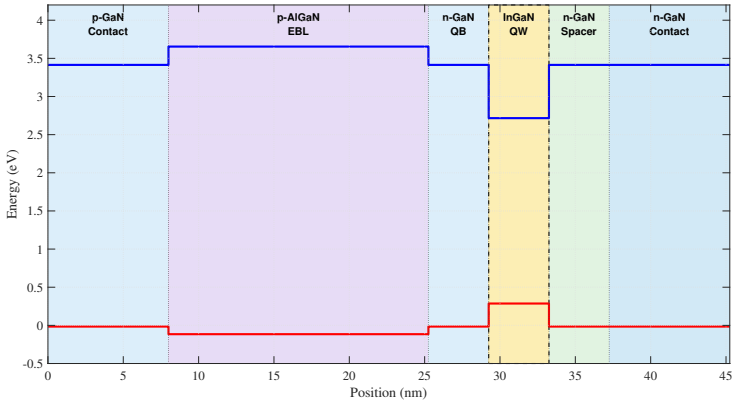


Figure 19: schematic representation of the band diagram, showing the conduction band in blue and the valence band in red.

Table 4: Physical and material parameters adopted in the DD and NEGF simulations [33] [34] [35] [36].

Name	**Symbol**	**Value**	**Unit**
Temperature	T	300	K
Energy of polar optical phonons	$\hbar\omega_{\mathrm{LO}}$	90	meV
Acoustic deformation potential	D_{aco}	10	eV
Lattice constant	a	0.5	nm
Material density	ρ_{mat}	6.15×10^{-3}	kg/cm^3
Sound velocity	u_l	6.56×10^3	m/s
Optical deformation potential	D_{op}	3×10^9	eV/cm
Maximum transverse wavevector	$k_t^{\max}$	2	nm^{-1}
Extinction coefficient	k_0	0.2831	–
GaN material parameters			
Relative dielectric constant	ε_r	9.7	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.2\,m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.4\,m_0$	kg
Band gap	E_g	3.43	eV
Valence band offset	VBO	-0.0157	eV
Spontaneous polarization	P_{sp}	-0.0340	C/m^2
Piezoelectric polarization	P_{pz}	0	C/m^2
$\mathbf{In_{0.195}Ga_{0.805}N}$ material parameters			
Relative dielectric constant	ε_r	10.76	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.2\,m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.4\,m_0$	kg
Band gap	E_g	2.43	eV
Valence band offset	VBO	0.2863	eV
Spontaneous polarization	P_{sp}	-0.0296	C/m^2
Piezoelectric polarization	P_{pz}	0.0222	C/m^2
$\mathbf{Al_{0.13}Ga_{0.87}N}$ material parameters			
Relative dielectric constant	ε_r	9.54	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.2\,m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.4\,m_0$	kg
Band gap	E_g	3.77	eV
Valence band offset	VBO	-0.1156	eV
Spontaneous polarization	P_{sp}	-0.0391	C/m^2
Piezoelectric polarization	P_{pz}	-0.0036	C/m^2

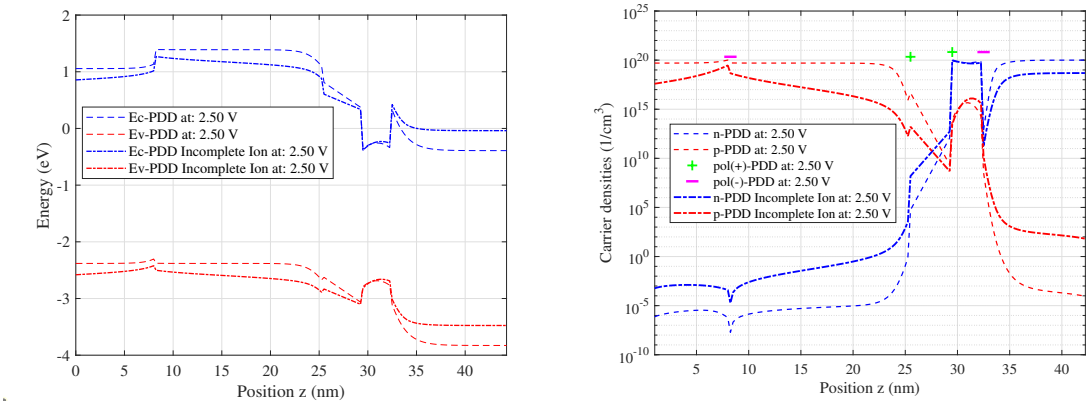


Figure 20: Comparison between the standard PDD solution (dashed lines) and the PDD solution including polarization charges and incomplete dopant ionization (dash-dotted lines) for the AlInGaN-based LED under a forward bias of 2.5 V. The band diagram (left) highlights the modifications induced by polarization charges and incomplete ionization, while the carrier density profiles (right) show that the largest variations occur near the contacts and in the electron blocking layer (EBL), whereas the quantum well is only weakly affected. The markers indicate the location and sign of the polarization charges, with + denoting positive polarization charges and − denoting negative polarization charges.

Poisson Drift Diffusion–Non-Equilibrium Green's Function

The Poisson Drift Diffusion–NEGF analysis focuses on the local density of states (LDOS) and the spectral current density. Figure 20 shows the differences between the PDD simulations with and without the inclusion of incomplete dopant ionization and polarization charges in the carrier density calculation. Polarization charges are localized at the interfaces between different materials, with the most relevant contributions occurring near the quantum well and the electron blocking layer (EBL). As shown in Fig. 22, the LDOS, calculated from Eq. 4.65, clearly reveals the quantum confinement within the quantum well through the discretization of the energy levels. Since the device contains a single quantum well, the LDOS exhibits well-defined confined electron and hole states with limited energy broadening, indicating relatively weak coupling between the confined states and the surrounding layers. For the following analysis, the position $z = 30.75\,\mathrm{nm}$, corresponding to the quantum well layer, is selected to extract the energy-resolved profiles. The non-equilibrium analysis is performed at a forward bias of 3.0 V, corresponding to the upper limit of the LED's operating range, which extends from 2.5 V to 3.0 V.

The electron blocking layer (EBL) strongly suppresses the electron contribution to the current, resulting in a dominant hole contribution to the current through the quantum well. The corresponding spectral current density, shown in Fig. 22, highlights the energy-resolved contributions of electrons and holes to the total current. In this analysis, recombination processes are temporarily neglected, allowing the transport properties and the role of the EBL to be examined independently of recombination effects.

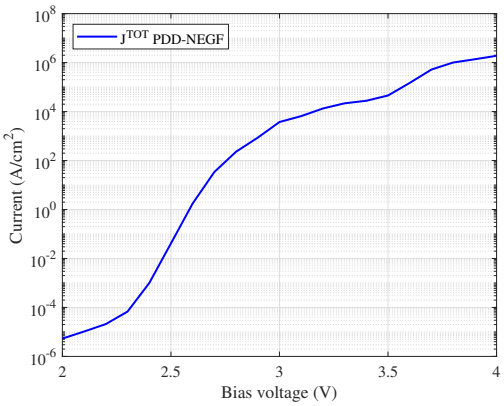
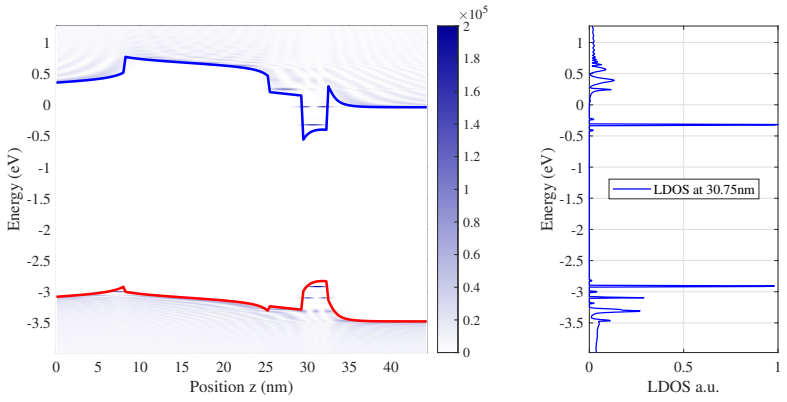
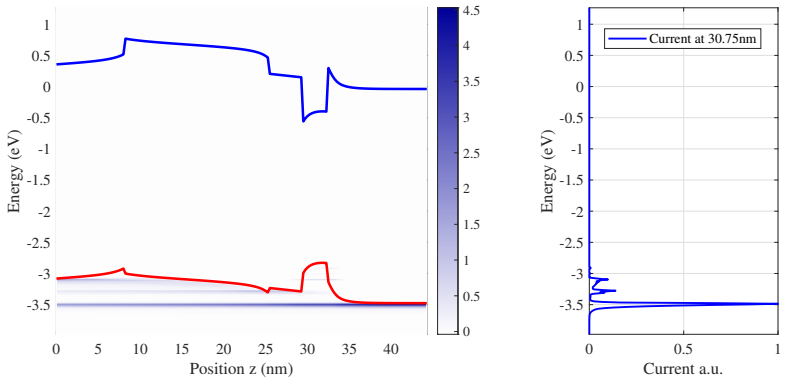


Figure 21: Current–voltage (J–V) characteristics of the AlInGaN-based LED obtained from frozen-field PDD–NEGF simulations for forward bias voltages ranging from 2 V to 4 V.



(a)



(b)

Figure 22: Frozen-field PDD–NEGF simulation of an AlInGaN-based LED under a forward bias of 3.0 V, including polarization charges and incomplete dopant ionization. (a) Local density of states (LDOS) for $\mathbf{k} = 0$ and the corresponding energy-resolved profile extracted at $z = 30.75$ nm. (b) Spectral current density and the corresponding energy-resolved profile at $z = 30.75$ nm. The LDOS illustrates the spatial and energetic distribution of the available states, while the spectral current density shows the dominant contribution of holes to the current, with recombination processes neglected. All energy-resolved profiles are normalized.

5.6 Recombination Current Analysis

The attention is shifted to the recombination current densities, which are investigated through the implementation of the ABC model introduced in Section 4.9. The NEGF analysis aims to understand the emission behaviour of the AlInGaN-based LED and to investigate the recombination mechanisms responsible for some of the well-known characteristics of III-nitride LEDs, such as the efficiency droop and the limitations associated with high carrier injection.

The frozen-field approach allows the inclusion of polarization and incomplete ionization effects in the electrostatic potential through the modified charge distribution obtained from the Poisson equation. Their influence on the recombination currents is firstly investigated. Subsequently, the role of the EBL is analyzed, affected by the variation of the Al molar fraction in the AlInGaN alloy, with particular focus on how the radiative recombination component, which directly contributes to photon generation in LED applications.

The recombination parameters adopted in the ABC model are reported in Tab. 5, focusing on the quantum well performance. The coefficients have been selected according to values reported in the literature. In particular, the Auger recombination coefficients are chosen within the range reported in [37], while the radiative recombination coefficient is selected considering the large variability reported for InGaN-based quantum wells [38]. The adopted value provides a reasonable estimation of the radiative recombination processes investigated in this work.

The recombination current density contributions are evaluated by integrating the local net recombination rates along the device thickness:

$$J_{\mathrm{RAD}} = q \int U^{\mathrm{RAD}}(z)\, dz \tag{5.13}$$

$$J_{\mathrm{SRH}} = q \int U^{\mathrm{SRH}}(z)\, dz \tag{5.14}$$

$$J_{\mathrm{AUG}} = q \int U^{\mathrm{AUG}}(z)\, dz \tag{5.15}$$

Table 5: Recombination parameters adopted in the ABC model for the $In_{0.195}Ga_{0.805}N$ quantum well layer.

Name	**Symbol**	**$In_{0.195}Ga_{0.805}N$**	**Unit**
Electron SRH lifetime	τ_n	1.0×10^{-9}	s
Hole SRH lifetime	τ_p	1.0×10^{-9}	s
SRH trap energy level	E_{trap}	0	eV
Radiative recombination coefficient	B_{rad}	1.0×10^{-17}	m^3/s
Electron Auger coefficient	C_{nnp}	2.7×10^{-43}	m^6/s
Hole Auger coefficient	C_{ppn}	2.7×10^{-43}	m^6/s

Polarization Effects on Recombination Mechanisms

The PDD–NEGF simulations show that polarization effects mainly modify the effective operating range of the LED, as reported in Fig. 23. The internal electric fields induced by polarization charges modify carrier transport and confinement, without significantly favoring any specific recombination mechanism, while resulting in an overall reduction in the current density. The limited impact of

polarization effects on the device characteristics may be attributed to the spatially localized nature of the polarization charges and the associated electric fields, which primarily affect the regions close to the corresponding heterointerfaces.

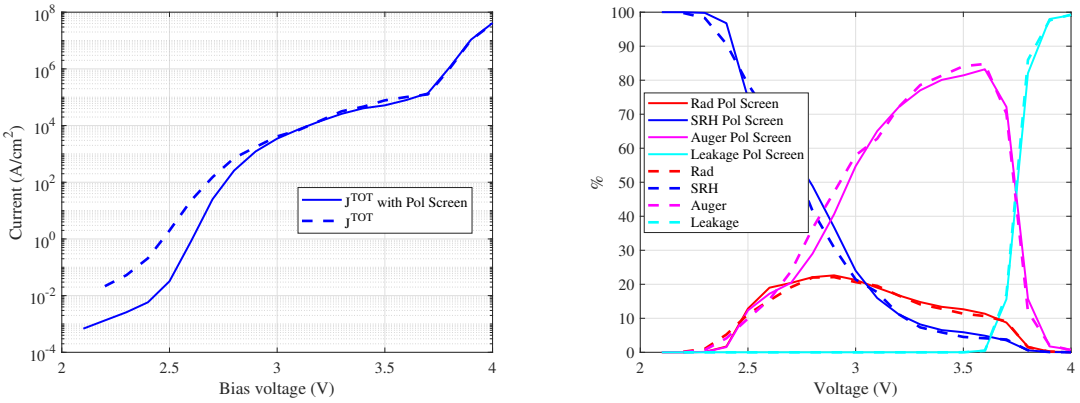


Figure 23: Comparison between PDD–NEGF simulations with and without polarization effects for the AlInGaN-based LED. On the left, the J–V characteristics show that polarization effects mainly influence the low-bias regime, whereas only minor differences in the current density are observed within the LED operating range. On the right, polarization charges do not significantly affect the relative contributions of the different recombination mechanisms to the total recombination current.

Incomplete Ionization Effects on Recombination Mechanisms

The implementation of incomplete ionization shows that the reduction of ionized majority carriers is accompanied by a redistribution of the carrier densities, with an increase of the minority carrier concentration in the affected regions. This behaviour partially compensates the variation of the carrier product np, limiting the impact on the overall recombination current densities in the J–V characteristics (Fig. 24).

A comparison between a PDD–NEGF simulation including polarization and incomplete ionization effects and a PDD–NEGF simulation assuming complete dopant ionization shows that, for the investigated structure and operating conditions, incomplete ionization produces a stronger modification of the LED behaviour than polarization effects. This result is obtained at room temperature ($T = 300\,\mathrm{K}$), where the dopant ionization degree is higher than at lower temperatures; nevertheless, the incomplete activation of dopants still significantly affects the carrier distribution. In the simulations, Si donors and Mg acceptors are considered, with activation energies of 20 meV and 210 meV [39], respectively. The inclusion of incomplete ionization mainly modifies the hole density distribution along the device, particularly near the contact regions, whereas the quantum well region exhibits only a limited redistribution of carriers.

The reduced availability of free carriers from dopant ionization modifies the carrier injection and transport mechanisms, changing the balance between radiative and non-radiative recombination contributions. At low current densities, the increased SRH contribution competes with the radiative recombination channel, reducing the fraction of carriers contributing to light emission and limiting the effective operating range of the LED. At higher current densities, leakage currents become progressively relevant, further affecting the recombination current behaviour.

