Full-wave discontinuous Galerkin time-domain methods for semiconductor device simulation

Massimiliano Montone

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

Méthodes Galerkine discontinues en domaine temporel pour la simulation full-wave de dispositifs à semi-conducteur

de

Massimiliano Montone

Centre Inria d'Université Côte d'Azur

Présentée en vue de l'obtention du grade de docteur en sciences pour l'ingénieur

Dirigée par: Stéphane Lanteri
Co-dirigée par: Claire Scheid
Date de soutenance: 30/03/2023

Devant le jury, composé de :
Claire Chainais-Hillairet, PR, Univ. Lille
Christophe Geuzaine, PR, Univ. Liège
Roland Masson, PR, Univ. Côte d'Azur
Eric Guichard, Senior VP and GM, TCAD Division of Silvaco, Inc
André Nicolet, PR, Univ. Aix-Marseille
Stéphane Lanteri, DR, Inria
Claire Scheid, MCF, Univ. Côte d'Azur

Inria SILVACO

Full-wave discontinuous Galerkin time-domain methods for semiconductor device simulation

by

Massimiliano Montone

Jury :

Président du jury

Roland Masson, Professeur, Univ. Côte d'Azur

Rapporteurs

Claire Chainais-Hillairet, Professeur, Univ. Lille

Christophe Geuzaine, Professeur, Univ. Liège

Examinateurs

Eric Guichard, Senior VP and GM, TCAD Division of Silvaco, Inc

André Nicolet, Professeur, Univ. Aix-Marseille

Invités

Stéphane Lanteri, Directeur de Recherche, Inria

Claire Scheid, Maître de Conférence, Univ. Côte d'Azur

Title: Full-wave discontinuous Galerkin time-domain methods for semiconductor device simulation

Abstract: In this dissertation we illustrate, from conception to implementation, a discontinuous Galerkin time-domain (DGTD) solver for optoelectronic simulation in the full-wave regime, with a particular focus on Photo-Conductive Antennas (PCAs) for THz radiation generation. These semiconductor devices absorb light and emit THz waves, and are of major relevance in a broad range of applications (from material characterization to detection of cancer tissue), but their simulation is a challenging task – it requires to concurrently describe electromagnetic field propagation and charge transport in complex geometries (usually featuring nanoparticles), whereas the two mechanisms occur at very different time and space scales. The usual mathematical model is composed of Maxwell's equations coupled with drift-diffusion (DD) equations. This system can be particularly difficult to tackle with usual numerical methods (e.g. FD, FEM, FV), so that it is typically not available in commercial solvers, if not in simplified versions that decouple the electrical and the optical part. But this happens at the detriment of accuracy and not necessarily at the advantage of implementation effort and computational cost. In the past two decades, the DGTD method has become a credible alternative to such usual methods, as it embeds and enhances some of their key features. In this work we discuss its application to the Maxwell-Drift-Diffusion (MDD) system for the simulation of PCAs. We start by introducing the model and its physical meaning, to then define a formal mathematical framework for the DG formulation – in which the presence of diffusion requires special treatment – and elaborate on key tasks such as numerical flux definition and explicit time integration with Low-Storage Runge-Kutta methods. We complement this with a few purely mathematical considerations (e.g. on well-posedness and asymptotic behavior). The MDD equations are then enhanced with Drude and Lorentz dispersion models to account for further important phenomena such as plasmon resonance and light absorption respectively. The model is coded in two dimensions and an itinerary through code verification is presented along with simulation and convergence results for relevant test cases. Finally, typical PCA geometries are simulated (classical, with anti-reflective coating, plasmon-enhanced). The process entails two steps: the steady state of the semiconductor is calculated with the aid of a commercial solver (Silvaco Atlas), then it is loaded (via a suitable interpolation technique) into the DGTD solver for the subsequent (transient) optoelectronic simulation. In the considered devices, the concurrent description of the physical components of the system and their interaction turns out to be particularly demanding in terms of space discretization; the DGTD method allows to effectively handle the constraint, thanks to the possibility of *hp-adaptive* refinement.

Keywords: Semiconductor simulation, discontinuous Galerkin, full-wave, Maxwell, Drift-Diffusion, THz.

Titre: Méthodes Galerkine discontinues en domaine temporel pour la simulation full-wave de dispositifs à semi-conducteur.

Résumé: L'objectif de la présente thèse est le développement d'un solveur Galerkine discontinu en domaine temporel (GDDT) pour la simulation en régime full-wave de dispositifs opto-électroniques tels que les antennes photo-conductives (APC) pour la génération d'ondes THz. Ces dispositifs à semi-conducteur absorbent la lumière et rayonnent des ondes THz, ce qui les rend extrêmement importants dans une vaste gamme d'applications (de la caractérisation de matériaux, à la détection de tissus cancéreux). Leur simulation est en revanche une tâche exigeante – elle entraîne la description simultanée, en géométries complexes (typiquement incluant des nano-particules), de deux phénomènes se déroulant sur des échelles spatiales et temporelles assez différentes, tels que la propagation du champ électromagnétique et le transport de charge électrique dans un semi-conducteur. Le modèle mathématique typiquement utilisé se compose des équations de Maxwell couplées avec celles de dérive-diffusion. Ce système peut s'avérer particulièrement difficile à résoudre à l'aide de méthodes numériques classiques (e.g. différences finies, éléments finis, volumes finis). Pour cette raison, en général, dans les logiciels commerciaux sa simulation passe par une simplification : la partie électrique et la partie optique sont découplées ; tout cela au détriment de la précision du modèle et pas forcément en simplifiant la conception ou en réduisant le coût computationnel. Au cours des deux dernières décennies, la méthode GDDT est devenue une alternative crédible aux méthodes numériques usuelles susmentionnées. Dans cette thèse on propose son application au système Maxwell-Dérive-Diffusion (MDD) pour la simulation d'APCs. Le point de départ est l'introduction du modèle avec sa signification physique, pour ensuite définir les fondations mathématiques de la formulation DG – dans laquelle la présence de diffusion nécessite une attention particulière – et discuter des tâches cruciales telles que la définition du flux numérique et l'intégration temporelle par des méthodes Low-Storage Runge-Kutta explicites. Des modèles de dispersion de Drude et Lorentz sont ensuite intégrés dans les équations MDD pour modéliser d'autres phénomènes importants, notamment la résonance plasmonique et l'absorption de lumière. Le modèle est codé en deux dimensions ; un parcours de vérification par étapes est proposé, ainsi que des résultats numériques et des analyses de convergence. Pour conclure, des géométries typiques d'APCs sont simulées (classique, avec couche antireflet, avec nano-particules métalliques). Le procès se déroule en deux étapes : d'abord on calcule l'état stationnaire du semi-conducteur à l'aide du solveur commercial Silvaco Atlas, ensuite on l'importe (par interpolation) dans le solveur DGTD pour la simulation opto-électronique. Dans ces dispositifs, la description simultanée et l'interaction des différentes composantes physiques du système se révèle particulièrement exigeante en terme de discrétisation en espace ; la méthode numérique considérée permet de gérer cette contrainte de façon optimale, grâce à une adaptation locale du degré d'interpolation et de la taille du maillage.

Mots clés: Simulation semi-conducteur, full-wave, Galerkine discontinu, Maxwell, dérive-diffusion, THz

Contents

Chapter 1

Introduction and motivations

Semiconductors constitute the heart and soul of integrated electronic systems, thus of modern technology. The extraordinary progress in circuit integration, since Moore's prediction from the Sixties, can certainly be majorly credited to the concurrent refinement of computer-aided design tools. Indeed, a higher scale of integration has regularly demanded a deeper understanding of semiconductor physics and more accurate mathematical models of devices to be numerically implemented and simulated. As a consequence, semiconductor device modeling has become complex to the degree that today it is an industry of its own.