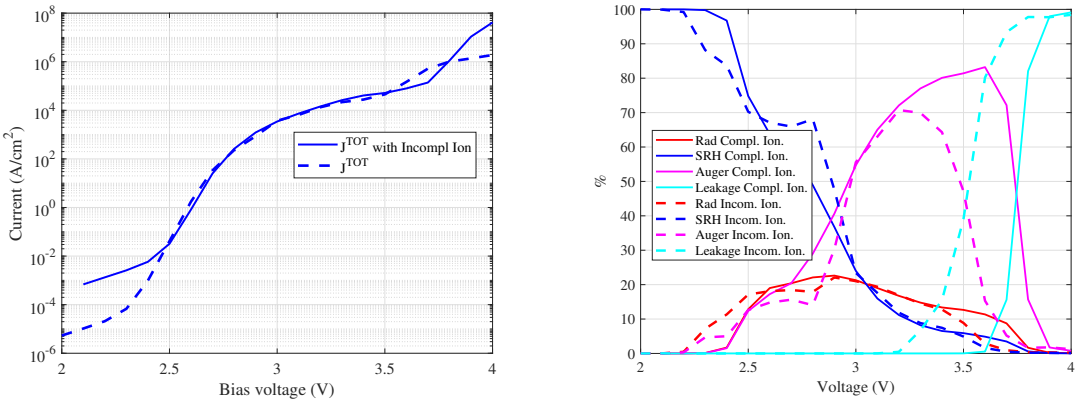


Figure 24: Comparison between PDD–NEGF simulations with and without incomplete ionization effects, including polarization screening, for the AlInGaN-based LED. The inclusion of incomplete ionization significantly extends the effective operating voltage range of the LED. Moreover, the leakage current becomes significant only at higher forward-bias voltages, delaying the onset of leakage-dominated transport.

EBL Influence on Recombination Mechanisms

To evaluate the influence of the electron blocking layer (EBL) on the recombination mechanisms, the Al molar fraction of the EBL is reduced from 0.13 to 0.05, thereby lowering the conduction-band barrier for electrons. As a consequence, electron overflow towards the p-side increases, leading to a larger coherent electron current through the device.

This behavior is clearly observed in Fig. 25, where the Al content is reduced to $x_{\mathrm{Al}} = 0.05$. The lower electron barrier allows a significantly larger coherent electron current to reach the p-side, demonstrating the reduced blocking capability of the EBL.

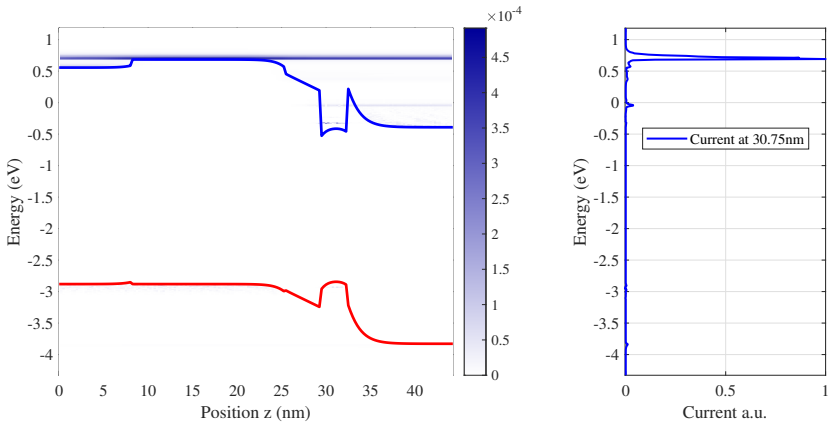


Figure 25: Frozen-field PDD–NEGF simulation including polarization charges for the AlInGaN-based LED with an electron blocking layer (EBL) having an Al molar fraction of $x = 0.05$. The spectral current density under a forward bias of 3.0 V and the corresponding energy-resolved profile extracted at $z = 30.75$ nm are shown. Recombination mechanisms are neglected to highlight the dominant contribution of electrons to the current density.

The recombination current analysis (Fig. 26) confirms this behavior. The weaker EBL enhances electron leakage, thereby reducing the radiative recombination current and consequently narrowing the effective operating range of the LED by lowering its upper bias limit.

Conversely, increasing the Al content of the EBL improves electron confinement inside the quantum well by suppressing electron overflow. However, a higher Al concentration also increases

the valence-band barrier, making hole injection into the active region more difficult. In addition, the larger spontaneous and piezoelectric polarization associated with Al-rich AlGaN enhances the internal electric fields, producing stronger band bending and quantum-confined Stark effect (QCSE). The resulting spatial separation of the electron and hole wavefunctions decreases their overlap, reducing the radiative recombination efficiency despite the improved electron confinement.

Table 6: Layer structure of the simulated `pin_AlInGaN` single–quantum–well LED with a modified electron blocking layer (EBL), highlighting the change in Al molar fraction in red.

Layer	**Material**	**Thickness (nm)**	**Type**	**Doping (cm^{-3})**
Bias contact	GaN	8.0	p	5×10^{19}
EBL	$Al_{0.05}Ga_{0.95}N$	17.25	p	5×10^{19}
QB	GaN	4.0	n	1×10^{16}
QW	$In_{0.195}Ga_{0.805}N$	4.0	n	1×10^{16}
Spacer	GaN	4.0	n	1×10^{20}
Ground contact	GaN	8.0	n	1×10^{20}

Table 7: Physical and material parameters adopted in the DD and NEGF simulations [33] [34] [35].

Name	**Symbol**	**Value**	**Unit**
$Al_{0.05}Ga_{0.95}N$ material parameters			
Relative dielectric constant	ε_r	9.64	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.2\,m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.4\,m_0$	kg
Band gap	E_g	3.56	eV
Valence band offset	VBO	-0.0541	eV
Spontaneous polarization	P_{sp}	-0.0359	C/m^2
Piezoelectric polarization	P_{pz}	-0.0013	C/m^2

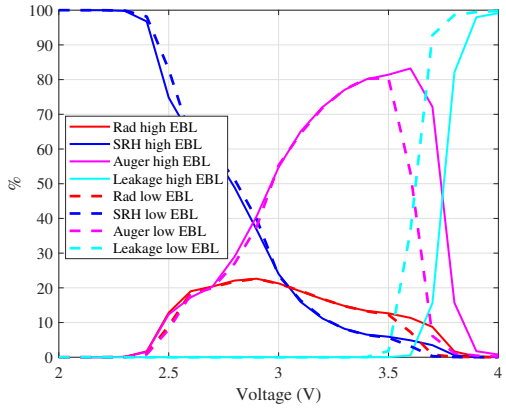


Figure 26: Frozen-field PDD–NEGF simulation including polarization charges for the AlInGaN-based LED with an electron blocking layer (EBL) having an Al molar fraction of $x = 0.05$. The recombination current contributions are shown as a function of the applied bias, highlighting the balance between radiative and leakage currents for the considered EBL composition.

5.7 Internal Quantum Efficiency

The internal quantum efficiency (IQE) is a fundamental parameter for evaluating LED performance and its evolution under different bias conditions. However, its experimental determination is challenging, since IQE is an internal quantity related to the competition among leakage, radiative and non-radiative recombination mechanisms inside the active region. Experimental estimations generally require assumptions on optical extraction, low-temperature efficiency and carrier distribution, which may introduce uncertainties.

Using the recombination parameters reported in Tab. 5, the maximum IQE estimated for the investigated structure is approximately 25%. This result has to be interpreted considering the limitations of the adopted model. Firstly, the investigated device consists of a single quantum well, while practical nitride LEDs generally employ multiple quantum well structures optimized for carrier confinement and optical emission. Secondly, the ABC coefficients are obtained from literature values and represent an approximation, since the exact radiative and non-radiative parameters strongly depend on material quality, defect density and device architecture.

The IQE is evaluated according to:

$$IQE = \frac{J_{\mathrm{RAD}}}{J_{\mathrm{RAD}} + J_{\mathrm{SRH}} + J_{\mathrm{AUG}} + J_{\mathrm{LEAK}}} \tag{5.16}$$

To evaluate the influence of the ABC parameters on the predicted optical performance, an additional set of recombination coefficients, representative of high-quality InGaN/GaN LED simulations, is considered. Figure 27 compares the IQE obtained using the parameters of Tab. 5 (solid line) and the modified parameters of Tab. 8 (dashed line) [40] [41]. The second parameter set leads to a significant improvement of the predicted IQE, mainly due to the increased SRH lifetime and radiative recombination coefficient.

This comparison highlights the strong dependence of the IQE prediction on the optical parameters adopted in the ABC model. A more accurate characterization of the radiative and non-radiative coefficients would therefore be required for a quantitative prediction of the photon generation efficiency in realistic nitride LEDs.

Table 8: Modified recombination parameters adopted in the ABC model for the $In_{0.195}Ga_{0.805}N$ quantum well layer [40] [41].

Name	**Symbol**	**$In_{0.195}Ga_{0.805}N$**	**Unit**
Electron SRH lifetime	τ_n	2.0×10^{-8}	s
Hole SRH lifetime	τ_p	2.0×10^{-8}	s
SRH trap energy level	E_{trap}	0	eV
Radiative recombination coefficient	B_{rad}	5.0×10^{-17}	m^3/s
Electron Auger coefficient	C_{nnp}	2.7×10^{-43}	m^6/s
Hole Auger coefficient	C_{ppn}	2.7×10^{-43}	m^6/s

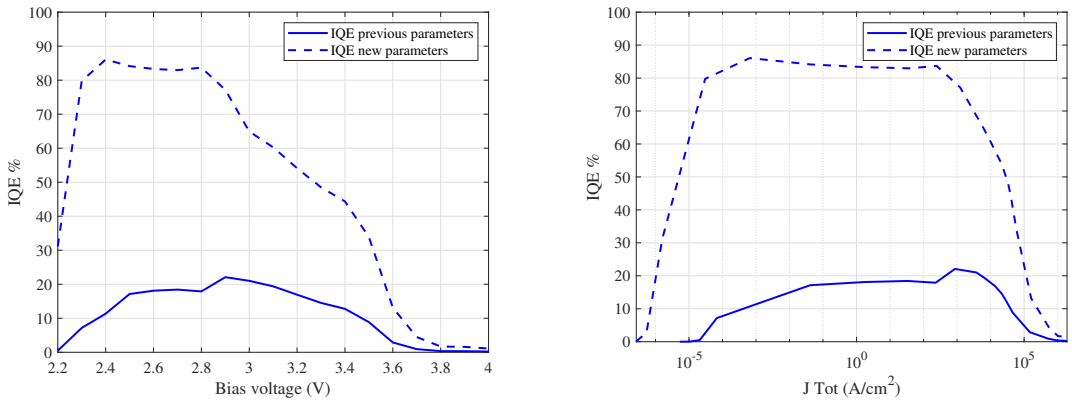


Figure 27: Internal Quantum Efficiency (IQE) of the AlInGaN-based LED obtained from frozen-field PDD–NEGF simulations using different sets of ABC recombination parameters. The comparison highlights the impact of the radiative and non-radiative coefficients on the predicted emission efficiency.

The results presented in this chapter demonstrate that the proposed PDD–NEGF approach represents an effective compromise between physical accuracy and computational efficiency. Despite the approximations introduced by the effective-mass model and, above all, by the frozen-field strategy, which removes the self-consistent Poisson solution from the NEGF iterations, the simulator preserves a detailed quantum-mechanical description of carrier transport and provides a physically consistent prediction of the recombination mechanisms and optical emission characteristics of the device.

The proposed methodology successfully captures the influence of polarization, incomplete ionization and EBL design on the transport and recombination processes, allowing a comprehensive analysis of the internal operation of the LED. Furthermore, the predicted optical power and internal quantum efficiency exhibit the expected trends and their dependence on the adopted recombination parameters, confirming the capability of the model to investigate realistic optoelectronic devices.

The reduction in computational cost achieved by the frozen-field implementation, together with the preservation of the essential physical mechanisms governing carrier transport and recombination, makes the PDD–NEGF framework a practical and reliable simulation tool for the analysis and optimization of nitride-based LEDs and, more generally, semiconductor optoelectronic devices.

6 Recursive Green's function

The recursive Green's function (RGF) technique can be regarded as an efficient numerical implementation of the Green's function formalism for layered and quasi-one-dimensional systems. The underlying idea is implicitly present in Datta's NEGF [10] framework, where the device is partitioned into a sequence of slices and the total system is progressively constructed by adding successive layers to the previously defined subsystems.

One of the main computational challenges in the implementation of the non-equilibrium Green's function (NEGF) formalism is the evaluation of the Green's functions. In particular, the retarded Green's function,

$$\underline{\underline{G}}^R(E) = \left[E\underline{\underline{M}} - \underline{\underline{H}} - \underline{\underline{\Sigma}}^R(E)\right]^{-1} \tag{6.1}$$

and the lesser Green's function, obtained through the Keldysh equation,

$$\underline{\underline{G}}^<(E) = \underline{\underline{G}}^R(E)\,\underline{\underline{\Sigma}}^<(E)\,\underline{\underline{G}}^A(E) \tag{6.2}$$

require the manipulation of large matrices for realistic device geometries. For a one-dimensional system discretized into N nodes, the direct evaluation of the retarded Green's function requires the inversion of an $N \times N$ matrix at each energy point. Since the computational complexity of matrix inversion scales as $\mathcal{O}(N^3)$, the computational effort rapidly increases with the device size, limiting the applicability of direct NEGF calculations to large systems.

The RGF algorithm overcomes this limitation by exploiting the block-tridiagonal structure of the 1D–discretized Hamiltonian, which naturally arises in layered semiconductor structures. By taking advantage of the local coupling between adjacent slices, the Green's function can be evaluated recursively without requiring the inversion of the full system matrix.

The recursive solution of both Dyson and Keldysh equations was systematically developed by Svizhenko *et al.* [42]. The method consists of a forward recursive sweep, where the influence of the left contact is propagated through the structure, followed by a backward sweep that incorporates the contribution of the right contact. This procedure provides the diagonal Green's function blocks and the required off-diagonal terms while avoiding the computational cost associated with the complete matrix inversion.

In this chapter, the RGF formalism is introduced and the corresponding numerical algorithm is described in detail. The method is initially tested on simple GaAs-based pn junction to identify the parameters affecting its computational performance, including the number of discretization nodes and the applied bias voltage. The structure is intentionally based on GaAs in order to minimize additional numerical complexities related to material properties.

The RGF algorithm is subsequently applied to a GaAs-based LED to quantify the computational speed-up with respect to the direct NEGF approach and to evaluate the accuracy loss associated with the approximation of the off-diagonal Green's function blocks.

Finally, the method is extended to an AlInGaN-based light-emitting diode to investigate its performance in wide-bandgap semiconductor devices. This case study highlights the advantages of the recursive approach for systems characterized by increased numerical stiffness, where conventional NEGF calculations become significantly more computationally demanding.

6.1 System partitioning

We introduce a formal block partitioning of the system into two coupled subsystems. For illustrative purposes, each $N \times N$ matrix is partitioned into four block matrices, leading to the following structure:

$$\left(\begin{array}{ccc|ccc} \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet & 0 & 0 \\ \hline 0 & 0 & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet & \bullet \end{array}\right) \begin{pmatrix} G_1^R & G_2^R \\ G_3^R & G_4^R \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6.3}$$

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} G_1^R & G_2^R \\ G_3^R & G_4^R \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6.4}$$

where

$$\begin{aligned} A_{i,i} &= E - H_{i,i} \\ A_{1,1} &= E - H_{1,1} - \Sigma_L \\ A_{N,N} &= E - H_{N,N} - \Sigma_R \end{aligned} \tag{6.5}$$

A_2 and A_3 represent the coupling matrices between the two subsystems, each subsystem is characterized by its isolated Green's function, denoted by g_1^R and g_4^R, while G_i^R denote the corresponding blocks of the Green's function of the fully coupled system.

$$\begin{cases} G_1^R = [(g_1^R)^{-1} - A_2 g_4^R A_3]^{-1} = [(g_1^R)^{-1} - \Sigma_4^R]^{-1} \\ G_4^R = [(g_4^R)^{-1} - A_3 g_1^R A_2]^{-1} = [(g_4^R)^{-1} - \Sigma_1^R]^{-1} \\ G_2^R = -g_1^R A_2 G_4^R \\ G_3^R = -g_4^R A_3 G_1^R \end{cases} \tag{6.6}$$

where

$$(A_1)^{-1} = g_1^R, \qquad (A_4)^{-1} = g_4^R \tag{6.7}$$

are the isolated system Green's functions. The quantities

$$\Sigma_4^R = A_2 g_4^R A_3 \tag{6.8}$$

$$\Sigma_1^R = A_3 g_1^R A_2 \tag{6.9}$$

can be interpreted as effective self-energies arising from the coupling between the two subsystems. A similar discussion can be done for the advanced Green's functions, seen as the adjoint of the Dyson equation:

$$A^\dagger G^A = \mathbb{I} \tag{6.10}$$

$$\begin{aligned} G_2^A &= -g_1^A A_3^\dagger G_4^A \\ G_3^A &= -g_4^A A_2^\dagger G_1^A \end{aligned} \tag{6.11}$$

Having established the theoretical formulation for the partitioning of a general system into two

coupled subsystems, we now specialize the discussion to the single-node implementation of the RGF algorithm.

6.2 Recursive retarded Green's function algorithm

Regardless of the selected node, the previous formulations require the inversion of large submatrices of A. Since these inversions dominate the computational cost, a different strategy is adopted in the RGF algorithm. The device is divided into elementary slices, each corresponding to a node of the one-dimensional discretization, and the Green's function is constructed recursively by progressively enlarging the subsystem. The first subsystem consists of the leftmost slice, whose Green's function already includes the effect of the left contact through the self-energy Σ_L^R. The subsystem is then enlarged by adding one slice at a time. At each recursion step, only the Green's function of the newly added slice is evaluated, yielding the *left-connected Green's function* $(g_L^R)_{q+1,q+1}$ (where q refers to the last node of the first subsystem). This quantity is exact for the subsystem extending from the left contact up to slice $q+1$, but it does not yet contain the influence of the remaining right-hand portion of the device. Consequently, the forward recursion propagates the left boundary condition through the structure, while the effect of the right contact is incorporated later by the backward recursion:

first subsystem partitioning ($q = 1$)

$$\underline{\underline{A}} = \begin{pmatrix} \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet & \bullet \end{pmatrix} \tag{6.12}$$

The first subsystem for each added layer consists of the first q slices of the device (red block), while the second subsystem is formed by the newly added slice $q+1$ node (black block). The matrices $A_{1:q,q+1}$ and $A_{q+1,1:q}$ (blue blocks)[2] describe the coupling between the two subsystems. In the present one-dimensional discretization, each slice contains a single degree of freedom, so the coupling between the two subsystems is described by scalar quantities. Consequently, the inversion of $A_{1:q,1:q}$ reduces to the inversion of a small matrix (or a scalar if $q = 1$). The left-connected Green's function $(g_L^R)_{q+1,q+1}$, from the system 6.6, is:

$$(g_L^R)_{q+1,q+1} = (A_{q+1,q+1} - A_{q+1,1:q} A_{1:q,1:q}^{-1} A_{1:q,q+1})^{-1} \tag{6.13}$$

where $A_{q+1,q+1}$ is a scalar quantity and $A_{1:q,1:q}^{-1}$ is the Green's function of the left-connected red block system $(g_L^R)_{1:q,1:q}$. Eq. 6.13 describes how the Green's function of the enlarged subsystem is obtained from that of the previous subsystem by incorporating the influence of the left-connected part through the effective self-energy $A_{q+1,1:q} A_{1:q,1:q}^{-1} A_{1:q,q+1}$. For this reason, the evaluation of $(g_L^R)_{q+1,q+1}$ is referred to as *forward propagation* and is solved iteratively for all the nodes till the

[2] Comparing with Eq. 6.3, the matrices $A_{1:q,q+1}$ and $A_{q+1,1:q}$ correspond respectively to A_2 and A_3, but defined for a reduced partitioning of the system. Instead of considering the entire device, the system is divided into slices, and the same partitioning is applied consistently to the coupling matrices.

last one.

The forward coupling matrices $A_{1:q,q+1}$ and $A_{q+1,1:q}$, represented with blue blocks, are clearly sparse vectors with only one non-zero element corresponding to the interface sites between the two subsystems, resulting from the hypothesis of tridiagonality of the Hamiltonian. As a consequence, the matrix product in Eq. 6.13 reduces to scalar quantities involving only the surface Green's function element of the corresponding left-connected subsystem matrix $(g_L^R)_{1:q,1:q}$:

$$\boxed{(g_L^R)_{q+1,q+1} = (A_{q+1,q+1} - A_{q+1,q}(g_L^R)_{q,q}A_{q,q+1})^{-1}} \tag{6.14}$$

This property significantly reduces computational cost, since only a node–to–node propagation is involved. In the implementation, the recursion starts from $q = 1$, ensuring that no full matrix inversion is required at any step of the algorithm.

The forward recursion alone does not provide the complete diagonal Green's function of the device, since each left-connected Green's function only accounts for the influence of the left contact and of the slices already incorporated into the subsystem. An exception is the last slice, whose Hamiltonian also contains the right-contact self-energy Σ_R^R. Consequently, once the forward recursion reaches the end of the device, the left- and right-contact boundary conditions are both included, and the last diagonal element coincides with that of the fully coupled system,

$$G_{N,N}^R = (g_L^R)_{N,N} \tag{6.15}$$

and to recover the full system Green's function, a backward propagation must be performed exploiting the complementary expression of G_1^R from the system 6.6:

first subsystem partitioning backward ($q = N$)

$$\underline{\underline{A}} = \begin{pmatrix} \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & \bullet & \boxed{\overset{q-1}{\bullet}} & \boxed{\bullet} \\ 0 & 0 & 0 & 0 & \boxed{\bullet} & \boxed{\overset{q}{\bullet}} \end{pmatrix} \tag{6.16}$$

$$G_{q-1,q-1}^R = (A_{q-1,q-1} - A_{q-1,q:N}A_{q:N,q:N}^{-1}A_{q:N,q-1})^{-1} \tag{6.17}$$

which propagates the influence of the right contact toward the left, where q labels the second subsystem propagating toward the first one at the interface node $(q-1, q-1)$. However, the matrices appearing in this expression do not yet incorporate the left-contact propagation effects generated during the forward recursion, and the evaluation of the term in parentheses would still require the inversion of an effective block associated with the right subsystem. The effect of the left contact can be introduced as perturbation of the right propagation wave through the *Woodbury identity*, which states for generic matrices:

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} \tag{6.18}$$

The *backward propagation* can be found starting from Eq. 6.17, which becomes:

$$G^R_{q-1,q-1} = A^{-1}_{q-1,q-1} + A^{-1}_{q-1,q-1}A_{q-1,q:N}(A_{q:N,q:N} - A_{q:N,q-1}A^{-1}_{q-1,q-1}A_{q-1,V})^{-1}A_{q:N,q-1}A^{-1}_{q-1,q-1} \tag{6.19}$$

$$\overset{\text{Eq. 6.14}}{=} A^{-1}_{q-1,q-1} + A^{-1}_{q-1,q-1}A_{q-1,q:N}(G^R)_{q:N,q:N}A_{q-1,q:N}A^{-1}_{q-1,q-1} \tag{6.20}$$

$$G^R_{q-1,q-1} = A^{-1}_{q-1,q-1} + A^{-1}_{q-1,q-1}A_{q-1,q:N}(G^R)_{q:N,q:N}A_{q:N,q-1}\,A^{-1}_{q-1,q-1} \tag{6.21}$$

where $(G^R)_{q:N,q:N}$ is the Green's function of the red block in Eq. 6.17.

This expression updates the Green's function of the left subsystem by incorporating the effect of the right subsystem through the forward–propagated Green's function matrix $(g^R_L)_{q,q}$. Consequently, once the forward recursion has reached the last slice, the backward recursion propagates the influence of the right contact toward the left, yielding the complete diagonal Green's function of the device.

Notice that the matrix product $A_{q-1,q:N}(G^R)_{q:N,q:N}A_{q:N,q-1}$ acts as an effective self-energy, as already observed during the forward propagation. Owing to the nearest-neighbor coupling introduced by the finite-difference discretization, the coupling matrices $A_{q-1,q:N}$ and $A_{q:N,q-1}$ reduce to sparse vectors containing a single non-zero element. Consequently, the matrix product simplifies to a scalar operation involving only the interface element of the Green's function block $(G^R)_{q:N,q:N}$, namely $(G^R)_{q,q}$:

$$G^R_{q-1,q-1} = A^{-1}_{q-1,q-1} + A^{-1}_{q-1,q-1}A_{q-1,q}(G^R)_{q,q}A_{q,q-1}\,A^{-1}_{q-1,q-1} \tag{6.22}$$

and using the Dyson equation:

$$\boxed{G^R_{q-1,q-1} = (g^R_L)_{q-1,q-1} + (g^R_L)_{q-1,q-1}A_{q-1,q}(G^R)_{q,q}A_{q,q-1}\,(g^R_L)_{q-1,q-1}} \tag{6.23}$$

The Dyson equation can be solved recursively, significantly reducing the computational cost of the simulations. In the conventional NEGF approach, the Green's function is obtained through direct full-matrix inversion, whose computational complexity scales as $\mathcal{O}(N^3)$. By contrast, the RGF algorithm exploits the block-tridiagonal structure of the Hamiltonian and recursively computes the required Green's-function blocks, reducing the computational complexity to $\mathcal{O}(N^2)$ for the diagonal Green's function.

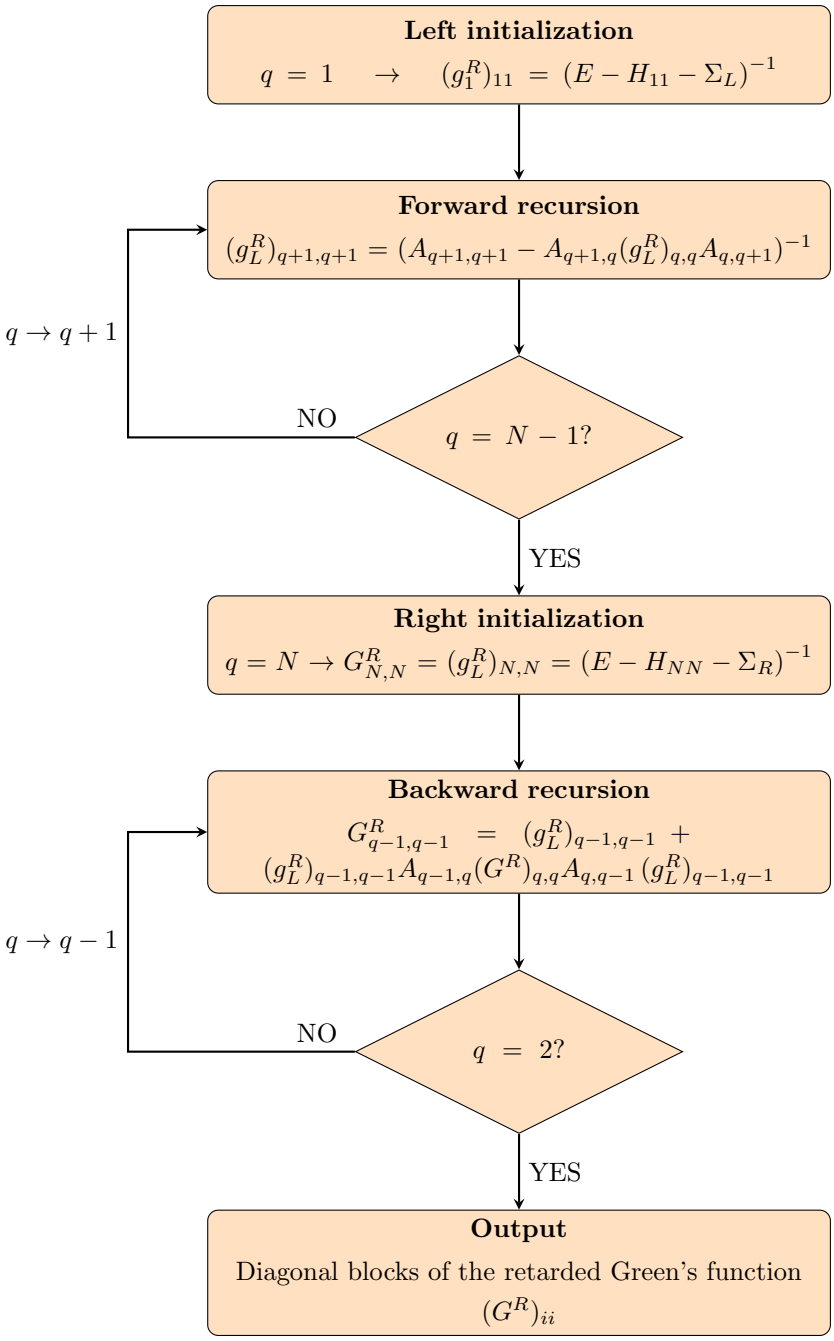


Figure 28: Flowchart of the recursive Green's function (RGF) algorithm used to compute the diagonal blocks of the retarded Green's function. The procedure consists of a forward recursion, in which the left-connected Green's functions are recursively evaluated from the source contact to the drain contact, followed by a backward recursion that reconstructs the diagonal blocks of the full retarded Green's function.

6.3 Recursive lesser Green's function algorithm

The evaluation of physical observables within the NEGF formalism requires the computation of the lesser Green's function through the Keldysh equation. Following the subsystem organization of Eq. 6.3:

$$\begin{pmatrix} G_1^< & G_2^< \\ G_3^< & G_4^< \end{pmatrix} = \begin{pmatrix} G_1^R & G_2^R \\ G_3^R & G_4^R \end{pmatrix} \begin{pmatrix} \Sigma_1^< & \Sigma_2^< \\ \Sigma_3^< & \Sigma_4^< \end{pmatrix} \begin{pmatrix} G_1^A & G_2^A \\ G_3^A & G_4^A \end{pmatrix} \tag{6.24}$$

by solving the matrix product above, the expression for the lesser Green's functions yields:

$$G_1^< = G_1^R\Sigma_1^<G_1^A + G_2^R\Sigma_3^<G_1^A + G_1^R\Sigma_2^<G_3^A + G_2^R\Sigma_4^<G_3^A \tag{6.25}$$

$$G_4^< = G_3^R\Sigma_1^<G_2^A + G_4^R\Sigma_3^<G_2^A + G_3^R\Sigma_2^<G_4^A + G_4^R\Sigma_4^<G_4^A \tag{6.26}$$

Taking into account the matrix construction of Eq. 6.12, the next step is to formulate the forward propagation of the left contact towards the right one, and, from the Keldysh equation, the *left-connected lesser Green's function* for the first slice is defined as:

$$g_{q,q}^< = g_{q,q}^R\Sigma_{q,q}^<g_{q,q}^A \tag{6.27}$$

where $q = 1$ according to our choice of selecting one single node for the first slice. By recalling the last two equations from the system 6.6 and the respective advanced Green's functions, written with the q index notation

$$g_{1:q,q+1}^R = -(g_L^R)_{1:q,1:q}\, A_{1:q,q+1}(g_L^R)_{q+1,q+1}, \quad g_{1:q,q+1}^A = (g_L^A)_{1:q,1:q}A_{1:q,q+1}^\dagger(g_L^A)_{q+1,q+1} \tag{6.28}$$

$$g_{q+1,1:q}^R = -(g_L^R)_{q+1,q+1}\, A_{q+1,1:q}\,(g_L^R)_{1:q,1:q}, \quad g_{q+1,1:q}^A = -(g_L^A)_{q+1,q+1}\, A_{q+1,1:q}^\dagger\,(g_L^A)_{1:q,1:q} \tag{6.29}$$