1.1 Semiconductor device simulation

The main phenomena to be predicted are transport, generation and recombination of electrons and holes, in order to derive higher-level quantities (e.g. electric current and voltage) that can be practically used for device simulation. The most celebrated model is represented by the Boltzmann transport equation (BTE), which is derived from arguments of statistical thermodynamics and quantum mechanics of electrons in crystal lattices [AM76],[Jün09]. The BTE governs the distribution function of the six-dimensional position-momentum phase space; due to its high complexity, it is typically solved with Monte Carlo methods, which simulate stochastic particle trajectories between collisions rather than solving partial differential equations. These methods allow to account for the several possible collision mechanisms and different energy band structures, hence they are attractive in terms of accuracy, but also computationally expensive [VG06].

A common approximation consists in restricting the attention to only certain moments of the distribution function. Applying the method of moments to the BTE yields a set of macroscopic transport equations; to the lowest order, these constitute the Drift-Diffusion (DD) model – a pair of non-linear advection-diffusion-reaction equations that express electric charge continuity in terms of electron and hole concentrations, drift (advection) and diffusion currents, carrier recombination and generation.

Since the 1950s, the DD model has been a staple of semiconductor simulation, thanks to its simplicity and the low computational cost compared to the more rigorous BTE [ABL98],[VG06]. A further reason might lie in the refinement that it has undergone over the course of such decades, to take into account higher-order effects due to device scaling, such as velocity saturation under high electric fields. Besides, the DD framework can serve as a foundation for more refined models – additional variables and equations can be introduced by considering higher-order moments of the

BTE, leading to so-called *energy-transport* (ET) and *hydrodynamic* (HD) models [Jün09].

For example, a fundamental assumption consists in the electron gas being in thermal equilibrium with the lattice (i.e. at the same temperature). Collisions play a fundamental role in maintaining such condition, which is valid as long as the mean free path (the distance a particle travels between two collisions) is small compared to the characteristic device size. The mean free path is of the order of 100 nm, hence the initial approximation is usually legitimate in devices whose diameter exceeds 1 μm, but a rigorous analysis should also take electric field intensity into account. In essence, if carriers are accelerated beyond equilibrium, it becomes necessary to add additional equations to the picture, in order to describe momentum and energy conservation [Jün09], [FGJJ95]). In the end, the aforementioned reasons should explain why the DD model is still the default option in most semiconductor simulation software.

The main interest here is in semiconductor devices that undergo electrical and/or optical excitation. As with any dynamical system, the goal can be to investigate a transient response or a steady state. The two options lead to essentially three different models, which are summarized below and will be illustrated in detail in the body of this work.

- Static Poisson-Drift-Diffusion – To describe the steady state of a semiconductor, the static DD equations are coupled to Poisson's equation for the electrostatic potential.
- Quasi-static Poisson-Drift-Diffusion – The dynamic DD equations are coupled to Poisson's equation for the electric potential, which is supposed to vary *slowly*, in that the shortest wavelength in its spectrum is much larger than the characteristic size of the device. If so, field propagation can be neglected inside the semiconductor and the laws of electrostatics approximately hold.
- Maxwell-Drift-Diffusion (MDD) – The (dynamic) DD model is coupled to Maxwell's equations. When the shortest wavelength in the spectrum of the electric field approaches the characteristic size of the device, field propagation becomes relevant and a rigorous model is required to be electrodynamic, or *full-wave*.

The Poisson-Drift-Diffusion (PDD) model, in its static and quasi-static versions, is the backbone of *dark* electronic device simulation; the necessity of describing wave propagation, on the other hand, usually arises in the field of optoelectronics, i.e. of electronic devices that interact with or simply emit light. Notable applications include CMOS image sensors [CBV$^+$09], solar cells [DFAA12]-[ISKZ14]-[MHH$^+$19], and THz radiation generation [MPG$^+$14]-[ZTL$^+$18]. Nevertheless, it is not uncommon for optoelectronic devices to be forced into a convenient steady state before undergoing optical excitation. In such cases, a static PDD solver would be required to calculate the steady-state prior to the transient simulation, which would be quasi-static or full-wave.