Now, we find the definition for the second subsystem left-connected lesser Green's function, starting from Eq. 6.26 adapted to the matrix notation:

$$\begin{aligned}(g_L^<)_{q+1,q+1} =& (g^R)_{q+1,1:q}\,\Sigma_{1:q,1:q}^<\,(g^A)_{1:q,q+1} \\ &+ (g_L^R)_{q+1,q+1}\,\Sigma_{q+1,1:q}^<\,(g^A)_{1:q,q+1} \\ &+ (g^R)_{q+1,1:q}\,\Sigma_{1:q,q+1}^<\,(g_L^A)_{q+1,q+1} \\ &+ (g_L^R)_{q+1,q+1}\,\Sigma_{q+1,q+1}^<\,(g_L^A)_{q+1,q+1}\end{aligned} \tag{6.30}$$

$$\begin{aligned}(g_L^<)_{q+1,q+1} \overset{\substack{\text{Eq. 6.28}\\ \text{Eq. 6.29}}}{=}& (g_L^R)_{q+1,q+1}^R A_{q+1,1:q}(g_L^R)_{1:q,1:q}\Sigma_{1:q,1:q}(g_L^A)_{1:q,1:q}A_{1:q,q+1}^\dagger(g_L^A)_{q+1,q+1} \\ &- (g_L^R)_{q+1,q+1}\Sigma_{q+1,1:q}(g_L^A)_{1:q,1:q}A_{1:q,q+1}^\dagger(g_L^A)_{q+1,q+1} \\ &- g_{q+1,q+1}^R A_{q+1,1:q}(g_L^R)_{1:q,1:q}\Sigma_{1:q,q+1}(g_L^A)_{q+1,q+1} \\ &+ (g_L^R)_{q+1,q+1}\Sigma_{q+1,q+1}(g_L^A)_{q+1,q+1}\end{aligned} \tag{6.31}$$

$$\begin{aligned}(g_L^<)_{q+1,q+1} =& (g_L^R)_{q+1,q+1}\times \\ &[A_{q+1,1:q}(g_L^<)_{q,q}A_{1:q,q+1}^\dagger - \Sigma_{q+1,1:q}^<(g_L^A)_{q,q}A_{1:q,q+1}^\dagger \\ &- A_{q+1,1:q}(g_L^R)_{q,q}\Sigma_{1:q,q+1}^< + \Sigma_{q+1,q+1}^<] \\ &\times (g_L^A)_{q+1,q+1}\end{aligned} \tag{6.32}$$

In the forward propagation, the four contributions inside the square brackets have distinct physical interpretations. The term

$$A_{q+1,1:q}(g_L^<)_{q,q}A^\dagger_{1:q,q+1}$$

describes the propagation of the non-equilibrium correlation function already established in the first subsystem into the newly added subsystem through the coupling matrices $A_{q+1,1:q}$ and $A^\dagger_{1:q,q+1}$. The terms

$$\Sigma^<_{q+1,1:q}(g_L^A)_{q,q}A^\dagger_{1:q,q+1}$$

and

$$A_{q+1,1:q}(g_L^R)_{q,q}\Sigma^<_{1:q,q+1}$$

represent the contribution associated with carrier injection mediated by the off-diagonal blocks of the lesser self-energy, describing the coupling between the two subsystems through non-local scattering or contact processes. Finally, the term

$$\Sigma^<_{q+1,q+1}$$

corresponds to the local source contribution of the newly added subsystem, accounting for carrier injection generated directly within that region, for example by contacts or local scattering mechanisms.

The sparse structure of the coupling matrices A significantly simplifies each matrix product involving $A_{q+1,1:q}$ and $A^\dagger_{1:q,q+1}$. In particular, these operations involve only the non-zero corner blocks associated with the coupling between adjacent subsystems, thereby reducing the number of matrix operations required at each recursive step. Moreover, within the general assumptions of the present NEGF model, the self-energy matrices are assumed to have a tridiagonal structure. Consequently, the lesser self-energy $\Sigma^<$ is also at most tridiagonal, further limiting the number of non-zero off-diagonal contributions entering the recursive propagation.

Therefore, Eq. 6.32 can be written in the simplified form:

$$\begin{aligned}(g_L^<)_{q+1,q+1} =&(g_L^R)_{q+1,q+1}\times \\ &[A_{q+1,q}(g_L^<)_{q,q}A^\dagger_{q,q+1} - \Sigma^<_{q+1,q}(g_L^A)_{q,q}A^\dagger_{q,q+1} \\ &- A_{q+1,q}(g_L^R)_{q,q}\Sigma^<_{q,q+1} + \Sigma^<_{q+1,q+1}] \\ &\times (g_L^A)_{q+1,q+1}\end{aligned} \tag{6.33}$$

Now, we turn our attention to the backward propagation, whose purpose is to reconstruct the full lesser Green's function once the forward recursion has been completed. Starting from Eq. 6.25, and recalling Eq. 6.28 and Eq. 6.29, the backward propagation of the Keldysh equation is given by:

$$
\begin{aligned}
(G^<)_{q-1,q-1} = (g_L^<)_{q-1,q-1} \\
&- (g_L^R)_{q-1,q-1} A_{q-1,q:N} (g_L^<)_{q:N,q-1} \\
&- (g_L^<)_{q-1,q:N} A^\dagger_{q-1,q:N} (g_L^A)_{q-1,q-1} \\
&+ (g^<)_{q-1,q-1} A^\dagger_{q:N,q-1} G^A_{q-1,q:N} \\
&+ G^R_{q-1,q:N} A_{q:N,q-1} (g_L^<)_{q-1,q-1} \\
&+ (g_L^R)_{q-1,q-1} A_{q-1,q:N} \, (G^<)_{q:N,q:N} \, A^\dagger_{q-1,q:N} (g_L^A)_{q-1,q-1}
\end{aligned}
\tag{6.34}
$$

where we adopted the following short-notations:

$$(g_L^<)_{q,q-1} = G^R_{q,q} \Sigma^<_{q,q-1} (g_L^A)_{q-1,q-1} \tag{6.35}$$

$$(g_L^<)_{q-1,q} = (g_L^R)_{q-1,q-1} \Sigma^<_{q-1,q} G^A_{q,q} \tag{6.36}$$

$$\tag{6.37}$$

By exploiting the sparsity of the coupling matrices and simplifying the resulting matrix products into scalar operations, the expression becomes:

$$
\boxed{
\begin{aligned}
(G^<)_{q-1,q-1} = (g_L^<)_{q-1,q-1} \\
&+ (g_L^R)_{q-1,q-1} A_{q-1,q} (g_L^<)_{q,q-1} + (g_L^<)_{q-1,q} A^\dagger_{q-1,q} (g_L^A)_{q-1,q-1} \\
&+ (g^<)_{q-1,q-1} A^\dagger_{q,q-1} G^A_{q-1,q} + G^R_{q-1,q} A_{q,q-1} (g_L^<)_{q-1,q-1} \\
&+ (g_L^R)_{q-1,q-1} A_{q-1,q} \, (G^<)_{q,q} \, A^\dagger_{q-1,q} (g_L^A)_{q-1,q-1}
\end{aligned}
}
\tag{6.38}
$$

The off-diagonal terms of the lesser Green's function matrix, since the evaluation of the current and other physical observables requires a complete description of the system, are determined during the backward propagation as:

$$
\begin{aligned}
G^<_{q-1,q} &= G^R_{q-1,q-1} \Sigma^<_{q-1,q-1} G^A_{q-1,q} + G^R_{q-1,q} \Sigma^<_{q,q} G^A_{q,q} \\
&\quad + G^R_{q-1,q-1} \Sigma^<_{q-1,q} G^A_{q,q} + G^R_{q-1,q} \Sigma^<_{q,q-1} G^A_{q,q+1} \\
G^<_{q,q} &= G^R_{q,q-1} \Sigma^<_{q-1,q-1} G^A_{q-1,q-1} + G^R_{q,q} \Sigma^<_{q,q-1} G^A_{q-1,q-1} \\
&\quad + G^R_{q,q-1} \Sigma^<_{q-1,q} G^A_{q,q-1} + G^R_{q,q} \Sigma^<_{q,q} G^A_{q,q-1}
\end{aligned}
\tag{6.39}
$$

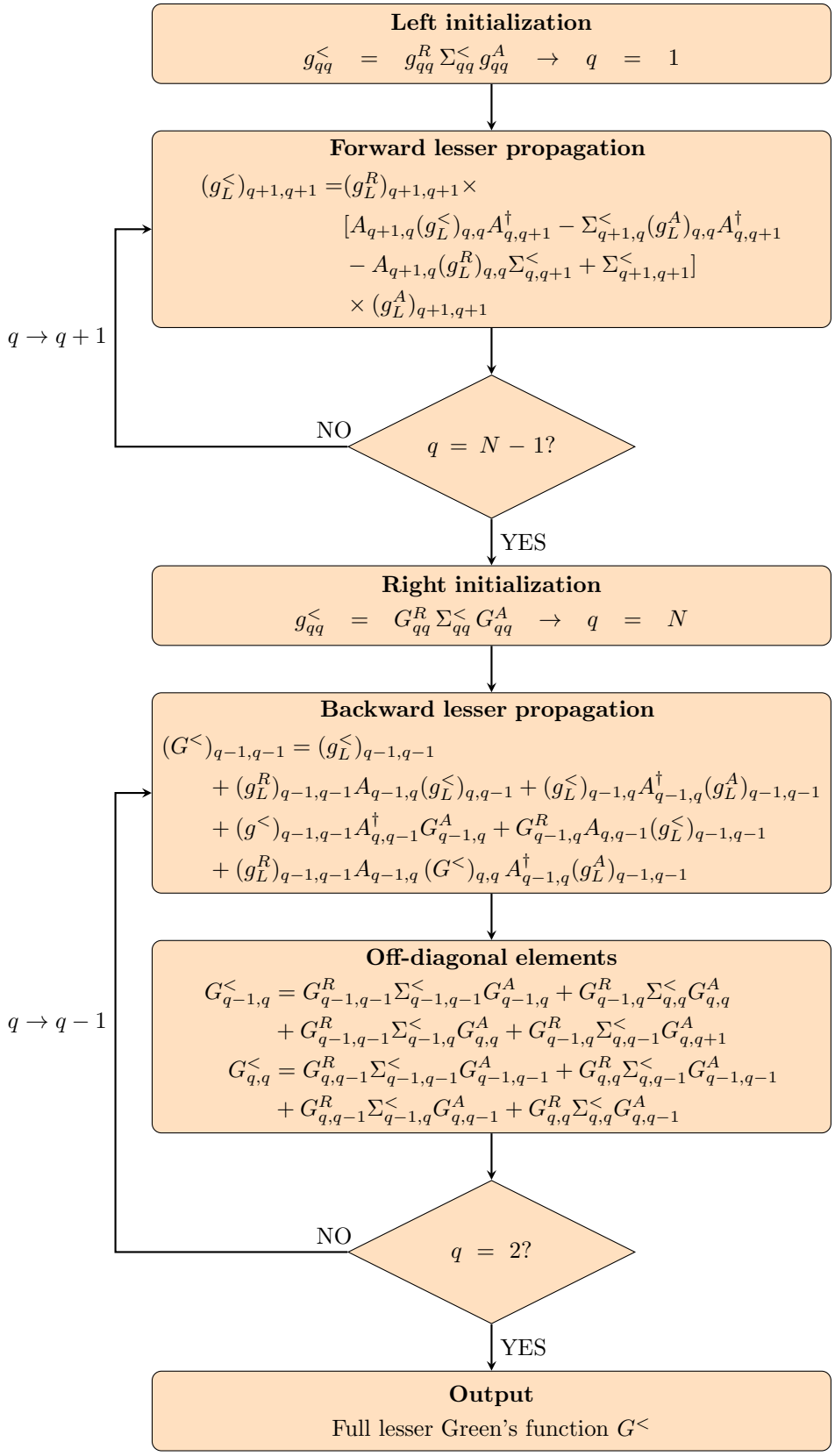


Figure 29: Recursive evaluation scheme of the lesser Green's function using the recursive green's function (RGF) technique. The algorithm proceeds through a forward recursion to build the left-connected Green's functions, followed by a backward sweep to reconstruct the full lesser Green's function. The procedure ensures efficient evaluation of $G^<$ by exploiting the tridiagonal structure of the Hamiltonian and is fully compatible with contact self-energies and scattering terms.

6.4 Büttiker Probes in the Recursive Green's Function Algorithm

An important feature of the RGF formulation is that it naturally enables a recursive evaluation of the derivatives entering the physical observables and the Newton–Raphson Jacobian. As discussed in Chapter 4.7, their direct evaluation involves computationally expensive matrix products and inversions over the full device Hamiltonian. By exploiting the recursive structure of the RGF algorithm, these derivatives can instead be evaluated recursively by operating on the local tridiagonal blocks, avoiding the resolution of the full matrix [43]. This significantly reduces both the computational cost and the overall simulation time, while preserving the accuracy of the solution.

In the following, the recursive expressions for the derivatives of the retarded (advanced) and lesser (greater) Green's functions with respect to the electrostatic potential are derived. The same recursive procedure can be straightforwardly extended to the derivatives with respect to the Büttiker probe quasi-Fermi levels.

The recursive evaluation starts from the derivative of the left-connected retarded Green's function. Recalling Eq. 6.14, one obtains

$$\frac{\partial (g_L^R)_{q+1,q+1}}{\partial U} = -(g_L^R)^2_{q+1,q+1}\left[\frac{\partial A_{q+1,q+1}}{\partial U} - A_{q+1,q}\frac{\partial (g_L^R)_{q,q}}{\partial U}A_{q,q+1}\right]. \tag{6.40}$$

The recursive relation above constitutes the forward-propagation step for the derivative of the retarded Green's function. The corresponding backward recursion, which yields the derivative of the diagonal block of the full retarded Green's function, is obtained by differentiating Eq. 6.23:

$$\begin{aligned}\frac{\partial G^R_{q-1,q-1}}{\partial U} = &\frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} \\ &+ 2\,\frac{\partial (g_L^R)_{q-1,q-1}}{\partial U}(g_L^R)_{q-1,q-1}A_{q-1,q}(G^R)_{q,q}A_{q,q-1} \\ &+ (g_L^R)^2_{q-1,q-1}A_{q-1,q}\frac{\partial (G^R)_{q,q}}{\partial U}A_{q,q-1}.\end{aligned} \tag{6.41}$$

Since the recursive formulation is applied to a single-node subsystem, the Green's function blocks reduce to scalar quantities, allowing the ordering of the terms in the matrix products to be simplified. The derivative of the advanced Green's function follows the same procedure by exploiting the Hermitian relation between the retarded and advanced Green's functions.

In a similar manner, the evaluation of the lesser Green's function derivative starts from the derivative of the left-connected lesser Green's function:

$$\begin{aligned}\frac{\partial (g_L^<)_{q+1,q+1}}{\partial U} = &\Big[|A_{q+1,q}|^2(g_L^<)_{q+1,q+1} - \Sigma^<_{q+1,q}(g_L^A)_{q,q}A^\dagger_{q,q+1} - \\ &A_{q+1,q}(g_L^R)_{q,q}\Sigma^<_{q,q+1} + \Sigma^<_{q+1,q+1}\Big] \times \\ &\left[\frac{\partial (g_L^R)_{q+1,q+1}}{\partial U}(g_L^A)_{q+1,q+1} + \frac{\partial (g_L^A)_{q+1,q+1}}{\partial U}(g_L^R)_{q+1,q+1}\right] \\ &+ (g_L^R)_{q+1,q+1}\left[|A_{q+1,q}|^2\frac{\partial (g_L^<)_{q+1,q+1}}{\partial U} - \Sigma^<_{q+1,q}\frac{\partial (g_L^A)_{q,q}}{\partial U}A^\dagger_{q,q+1}\right. \\ &\left. - A_{q+1,q}\frac{\partial (g_L^R)_{q,q}}{\partial U}\Sigma^<_{q,q+1} + \frac{\partial \Sigma^<_{q+1,q+1}}{\partial U}\right](g_L^A)_{q+1,q+1}\end{aligned} \tag{6.42}$$

The diagonal structure of the lesser self-energies significantly simplifies the resulting expression. Furthermore, by exploiting the Hermitian relation between the retarded and advanced Green's

functions,

$$(g_L^A)_{q+1,q+1} = \left[(g_L^R)_{q+1,q+1}\right]^\dagger,$$

together with

$$\frac{\partial (g_L^A)_{q+1,q+1}}{\partial U} = \left[\frac{\partial (g_L^R)_{q+1,q+1}}{\partial U}\right]^\dagger,$$

the final recursive expression for the derivative of the left-propagated lesser Green's function can be obtained as:

$$\begin{aligned}\frac{\partial (g_L^<)_{q+1,q+1}}{\partial U} = &\left[|A_{q+1,q}|^2 (g_L^<)_{q+1,q+1} + \Sigma^<_{q+1,q+1}\right] \times \\ &\left[2\mathrm{Re}\left\{\frac{\partial (g_L^R)_{q+1,q+1}}{\partial U}(g_L^A)_{q+1,q+1}\right\}\right] \\ &+ |(g_L^R)_{q+1,q+1}|^2 \left[|A_{q+1,q}|^2 \frac{\partial (g_L^<)_{q+1,q+1}}{\partial U} + \frac{\partial \Sigma^<_{q+1,q+1}}{\partial U}\right]\end{aligned} \tag{6.43}$$

The derivative of the lesser Green's function of the full system, obtained by recalling Eq. 6.38, is evaluated according to the following recursive expression:

$$\begin{aligned}\frac{\partial (G^<)_{q-1,q-1}}{\partial U} = &\frac{\partial (g_L^<)_{q-1,q-1}}{\partial U} \\ &+ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} A^\dagger_{q,q-1} (G^A)_{q-1,q} + (g_L^R)_{q-1,q-1} A^\dagger_{q,q-1} \frac{\partial (G^A)_{q-1,q}}{\partial U} \\ &+ \frac{\partial (G^R)_{q-1,q}}{\partial U} A_{q,q-1} (g_L^<)_{q-1,q-1} + (G^R)_{q-1,q} A_{q,q-1} \frac{\partial (g_L^<)_{q-1,q-1}}{\partial U} \\ &+ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} A_{q-1,q} (G^<)_{q,q} A^\dagger_{q,q-1} (g_L^A)_{q-1,q-1} \\ &+ (g_L^R)_{q-1,q-1} A_{q-1,q} \frac{\partial (G^<)_{q,q}}{\partial U} A^\dagger_{q,q-1} (g_L^A)_{q-1,q-1} \\ &+ (g_L^R)_{q-1,q-1} A_{q-1,q} (G^<)_{q,q} A^\dagger_{q,q-1} \frac{\partial (g_L^A)_{q-1,q-1}}{\partial U}\end{aligned} \tag{6.44}$$

where the off-diagonal term $g_L^<$ and its partial derivative have been set to zero since the $\Sigma^<$ is a diagonal matrix in the Büttiker-probe technique. By exploiting the Hermitian relation between the retarded and advanced Green's functions,

$$G^A = (G^R)^\dagger,$$

together with the anti-Hermitian property of the lesser Green's function,

$$G^{<\dagger} = -G^<,$$

the conjugate terms appearing in the recursive derivative expression can be combined. Therefore, the resulting expression can be rewritten in terms of the imaginary part of a single matrix product:

$$\begin{aligned}\frac{\partial (G^<)_{q-1,q-1}}{\partial U} &= \frac{\partial (g_L^<)_{q-1,q-1}}{\partial U} \\ &+ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} A^\dagger_{q,q-1} (G^A)_{q-1,q} + (g_L^R)_{q-1,q-1} A^\dagger_{q,q-1} \frac{\partial (G^A)_{q-1,q}}{\partial U} \\ &+ \frac{\partial (G^R)_{q-1,q}}{\partial U} A_{q,q-1} (g_L^<)_{q-1,q-1} + (G^R)_{q-1,q} A_{q,q-1} \frac{\partial (g_L^<)_{q-1,q-1}}{\partial U} \\ &+ 2\mathrm{Re}\left\{ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} (g_L^A)_{q-1,q-1} \right\} |A_{q-1,q}|^2 \, (G^<)_{q,q} \\ &+ |(g_L^R)_{q-1,q-1}|^2 |A_{q-1,q}|^2 \, \frac{\partial (G^<)_{q,q}}{\partial U}\end{aligned} \tag{6.45}$$

$$\begin{aligned}\frac{\partial (G^<)_{q-1,q-1}}{\partial U} &= \frac{\partial (g_L^<)_{q-1,q-1}}{\partial U} \\ &- 2\mathrm{Re}\left\{ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} A^\dagger_{q,q-1} (G^A)_{q-1,q} \right\} \\ &- 2\mathrm{Re}\left\{ (g_L^R)_{q-1,q-1} A^\dagger_{q,q-1} \frac{\partial (G^A)_{q-1,q}}{\partial U} \right\} \\ &+ 2\mathrm{Re}\left\{ \frac{\partial (g_L^R)_{q-1,q-1}}{\partial U} (g_L^A)_{q-1,q-1} \right\} |A_{q-1,q}|^2 \, (G^<)_{q,q} \\ &+ |(g_L^R)_{q-1,q-1}|^2 |A_{q-1,q}|^2 \, \frac{\partial (G^<)_{q,q}}{\partial U}\end{aligned} \tag{6.46}$$

The derivative of the greater Green's function follows an analogous procedure, obtained by replacing the lesser self-energy with the corresponding greater self-energy.

The final step consists of evaluating the derivatives of the off-diagonal components of the lesser Green's function:

$$G^<_{q-1,q} = -G^R_{q,q} A_{q,q-1} g^<_{q-1,q-1} - G^<_{q,q} A^\dagger_{q-1,q} (g_L^A)_{q-1,q-1} \tag{6.47}$$

$$\begin{aligned}\frac{\partial G^<_{q-1,q}}{\partial U} &= -A_{q,q-1} \left[\frac{\partial G^R_{q,q}}{\partial U} g^<_{q-1,q-1} - G^R_{q,q} \frac{\partial g^<_{q-1,q-1}}{\partial U} \right] \\ &- A^\dagger_{q-1,q} \left[\frac{\partial G^<_{q,q}}{\partial U} (g_L^A)_{q-1,q-1} - G^<_{q,q} \frac{\partial (g_L^A)_{q-1,q-1}}{\partial U} \right]\end{aligned} \tag{6.48}$$

The derivatives evaluated so far can be directly extended to derivatives with respect to the Fermi energies of the contacts and Büttiker probes. The only difference with respect to the electrostatic potential derivative is that the device Hamiltonian is independent of the Fermi energies. Consequently, all terms involving the derivative of the Hamiltonian vanish, leading to a simplified recursive formulation of the Green's function derivatives:

$$\frac{\partial h_D}{\partial \mu_N} = 0. \tag{6.49}$$

The recursive formulation of the Green's function derivatives allows the simplification of the expressions required for the evaluation of physical observables and the construction of the residual Jacobians:

- The derivative of the spectral electron density, expressed as

$$n(E,q) = -\sigma_{\text{spin}}\text{Re}\left\{iG^{<}_{q,q}(E)\right\}, \tag{6.50}$$

is obtained as

$$\frac{\partial n(E,q)}{\partial U} = -\sigma_{\text{spin}}\text{Re}\left\{i\frac{\partial G^{<}_{q,q}(E)}{\partial U}\right\}. \tag{6.51}$$

The same procedure is applied to the hole density by replacing the lesser Green's function with the corresponding greater Green's function.

- The derivative of the spectral current divergence residual, defined as

$$r_e(E,q) = \frac{q}{h}\sigma_{\text{spin}}\left[2\text{Re}\left\{h_{q,q+1}G^{<}_{q+1,q} - h_{q-1,q}G^{<}_{q,q-1}\right\}\right], \tag{6.52}$$

is given by

$$\frac{\partial r_e(E,q)}{\partial U} = \frac{q}{h}\sigma_{\text{spin}}\left[2\text{Re}\left\{h_{q,q+1}\frac{\partial G^{<}_{q+1,q}(E)}{\partial U} - h_{q-1,q}\frac{\partial G^{<}_{q,q-1}(E)}{\partial U}\right\}\right]. \tag{6.53}$$

The implementation of the derivative scheme completes the basic structure of the RGF algorithm. After implementing the equations described in the previous sections, the simulator can be tested on simple structures to validate its numerical behaviour. The analysis begins with GaAs-based p–n junctions, which provide a controlled environment for evaluating the performance of the recursive approach and identifying the main parameters affecting its computational efficiency.

6.5 RGF Validation on a GaAs p–n Junction

The RGF algorithm is first validated on a GaAs-based p–n junction, which provides a representative test case involving the simultaneous transport of electrons and holes. This structure enables the performance of the recursive approach to be assessed under bipolar transport conditions and, in particular, allows the computational efficiency of the RGF method to be evaluated when the Newton–Raphson cycle requires the self-consistent treatment of both carrier types. The analysis aims to determine whether the RGF approach becomes more computationally efficient than the conventional approach, which relies on the direct inversion of matrices to solve the Dyson and Keldysh equations.

The simulated device consists of two regions associated with the ground and bias contacts, respectively, as reported in Tab. 9. To investigate the scaling of the computational cost with device size, the thickness of both regions is progressively increased, resulting in a corresponding increase in the total number of spatial discretization nodes.

The thickness of each region is varied from 10 nm to 30 nm, while the applied forward bias ranges from 0.1 V to 0.5 V. Complete dopant ionization is assumed throughout the simulations, as the relatively low dopant activation energies in GaAs lead to nearly complete ionization at room temperature. A uniform spatial discretization step of 0.25 nm is employed in all simulations.

To quantify the computational performance difference between the RGF (T_{RGF}) and conventional NEGF (T_{classic}) approaches, two metrics are introduced:

- *speed–up time*, defined as

$$\%T_{\text{speed-up}} = \frac{T_{\text{classic}}}{T_{\text{RGF}}} \times 100 \tag{6.54}$$

Table 9: Layer structure of the GaAs-based p–n junction adopted for the RGF performance analysis.

Layer	**Material**	**Thickness (nm)**	**Type**	**Doping (cm^{-3})**
Bias contact	GaAs	L	p	1×10^{19}
Ground contact	GaAs	L	n	1×10^{19}

which quantifies how much time faster the RGF method is compared with the conventional NEGF approach;

- *time saving*, defined as

$$\%T_{\mathrm{saving}} = \left(1 - \frac{T_{\mathrm{RGF}}}{T_{\mathrm{classic}}}\right) \times 100 \tag{6.55}$$

which represents the percentage reduction in computational time achieved by employing the RGF method instead of the conventional NEGF approach.

Equilibrium Case

As shown in Fig. 30, the electrostatic potential and Green's function updates obtained with the RGF method are consistent with those of the conventional NEGF approach, while the recursive formulation requires significantly less computational time.

The dependence of the computational savings on the device size is further investigated in Fig. 31. The RGF approach provides a computational time reduction ranging from approximately 40% to 60%, with the computational advantage increasing as the number of discretization nodes increases. This result confirms the scalability of the recursive approach also for a structure involving both electron and hole transport.

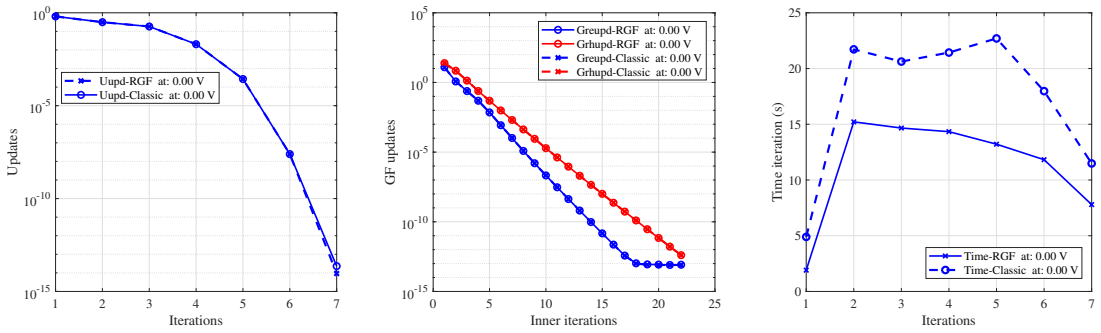


Figure 30: Comparison between the conventional NEGF and recursive Green's function (RGF) approaches under thermodynamic equilibrium conditions for a GaAs-based p–n junction. The left and central panels show the close agreement between the conventional NEGF and RGF results for the device with 10 nm-thick layers. The right panel highlights the computational advantage of the recursive approach, which substantially reduces the execution time while maintaining the accuracy of the conventional NEGF solution.

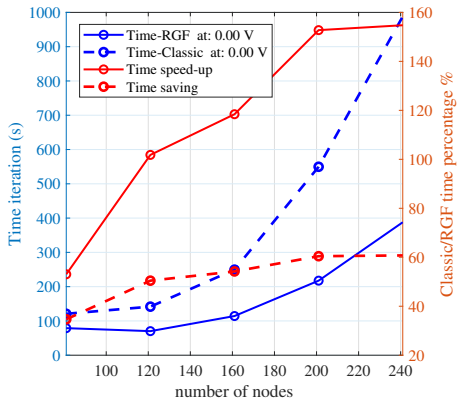


Figure 31: Comparison of the computational performance of the conventional NEGF and RGF approaches for a GaAs-based p–n junction at thermodynamic equilibrium as a function of the number of discretization nodes (the blue curve refers to the left y-axis, while the red curve refers to the right y-axis). The RGF algorithm provides an increasing computational advantage with device size, achieving a time reduction of approximately 60% for the largest structure.

Non-Equilibrium Case

Under non-equilibrium conditions, the update terms converge to slightly different values and exhibit small oscillations, as shown in Fig. 32. Despite these differences, the overall convergence behaviour of the conventional NEGF and RGF approaches remains consistent.

From a computational point of view, the execution time of both approaches increases with the number of discretization nodes, while the absolute time difference between the two methods becomes progressively larger for larger device sizes. This results in computational savings ranging from approximately 40% to 60% over the investigated device sizes.

These results further confirm the suitability of the RGF approach for semiconductor optoelectronic devices under both equilibrium and forward-bias conditions. The simultaneous treatment of electron and hole transport does not significantly affect the computational advantages provided by the recursive algorithm, demonstrating that the RGF approach remains effective also for bipolar transport conditions.

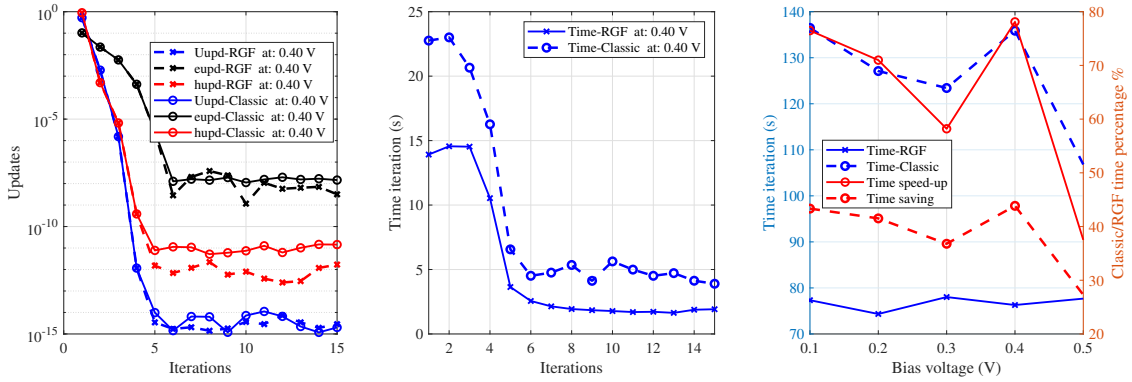


Figure 32: Comparison of the conventional NEGF and recursive Green's function (RGF) approaches for the GaAs-based p–n junction under forward bias. The left panel shows the residuals of the Poisson and carrier continuity equations, while the right panel reports the execution time and computational time saving. For a 10 nm layer thickness, the RGF approach reduces the computational time by approximately 30%–40%, depending on the applied bias.

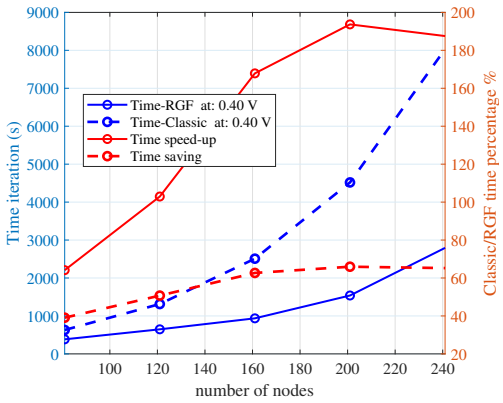


Figure 33: Comparison of the computational performance of the conventional NEGF and RGF approaches for a GaAs-based p–n junction under a forward bias of 0.4 V as a function of the number of discretization nodes (the blue curve refers to the left y-axis, while the red curve refers to the right y-axis). The computational time difference between the conventional NEGF and RGF approaches increases with device size, favoring the optimized RGF method for larger structures.

6.6 RGF Validation on a Single–Quantum–Well InGaN LED

The RGF approach is finally applied to a single–quantum–well light-emitting diode (LED) to assess its computational performance in a realistic optoelectronic device. The analysis is performed on the InGaN/GaN-based LED introduced in Section 5.5, focusing specifically on the computational cost associated with the evaluation of the Green's functions.

The main challenge is represented by the relatively high computational cost required by the NEGF simulations of the InGaN/GaN-based LED. The wide-bandgap materials and the associated numerical stiffness require high numerical precision to obtain a reliable description of quantum transport, making the evaluation of the Green's functions particularly demanding. Therefore, the simulations are performed using quadruple precision rather than double precision, providing a significantly larger number of significant digits and reducing round-off errors in the numerical calculations. This higher numerical precision improves the numerical stability and reliability of the simulations, at the cost of increased computational time and memory requirements. Since the evaluation of the Green's functions is performed repeatedly within the self-consistent procedure, its computational cost has a significant impact on the overall simulation time.

In this context, the RGF algorithm is employed to reduce the computational cost associated with the evaluation of the retarded and lesser Green's functions while preserving the physical description provided by the conventional NEGF approach. The analysis therefore focuses on the computational speed-up achieved by the recursive method.

This application represents the final validation of the RGF approach on a realistic optoelectronic device, extending the results obtained for the simpler GaAs-based structures to a wide-bandgap LED characterized by a significantly more demanding numerical problem.

The analysis focuses on the equilibrium condition, where the physical quantities obtained with the RGF approach are compared with those calculated using the conventional NEGF method. This comparison allows the consistency of the two numerical approaches to be verified before investigating the computational performance of the RGF algorithm. The simulated device structure is reported in Tab. 3.

Equilibrium Case

The results reported in Fig. 34 highlight the significant computational advantage provided by the RGF approach for the InGaN/GaN-based single–quantum–well LED. A computational time reduction of approximately 93% is achieved, with the execution time of a single outer iteration decreasing from almost 7500 s with the conventional NEGF approach to approximately 500 s when the RGF algorithm is employed. This result clearly demonstrates the dominant contribution of the Green's function evaluation to the overall computational cost of a complex wide-bandgap optoelectronic device.

Despite the substantial reduction in computational time, the two approaches provide highly consistent results. The electrostatic potential exhibits the same convergence behaviour in both simulations, while the updates of the Green's functions also show excellent agreement. The Green's function update is introduced as a diagnostic quantity to assess the accuracy of the Dyson equation evaluation. Since the Green's functions are not independent unknowns of the Newton–Raphson scheme, their updates do not constitute an additional convergence criterion for the self-consistent solution. Consequently, an apparent lack of convergence in the Green's function update, particularly for the hole component, does not imply a lack of convergence of the underlying drift-diffusion equations. Indeed, the hole continuity equation and the corresponding quasi-Fermi level converge to the same solution in both approaches. Therefore, the observed behaviour is not associated with a difference in the convergence of the physical solution, but rather with the numerical update of the Green's functions within each Newton–Raphson iteration.

Overall, the RGF algorithm significantly reduces the computational cost while preserving the convergence behaviour and accuracy of the self-consistent solution, with no appreciable differences compared with the conventional NEGF approach.

The preservation of the physical description is further verified by comparing the localized density of states (LDOS) obtained with the two approaches. As shown in Fig. 35, the spatial and energetic distribution of the LDOS is consistently reproduced by the RGF method, confirming that the recursive evaluation of the Green's functions does not compromise the physical observables of the simulation.

These results demonstrate that the RGF algorithm can provide a highly effective optimization strategy for NEGF simulations of complex optoelectronic devices, achieving a substantial reduction in computational time while preserving the accuracy of the quantum transport description.

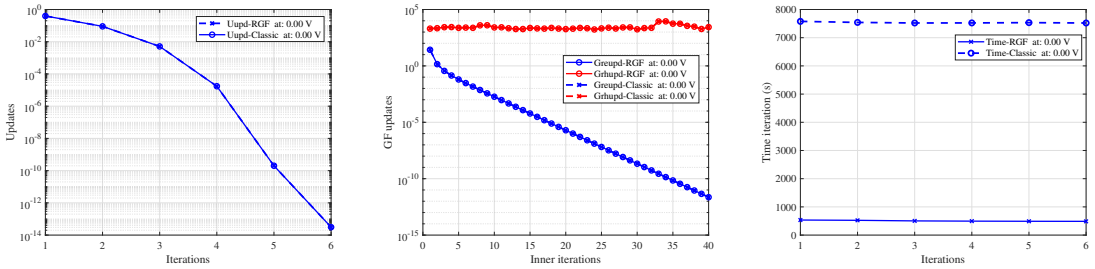


Figure 34: Comparison between the conventional NEGF and recursive Green's function (RGF) approaches under thermodynamic equilibrium conditions for a single-well InGaN/GaN-based LED. The left panel shows that the electrostatic potential converges to the same solution in both simulations, independently of the Green's function evaluation algorithm, while the corresponding Green's functions are equivalent. The right panel shows a 93% reduction in computational time when using the RGF approach, corresponding to a computational speedup of approximately 14× compared with the conventional NEGF method.

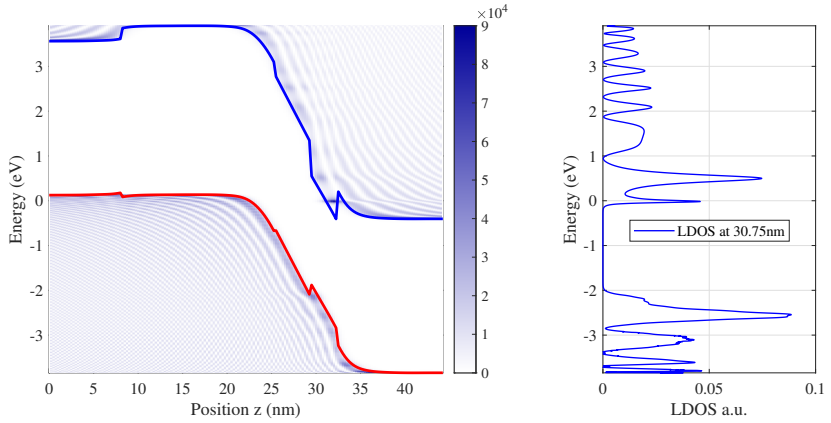


Figure 35: RGF simulation of a single-well InGaN/GaN-based LED under equilibrium conditions. The localized density of states (LDOS) for $\mathbf{k} = 0$ and the corresponding energy-resolved profile at $z = 30.75\,\text{nm}$ highlight the discretization of the quantum-well states, consistent with the results obtained from a conventional NEGF simulation.

6.7 RGF Validation on a Type–2 Superlattice

Another promising application of the RGF approach is the analysis of type-II superlattice (T2SL) structures [44], where transport is strongly influenced by quantum-mechanical tunneling. In this context, the NEGF formalism provides an accurate framework for describing carrier transport, as tunneling is naturally included in the Green's function formulation without requiring additional phenomenological assumptions. This makes T2SLs a relevant test case for assessing the applicability of the RGF algorithm to devices characterized by complex quantum transport.

Infrared photodetectors are widely based on mercury cadmium telluride (HgCdTe), a well-established material system for detection in the mid-wave infrared (MWIR) and long-wave infrared (LWIR) spectral ranges. Despite its excellent optoelectronic properties, HgCdTe technology presents significant challenges associated with material growth, fabrication complexity, and manufacturing cost. Furthermore, defects and material non-uniformities introduced during the fabrication process can adversely affect the performance and reliability of the resulting devices. These limitations have motivated the investigation of alternative material platforms, among which type-II superlattices have emerged as a promising candidate for infrared detection.

In this context, antimonide-based type-II superlattices (T2SLs) have been extensively investigated as a potential alternative to HgCdTe infrared detectors. Superlattice structures, originally introduced as a means of engineering the electronic properties of semiconductor heterostructures [44], exploit the band alignment between antimonide and arsenide compounds to engineer the electronic structure and tailor the optical response over the infrared spectral range.

The device considered in this work is a type-II superlattice (T2SL) based on GaSb, InAs, and AlSb layers. The complete layer structure, together with the corresponding material, thickness, and doping parameters, is reported in Tab. 10. The structure consists of 12 GaSb/InAs periods in the top contact region, followed by 14 periods in the absorber region. These are followed by 9 GaSb/AlSb barrier periods and 8 GaSb/InAs periods in the bottom-contact region, with an additional GaSb layer completing the structure. The material parameters are reported in Tab. 11.

We focus on an nBn configuration, consisting of n-doped layers separated by a slightly doped barrier, in which holes provide the dominant contribution to the current density flowing through the device [45] [46]. The electron flow towards the positive pole (left contact) is controlled by the

barrier, which suppresses electron transport and thereby reduces the dark current, improving the device sensitivity to photon absorption. However, the introduction of the barrier does not directly suppress SRH recombination processes. Consequently, the control of the dark current provided by the barrier is only partial.

The type-II band alignment of the superlattice leads to a spatial separation of the electron and hole states, resulting in weak overlap between the corresponding wave functions. As a consequence, the electron and hole quantum-well states can be tuned by independently varying the thicknesses of the GaSb and InAs layers, providing an additional degree of freedom for controlling the electronic structure of the superlattice.

The electron states are primarily localized in the InAs alloy, whereas the hole states are localized in the GaSb layers. By controlling the thicknesses of the constituent layers, the energy of the confined states can be tuned, allowing the photodetector to operate in the short-wave infrared (SWIR), mid-wave infrared (MWIR), and long-wave infrared (LWIR) spectral ranges.

For suitable layer thicknesses, the coupling between adjacent quantum wells can lead to the formation of Bloch states and, consequently, to the formation of minibands in both the conduction and valence bands. At the same time, the relatively high electron barrier in the conduction band provides additional control over electron leakage currents, potentially reducing non-radiative recombination processes.

The T2SL represents a significantly more complex test case for the RGF algorithm than the previously investigated GaAs-based junctions. The large number of repeated layers results in a substantial increase in the number of discretization nodes and therefore in the computational cost associated with the Green's function evaluation. This structure therefore provides a suitable test case to assess whether the computational advantage of the recursive approach observed in simpler devices is maintained in a realistic multilayer infrared optoelectronic structure.

Table 10: Layer structure of the simulated type-II superlattice (T2SL).

Region	**Material**	**Thickness (nm)**	**Type**	**Doping (cm^{-3})**	**Repetitions**
Top contact	GaSb	1.8	n	1×10^{18}	12
Top contact	InAs	2.4	n	1×10^{18}	12
Absorber	GaSb	1.8	n	1×10^{15}	14
Absorber	InAs	2.4	n	1×10^{15}	14
Barrier	GaSb	1.8	n	1×10^{15}	9
Barrier	AlSb	2.4	n	1×10^{15}	9
Bottom contact	GaSb	1.8	n	1×10^{18}	8 + 1
Bottom contact	InAs	2.4	n	1×10^{18}	8

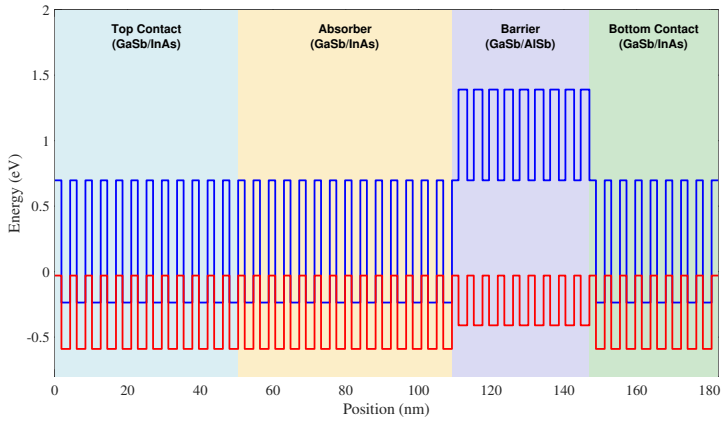


Figure 36: Schematic representation of the band diagram of an infrared photodetector based on a type-II superlattice, showing the conduction band in blue and the valence band in red.

Table 11: Physical and material parameters adopted in the DD and NEGF simulations of the type-II superlattice [47].

Name	**Symbol**	**Value**	**Unit**
Temperature	T	300	K
Energy of polar optical phonons	$\hbar\omega_{\mathrm{LO}}$	30	meV
Acoustic deformation potential	D_{aco}	10	eV
Lattice constant	a	0.5	nm
Material density	ρ_{mat}	5.32×10^{-3}	kg/cm^3
Sound velocity	u_l	6.56×10^3	m/s
Optical deformation potential	D_{op}	3×10^9	eV/cm
Maximum transverse wavevector	$k_t^{\max}$	2	nm^{-1}
Extinction coefficient	k_0	0.2831	–
GaSb material parameters			
Relative dielectric constant	ε_r	15.70	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.035\, m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.280\, m_0$	kg
Band gap	E_g	0.727	eV
Valence band offset	VBO	-0.030	eV
InAs material parameters			
Relative dielectric constant	ε_r	13.10	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.035\, m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.280\, m_0$	kg
Band gap	E_g	0.354	eV
Valence band offset	VBO	-0.590	eV
AlSb material parameters			
Relative dielectric constant	ε_r	15.7	–
Electron effective mass	$m_{\mathrm{eff},n}$	$0.035\, m_0$	kg
Hole effective mass	$m_{\mathrm{eff},p}$	$0.280\, m_0$	kg
Band gap	E_g	1.80	eV
Valence band offset	VBO	-0.410	eV

Equilibrium Case

The results reported in Fig. 37 demonstrate that the RGF approach can be effectively applied to structures with a high degree of complexity, such as type-II superlattice infrared photodetectors. In particular, the localized density of states obtained with the RGF-based NEGF implementation shows a high degree of agreement with the results obtained using the conventional approach, confirming that the recursive evaluation of the Green's functions does not introduce appreciable differences in the physical description of the device.

It is worth noting that the hole Green's function update shown in the central panel does not appear to reach the same degree of convergence observed for the electrostatic potential. This behaviour does not necessarily indicate a lack of convergence of the overall simulation. The Green's functions are not direct unknowns of the Newton–Raphson system, but are recalculated at each iteration as a function of the electrostatic potential and the corresponding self-consistent solution. Therefore, a relatively large Green's function update can indicate that the carrier Green's functions are still adjusting to the changes in the electrostatic potential during the self-consistent procedure. The convergence of the simulation is instead assessed through the residuals of the equations included in the Newton–Raphson scheme, with particular emphasis on the electrostatic potential and carrier-density equations.

From the computational point of view, the RGF approach provides a significant reduction in the simulation time even for this considerably more complex structure. A computational saving of approximately 21.6% is obtained, while the speed–up associated with the recursive procedure is approximately 27.5%. These results demonstrate that the RGF algorithm remains effective when applied to multilayer infrared photodetectors, confirming its capability to reduce the computational cost of Green's function evaluation also for large and complex structures.

From Fig. 38, the energy separation between the dominant electron and hole states is approximately 0.30 eV. This energy corresponds to an optical transition wavelength of

$$\lambda = \frac{hc}{E} \simeq \frac{1.24\,\text{eV}\,\mu\text{m}}{0.30\,\text{eV}} \simeq 4.13\,\mu\text{m}, \tag{6.56}$$

which lies within the mid-wave infrared (MWIR) spectral range. The calculated band structure therefore indicates that the considered T2SL is designed to operate in the MWIR region, with the relevant optical transitions extending approximately over the $3-5\,\mu$m wavelength range.

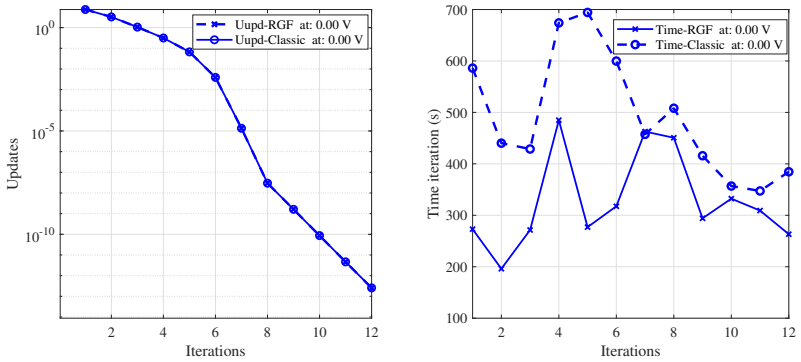


Figure 37: Comparison of the conventional NEGF and RGF approaches for the infrared photodetector based on an nBn type-II superlattice under thermodynamic equilibrium conditions. The left panel shows that the electrostatic potential corrections exhibit the same convergence behavior for both approaches. The RGF approach achieves a computational time reduction of approximately 33% over the outer Newton–Raphson cycles.

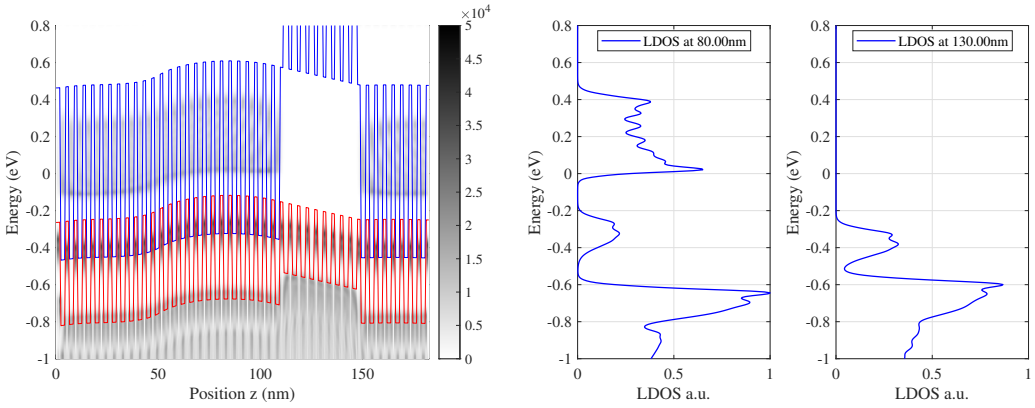


Figure 38: RGF simulation of the infrared photodetector based on an nBn type-II superlattice under thermodynamic equilibrium conditions. The LDOS at $\mathbf{k} = 0$ shows the formation of minibands within the superlattice structure. The corresponding energy-resolved profiles at $z = 80\,\text{nm}$ highlight the presence of both electron and hole minibands, whereas the cut at $z = 130\,\text{nm}$ illustrates the effect of the electron barrier, which suppresses electron miniband transport across the device while preserving the hole transport path. The results are consistent with those obtained using the conventional NEGF approach.

Non-Equilibrium Case

The non-equilibrium behaviour of the infrared photodetector is investigated by applying a forward bias to the device. For the considered T2SL structure, the simulations are performed over the $0.10-0.20\,\text{V}$ bias range, with a bias of $V = 0.10\,\text{V}$ selected as a representative case for the analysis presented below.

In Fig. 39, the convergence of the electrostatic potential and the electron and hole quasi-Fermi levels is compared for the conventional NEGF and RGF approaches under a forward bias of $0.10\,\text{V}$. The corresponding updates exhibit nearly identical convergence behaviour in the two simulations, confirming that the recursive treatment of the Green's functions does not introduce appreciable differences in the self-consistent solution. The Green's function updates also show consistent convergence behaviour for both approaches.

At the operating bias of $0.10\,\text{V}$, the RGF approach reduces the computational time by approximately 35% compared with the conventional NEGF calculation over the outer Newton–Raphson cycles. Moreover, both approaches reach the imposed convergence tolerance of 10^{-12} within the same number of iterations. This result demonstrates that the computational advantage of the RGF algorithm is achieved without affecting the convergence of the self-consistent electrostatic potential and carrier distributions.

Figure 40 shows the spatially and energy-resolved distribution of the spectral current at $V = 0.10\,\text{V}$. Despite the clear formation of electron minibands in the conduction band, the current is predominantly governed by the hole contribution. Electron transport is strongly affected by the barrier region, which limits the propagation of the electron minibands across the device. In contrast, the hole states provide the main transport path through the structure.

The figure also highlights the role of generation and recombination mechanisms in electron transport. In particular, trap-assisted processes mediated by intragap states contribute to the recovery of the electron population beyond the barrier, as shown in Fig. 41, towards the positive pole on the left. This behaviour indicates that defect-related states can provide additional transport pathways across the barrier. Consequently, an improvement in material quality, associated with a reduction in the density of trap states arising from lattice defects, could modify these generation

and recombination processes and reduce the contribution of trap-assisted transport to the dark current.

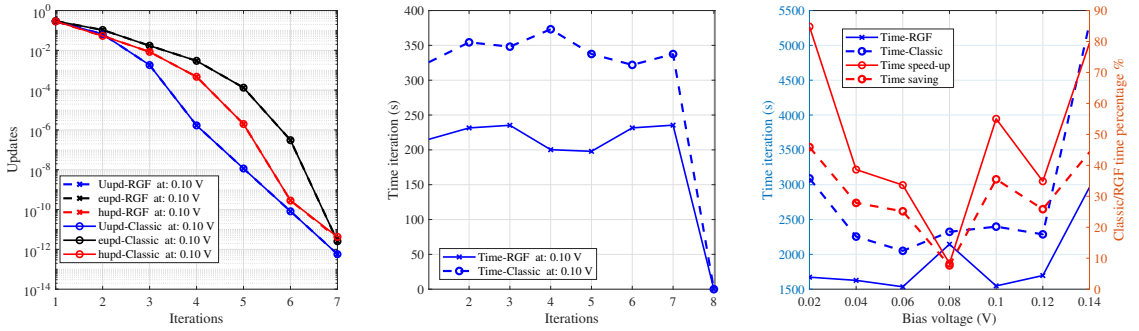


Figure 39: Comparison of the conventional NEGF and RGF approaches for the infrared photodetector based on an nBn type-II superlattice under a forward bias of 0.10 V. The left panel shows that the electrostatic potential and electron and hole quasi-Fermi-level corrections exhibit the same convergence behaviour in both approaches. The middle panel reports the computational time required by the outer Newton–Raphson cycles, while the right panel shows its dependence on the applied bias. The RGF approach provides a computational time reduction of approximately 35% at the operating bias of 0.10 V, with the advantage becoming more pronounced at some higher bias voltages.

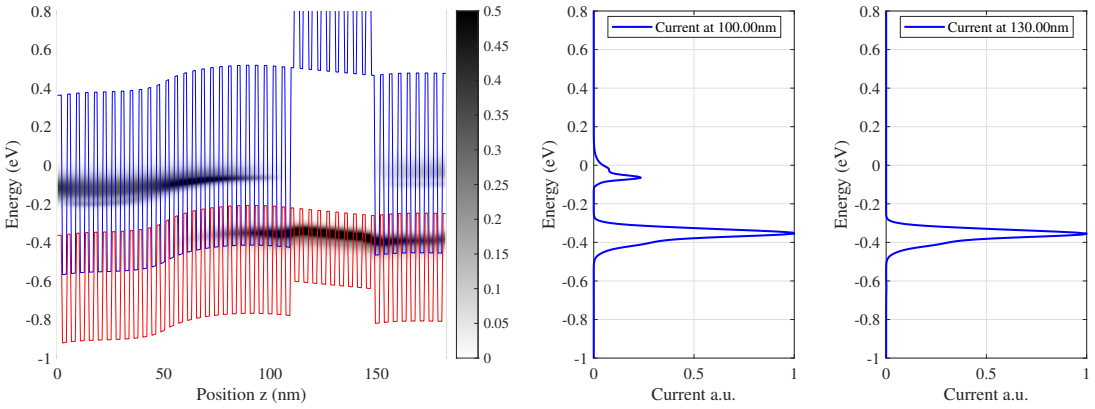


Figure 40: RGF simulation of the infrared photodetector based on an nBn type-II superlattice under a forward bias of $V = 0.10$ V. The spatial and energy-resolved spectral current density is shown together with its profile at $z = 100$ nm and $z = 130$ nm. Although electron minibands are formed within the superlattice, the electron contribution is strongly suppressed by the barrier, resulting in a current predominantly governed by hole transport.

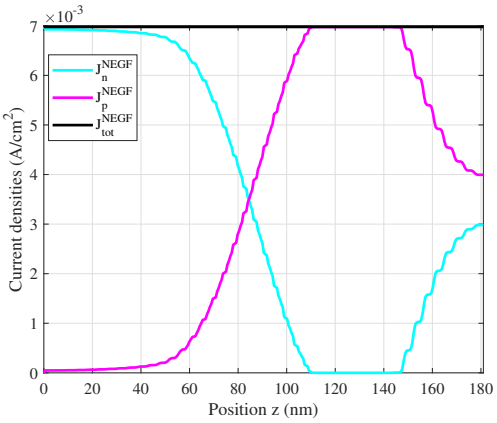


Figure 41: RGF simulation of the infrared photodetector based on an nBn type-II superlattice under a forward bias of $V = 0.10\,\mathrm{V}$. The electron and hole current density contributions are shown along the longitudinal axis of the device, with electrons flowing towards the left contact. The electronic barrier strongly suppresses the electron current, while generation and recombination mechanisms restore the electron contribution beyond the barrier, bringing the carrier concentration towards the value of the doped regions.

7 Conclusion and Outlook

This thesis investigated numerical strategies for improving the computational efficiency of non-equilibrium Green's function (NEGF) simulations for nanoscale semiconductor devices. NEGF provides a rigorous quantum-mechanical framework for describing transport phenomena such as tunneling, quantum confinement, non-equilibrium carrier dynamics, and scattering. However, its application to realistic device structures is computationally demanding, mainly due to the solution of large matrix systems over dense spatial and energy grids. Reducing this computational burden is therefore essential for extending quantum-transport simulations to increasingly complex nanostructures.

To address these limitations, several complementary strategies were developed, implemented, and validated throughout the thesis:

1. **drift–diffusion integration and quantum corrections:** the semiclassical drift–diffusion (DD) framework was integrated with quantum corrections based on the Schrödinger equation to provide efficient initial electrostatic solutions for NEGF simulations. The resulting electrostatic potential and carrier distributions were used to initialize the self-consistent Poisson–NEGF procedure, reducing the computational effort required to reach convergence. The comparison between different DD formulations also highlighted the importance of maintaining consistency between the electrostatic approximation and the subsequent quantum-transport calculation.

2. **Frozen-field NEGF approach:** the Frozen-Field approach was investigated as an effective compromise between fully self-consistent Poisson–NEGF simulations and computational efficiency. In particular, its application to single-quantum-well InGaN/GaN LEDs allowed the quantum transport problem to be solved using an electrostatic potential obtained from a preliminary DD calculation. This approach considerably reduces the computational cost of the simulation while retaining the quantum-mechanical carrier distributions required to investigate transport and recombination-related quantities. The analysis also demonstrated the sensitivity of the NEGF results to the electrostatic model used to generate the frozen potential, including the effects of incomplete ionization and polarization in III-nitride heterostructures.

3. **Recursive Green's function algorithm with Büttiker probes:** a recursive Green's function (RGF) algorithm was implemented for the efficient calculation of retarded and lesser Green's functions in one-dimensional layered structures. The method was further combined with the Büttiker probe model to account phenomenologically for dephasing and energy-relaxation processes. By exploiting the block-tridiagonal structure of the Hamiltonian, the recursive formulation avoids the direct inversion of the complete device matrix at each energy point, reducing the computational complexity associated with the Green's function calculation and significantly lowering memory requirements. The implementation was validated on progressively more complex structures, including GaAs-based junctions, single-quantum-well LEDs, and type-II superlattices. The successful application to the multilayer T2SL structure demonstrates the scalability of the recursive approach to devices with a large number of discretization nodes.

Overall, the results demonstrate that the combination of efficient electrostatic initialization, frozen-field calculations, and recursive Green's function techniques provides an effective strategy for reducing the computational requirements of NEGF simulations. Rather than replacing

the quantum-transport formalism, these approaches allow its application to increasingly complex structures while preserving the essential quantum-mechanical description of carrier transport. The validation on devices characterized by different physical mechanisms and structural complexity supports the suitability of the developed framework as a computational tool for nanoscale semiconductor device modeling.

The methodologies developed in this thesis provide a foundation for further improvements in the accuracy, scalability, and physical completeness of quantum-transport simulations. Several extensions can be considered for future work:

- **Multi-band quantum-transport models:** the extension from the single-band effective-mass approximation to multi-band approaches, such as $k \cdot p$ or tight-binding models, would allow a more accurate description of non-parabolic band structures, valence-band mixing, and hole transport in III-nitride heterostructures and superlattices.
- **Microscopic scattering models:** the Büttiker probe description could be complemented or replaced, where required, by microscopic scattering models based on the Self-Consistent Born Approximation (SCBA). In combination with the RGF formalism, this would enable a more rigorous treatment of mechanisms such as polar optical phonon scattering, acoustic phonon scattering, interface roughness, and other device-specific interactions.
- **Quantum treatment of generation–recombination mechanisms:** the semiclassical treatment of recombination processes could be extended toward a fully quantum-mechanical description through appropriate non-local self-energies. This would provide a more rigorous treatment of processes such as radiative recombination, Shockley–Read–Hall recombination, Auger recombination, and trap-assisted tunneling within the NEGF framework.
- **Extension to higher-dimensional structures:** the recursive and partitioning strategies developed in this work could be extended to two- and three-dimensional geometries. Such an extension would open the possibility of applying the methodology to more realistic device architectures, including quantum cascade lasers, infrared photodetectors, and complex III-nitride optoelectronic structures.

In this perspective, the combination of recursive algorithms, efficient electrostatic strategies, and progressively more complete physical models represents a promising route toward practical quantum-based device simulation. The results of this thesis provide a basis for continuing this development toward more accurate and computationally scalable NEGF-based TCAD methodologies.

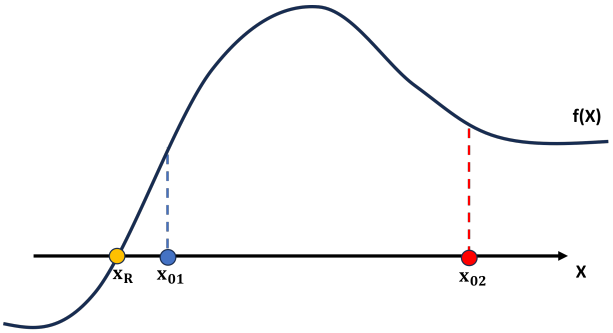


Figure A.1: Illustration of a good initial guess (blue) and a poor initial guess (red) for the determination of the root (yellow).

A Newton-Raphson algorithm

The drift-diffusion and non-equilibrium Green's function models require the solution of nonlinear systems of equations, which can only be solved numerically. We therefore begin by introducing the Newton-Raphson algorithm, a widely used iterative method for finding the roots of a system of nonlinear equations. For simplicity, let's begin by considering a single equation $f(x) = 0$ with one unknown, x. To solve this equation, we approximate the function using a first-order Taylor expansion around an initial *guess* x_0, which is expressed as follows:

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0) \tag{A.1}$$

Let x_1 denote the root of the linearized approximation. By imposing $f(x_1) \approx 0$ in A.1, we obtain:

$$f(x_1) \approx f(x_0) + (x_1 - x_0)f'(x_0) = 0 \tag{A.2}$$

from which we find that:

$$\begin{aligned} x_1 - x_0 = \Delta x &= -\frac{f(x_0)}{f'(x_0)} \\ x_1 &= x_0 + \Delta x \end{aligned} \tag{A.3a}$$

This procedure provides an increasingly accurate approximation of the solution by iteratively repeating the previous steps. However, convergence is guaranteed only if the initial guess x_0 is sufficiently close to the actual root (see Fig. A.1). The iterative process is considered converged when the update Δx becomes smaller than a prescribed tolerance[3]. The Newton-Raphson iteration can therefore be written as

$$\boxed{\begin{aligned} \Delta x_k &= -\frac{f(x_{k-1})}{f'(x_{k-1})} \\ x_k &= x_{k-1} + \Delta x_k \end{aligned}} \tag{A.4}$$

The previous derivation can now be generalized to a system of n nonlinear equations with n unknowns:

[3]The convergence tolerance should be selected according to the expected magnitude of the solution. For example, an update of 10^{-13} is negligible for a solution of order unity, but it corresponds to a relative correction of 10^{-3} for a solution of order 10^{-10}. Consequently, both absolute and relative convergence criteria are often employed in practical implementations.

$$\begin{cases} f_1(x_1, x_2, ..., x_n) = 0 \\ f_2(x_1, x_2, ..., x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, ..., x_n) = 0 \end{cases} \tag{A.5}$$

A first-order Taylor expansion is performed for each equation in the system:

$$f_i(x_1, x_2, ..., x_n) \approx f_i(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) + \frac{\partial f_i}{\partial x_1}(x_1 - x_1^{(k)}) + ... + \frac{\partial f_i}{\partial x_n}(x_n - x_n^{(k)}) \tag{A.6}$$

All partial derivatives are evaluated at the point $\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right)$. The quantities $x_i^{(k)}$ denote the values of the unknowns at the k-th iteration (for $k = 0$, they correspond to the initial guesses), whereas $x_i^{(k+1)}$ are the updated values to be determined. By imposing the linearized equations to be equal to zero, the nonlinear system is transformed into the following linear system:

$$\begin{cases} f_1(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) + \frac{\partial f_1}{\partial x_1}\Delta x_1^{(k+1)} + ... + \frac{\partial f_1}{\partial x_n}\Delta x_n^{(k+1)} = 0 \\ f_2(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) + \frac{\partial f_2}{\partial x_1}\Delta x_1^{(k+1)} + ... + \frac{\partial f_2}{\partial x_n}\Delta x_n^{(k+1)} = 0 \\ \vdots \\ f_n(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) + \frac{\partial f_n}{\partial x_1}\Delta x_1^{(k+1)} + ... + \frac{\partial f_n}{\partial x_n}\Delta x_n^{(k+1)} = 0 \end{cases} \tag{A.7}$$

where $\Delta x_i^{(k+1)} = x_i^{(k+1)} - x_i^{(k)}$ is referred to as the update of the i-th unknown. The system in Eq. (A.7) can be rearranged and written in matrix form:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & ... & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & ... & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & ... & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} \Delta x_1^{(k+1)} \\ \Delta x_2^{(k+1)} \\ \vdots \\ \Delta x_n^{(k+1)} \end{pmatrix} = - \begin{pmatrix} f_1(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \\ f_2(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \\ \vdots \\ f_n(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \end{pmatrix} \tag{A.8}$$

In a more compact way:

$$\boxed{\underline{\underline{J}} \cdot \underline{\Delta x}^{(k+1)} = -\underline{r}^k} \tag{A.9}$$

The coefficient matrix is the Jacobian matrix, whose entries are given by the partial derivatives of the residual functions with respect to the unknowns. Throughout this work, double underlining denotes matrices ($\underline{\underline{J}}$) whereas single underlining denotes vectors ($\underline{\Delta x}$). The vector $\underline{r}$ contains the residuals, namely the values of the nonlinear equations evaluated at the solution obtained from the previous iteration. The residual norm is commonly used as an additional convergence criterion.[4] Therefore, solving a nonlinear system by the Newton-Raphson method requires the evaluation of the Jacobian matrix and the residual vector. The update is then obtained by solving:

$$\boxed{\underline{\Delta x}^{(k+1)} = -\underline{\underline{J}}^{-1}\, \underline{r}^k} \tag{A.10}$$

The solution vector is then updated according to

$$\underline{x}^{k+1} = \underline{x}^k + \underline{\Delta x}^{(k+1)} \tag{A.11}$$

[4] Previously, convergence was defined in terms of the update norm. In practice, a solution may also be considered converged when the residuals are sufficiently close to zero.

Bibliography

[1] W. Van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *The Bell System Technical Journal*, 29(4):560–607, 1950.

[2] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. *"Band parameters for III-V compound semiconductors and their alloys"*. International series of monographs on physics. J. Appl. Phys, 2001.

[3] Alberto Tibaldi, Michele Goano, Francesco Bertazzi, and Enrico Bellotti. From nonequilibrium green's functions to quantum-corrected drift-diffusion: Modeling infrared superlattice photodetectors. *Physical Review Applied*, 16:044024, Oct 2021.

[4] C. de Falco, E. Gatti, A. L. Lacaita, and R. Sacco. Quantum-corrected drift-diffusion models for transport in semiconductor devices. *Journal of Computational Physics*, 204(2):533–552, 2005.

[5] Ioffe Physico-Technical Institute. Physical properties of semiconductors.

[6] A. Kumar S. E. Laux and M. V. Fischetti. *"Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects"*. J. Appl. Phys,vol. 95, no. 10, pp. 5545–5582, 2004.

[7] Julian S. Schwinger. Brownian motion of a quantum oscillator. *J. Math. Phys.*, 2:407–432, 1961.

[8] Leo P. Kadanoff and Gordon Baym. *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems.* W. A. Benjamin, New York, 1962.

[9] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Soviet Physics JETP*, 20(4):1018–1026, 1965.

[10] Supriyo Datta. *Electronic Transport in Mesoscopic Systems.* Cambridge University Press, Cambridge, 1995.

[11] Takeo Matsubara. A new approach to quantum-statistical mechanics. *Progress of Theoretical Physics*, 14(4):351–378, 1955.

[12] D. Langreth. *Linear and Non-linear Electron Transport in solids*, volume 17. Plenum Press, 1976.

[13] Urs Aeberhard. *A Microscopic Theory of Quantum Well Photovoltaics.* Ph.d. thesis, Eidgenössische Technische Hochschule Zürich, Zürich, Switzerland, 2008.

[14] C. Caroli, R. Combescot, P. Nozières, and D. Saint-James. Direct calculation of the tunneling current. *Journal of Physics C: Solid State Physics*, 4:916–929, 1971.

[15] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic. Single and multiband modeling of quantum transport in nanostructures. *Journal of Applied Physics*, 81:7845–7862, 1997.

[16] Supriyo Datta. Nanoscale device modeling: The green's function approach. *Superlattices and Microstructures*, 28:253–278, 2000.

[17] Hartmut Haug and Antti-Pekka Jauho. *Quantum Kinetics in Transport and Optics of Semiconductors*. Springer, Berlin, Heidelberg, 2 edition, 2008.

[18] Jesse A. Vaitkus and Jared H. Cole. Büttiker probes and the recursive green's function: An efficient approach to include dissipation in general configurations. *Physical Review B*, 97(8):085149, 2018.

[19] Peter Greck, Stefan Birner, Bernhard Huber, and Peter Vogl. Efficient method for the calculation of dissipative quantum transport in quantum cascade lasers. *Optics Express*, 23(5):6587–6600, 2015.

[20] Shuji Nakamura, Masayuki Senoh, and Takashi Mukai. High-power ingan/gan double-heterostructure violet light emitting diodes. *Japanese Journal of Applied Physics, Part 2*, 32:L8–L11, 1993.

[21] N. I. Bochkareva, V. V. Voronenkov, R. I. Gorbunov, A. S. Zubrilov, Y. S. Lelikov, P. E. Latyshev, Y. T. Rebane, A. I. Tsyuk, and Y. G. Shreter. Defect-related tunneling mechanism of efficiency droop in iii-nitride light-emitting diodes. *Applied Physics Letters*, 96(13):133502, 2010.

[22] G. Verzellesi, D. Saguatti, M. Meneghini, F. Bertazzi, M. Goano, G. Meneghesso, and E. Zanoni. Efficiency droop in ingan/gan blue light-emitting diodes: Physical mechanisms and remedies. *Journal of Applied Physics*, 114(7):071101, 2013.

[23] A. David, N. G. Young, C. A. Hurni, and M. D. Craven. Quantum efficiency of iii-nitride emitters: Evidence for defect-assisted nonradiative recombination and its effect on the green gap. *Phys. Rev. Appl.*, 11:031001, 2019.

[24] A. Hurni, A. David, M. J. Cich, R. I. Aldaz, B. Ellis, K. Huang, A. Tyagi, R. A. DeLille, M. D. Craven, and F. M. Steranka. Bulk GaN flip-chip violet light-emitting diodes with optimized efficiency for high-power operation. *Applied Physics Letters*, 106(3):031101, 2015.

[25] L. Y. Kuritzky, C. Weisbuch, and J. S. Speck. Prospects for 100% wall-plug efficiency III-nitride LEDs. *Optics Express*, 26:16600–16608, 2018.

[26] Jong-In Shim and Dong-Soo Shin. Measuring the internal quantum efficiency of light-emitting diodes: Towards accurate and reliable room-temperature characterization. *Nanophotonics*, 7(10):1601–1615, 2018.

[27] A. David, N. G. Young, C. Lund, and M. D. Craven. Review—the physics of recombinations in III-nitride emitters. *ECS Journal of Solid State Science and Technology*, 9(1):016021, 2020.

[28] C. Onwukaeme, B. Lee, and Han-Youl Ryu. Temperature dependence of electron leakage current in InGaN blue light-emitting diode structures. *Nanomaterials*, 12(14):2405, 2022.

[29] M. Mandurrino, G. Verzellesi, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, M. Meneghini, G. Meneghesso, and E. Zanoni. Physics-based modeling and experimental implications of trap-assisted tunneling in InGaN/GaN light-emitting diodes. *Physica Status Solidi (a)*, 212(5):947–953, 2015.

[30] M. Auf der Maur, B. Galler, I. Pietzonka, M. Strassburg, H. Lugauer, and A. Di Carlo. Trap-assisted tunneling in ingan/gan single-quantum-well light-emitting diodes. *Applied Physics Letters*, 99(18):183504, 2011.

[31] M. Goano, F. Bertazzi, X. Zhou, M. Mandurrino, S. Dominici, M. Vallone, G. Ghione, A. Tibaldi, M. Calciati, P. Debernardi, F. Dolcini, F. Rossi, G. Verzellesi, M. Meneghini, N. Trivellin, C. De Santi, E. Zanoni, and E. Bellotti. Challenges towards the simulation of gan-based leds beyond the semiclassical framework. In B. Witzigmann, M. Osiński, and Y. Arakawa, editors, *Physics and Simulation of Optoelectronic Devices XXIV*, volume 9742 of *Proceedings of the SPIE*, page 974202, San Francisco, CA, Feb 2016. SPIE.

[32] Jesús Alberto González Montoya. *Full-band NEGF Modeling of Optoelectronic Devices*. Ph.d. thesis, Politecnico di Torino, Turin, Italy, June 2023.

[33] S. L. Chuang. *Physics of Photonic Devices*. John Wiley & Sons, Hoboken, New Jersey, 2nd edition, 2009.

[34] Fabio Bernardini, Vincenzo Fiorentini, and David Vanderbilt. Spontaneous polarization and piezoelectric constants of iii-v nitrides. *Physical Review B*, 56(16):R10024–R10027, 1997.

[35] Akshay Shedbalkar, Zhelio Andreev, and Bernd Witzigmann. Simulation of an indium gallium nitride quantum well light-emitting diode with the non-equilibrium green's function method. *Physica Status Solidi B*, 253(1):158–163, 2016.

[36] F. Bertazzi, M. Moresco, and E. Bellotti. Theory of high field carrier transport and impact ionization in wurtzite GaN. Part I: A full band Monte Carlo model. *Journal of Applied Physics*, 106(6):063718, Sep 2009.

[37] W. Liu et al. Carrier leakage effect on efficiency droop in ingan/gan light-emitting diodes. *Journal of Applied Physics*, 2016.

[38] Morteza Monavarian, Arman Rashidi, Andrew A. Aragon, Sang H. Oh, Mohsen Nami, Steve P. DenBaars, and Daniel F. Feezell. Explanation of low efficiency droop in semipolar $(20\bar{2}\bar{1})$ ingan/gan leds through evaluation of carrier recombination coefficients. *Optics Express*, 25(16):19343–19353, 2017.

[39] K. B. Nam, J. Li, M. L. Nakarmi, J. Y. Lin, and H. X. Jiang. Mg acceptor level in aln probed by deep ultraviolet photoluminescence. *Applied Physics Letters*, 83(5):878–880, 2003.

[40] T. J. Badcock, M. Ali, Tongtong Zhu, M. Pristovsek, R. A. Oliver, and A. J. Shields. Radiative recombination mechanisms in polar and non-polar ingan/gan quantum well led structures. *Applied Physics Letters*, 109(15):151110, 2016.

[41] Matthias Auf der Maur, Fabio Sacconi, and Aldo Di Carlo. A parametric study of ingan/gan nanorod core-shell leds. *IEEE Transactions on Electron Devices*, 60(1):171–177, 2013.

[42] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal. Two-dimensional quantum mechanical modeling of nanotransistors. *Journal of Applied Physics*, 91(4):2343–2354, 2002.

[43] Jesse A. Vaitkus and Jared H. Cole. Büttiker probes and the recursive green's function: An efficient approach to include dissipation in general configurations. *Physical Review B*, 97(8):085149, 2018.

[44] G. A. Sai-Halasz, R. Tsu, and L. Esaki. A new semiconductor superlattice. *Applied Physics Letters*, 30(12):651–653, 1977.

[45] A. Jang, H. J. Lee, Y. H. Kim, H. Jung, P. Bidenko, and S. Kim. Effect of barrier layer on InAs/GaSb type-II superlattice nBn detector. In *2021 International Conference on Electronics, Information, and Communication (ICEIC)*, pages 1–3. IEEE, Feb 2021.

[46] David Z.-Y. Ting, Alexander Soibel, Linda Höglund, Jean Nguyen, Cory J. Hill, Arezou Khoshakhlagh, and Sarath D. Gunapala. Type-II superlattice infrared detectors. In Krishna Sanjay and Chennupati Jagadish, editors, *Advances in Infrared Photodetectors*, volume 84 of *Semiconductors and Semimetals*, chapter 1, pages 1–57. Academic Press, San Diego, 2011.

[47] Ioffe Institute. New semiconductor materials: Characteristics and properties. NSM Archive. Electronic database of semiconductor material parameters.