The choice between quasi-static and full-wave simulation is the result of a trade-off between numerical complexity (design effort), computational load and accuracy. A recognizable trend in commercial software consists in adopting a modular approach. Typically, a PDD solver predicts carrier dynamics and interacts, according to a certain protocol, with a separate optical module that solves just Maxwell's equations (with no DD equations). For example, the two might be alternately executed and exchange information at the end of each call, until a stop criterion is fulfilled (e.g. a maximum number of iterations is reached).[1]

[1]The PDD solver in question can be static or quasi-static, depending on particular implementation choices. Ex-

The approach avoids the effort required by a single multiphysics solver, e.g. designing and testing a numerical scheme capable of handling, at once, two very different characteristic time/space scales like those associated to charge transport and electromagnetic wave propagation. However, the separate resolution of the two phenomena weakens the coupling between them and inevitably yields a model of lower accuracy. Besides, considerable design effort (and computational overhead) is likely to be required if the two modules are based – as it happens – on different numerical schemes/meshes (e.g. finite elements at the electrical side, finite differences at the optical one). Let us elaborate on this point with an example.

In the design of photo-conductive devices, i.e. devices in which conductivity increases as a result of light absorption, the major goal is to maximize the amount of light that penetrates the semiconductor. However, this amount is reduced as the conductivity increases – an effect known as *screening*. Neglecting this aspect can lead to significantly overestimate efficiency [KD16],[BES17].

Capturing mechanisms such as screening requires prompt interaction between the solution to Maxwell's equations and that to the DD ones. Let us analyze how this would be done in the two approaches.

- In the modular approach, the two solvers would have to exchange data at a high rate; the computational cost associated to the exchange itself would increase, along with the design effort required to interface the solvers. Furthermore, by solving both Poisson's and Maxwell's equations, two versions of the time-varying electric field would be calculated: a slow-varying one and an optical one.

- In the MDD approach, interaction would be instantaneous, no interfacing between different solvers (or methods) would be needed, and one time-varying electric field would be calculated. The design effort for one multiphysics solver would be high, as outlined earlier, but the choice would pay off, in terms of both accuracy and computational efficiency.

Throughout the comparison we have not mentioned a fundamental aspect: the MDD solver should evidently rely on an appropriate numerical method. The next section is devoted to this topic.

1.2 The choice of an appropriate numerical method

The finite-difference time-domain (FDTD) method [TH05] is perhaps the most commonly employed technique for simulating optical devices. It owes its success in academic and commercial contexts to the simplicity of the algorithm. However, this simplicity is also its main pitfall – the constraint of using a Cartesian mesh entails at least two major limitations [NPB09].

- Refinement cannot be local, so that high computational resources may be easily demanded to represent structures in which the electromagnetic field exhibits strong spatial variations (which is often the case in nanophotonics). Furthermore, since time integration is explicit, tight space discretization implies small timesteps. One way to fix such a flaw is to implement nested areas of varying grid density, but this generally happens at the detriment of the original algorithm's simplicity.

amples will be provided in the course of the dissertation.

- The Cartesian grid environment is also incompatible with curved objects – the resulting staircasing effect can lead to significant inaccuracy; here as well, mitigating the problem entails a more complex implementation.

In the end, a viable alternative to the FDTD method should be able to: a) accurately describe complex geometries without staircasing (hence support unstructured meshes); b) allow explicit time-stepping; c) offer higher order spatial discretization to limit mesh refinement (and its impact on the maximum stable time step). Among other possible numerical frameworks, e.g. the Finite Volume Method (FVM) or the Finite Element Method (FEM), the Discontinuous Galerkin Time-Domain (DGTD) method complies with all such criteria, as outlined through the following points [NPB09], [DDL$^+$13].

- It supports unstructured (and even non-conforming or hybrid) meshes.
- Following the Galerkin approach, the unknowns are expanded into a polynomial basis, yielding a set of coupled ordinary differential equations for the time-dependent expansion coefficients. The key difference with respect to the classical FEM, however, is that the expansion is local to each mesh element.[2]
- Different interpolation degrees are allowed in different elements; thus, so-called hp-adaptive discretization is possible (i.e. the mesh size and the interpolation degree can change locally, independently, where needed).
- When explicit time integration is used, the global matrix is block diagonal and the local problems are solved independently at each time step, yielding a natural adaptability to parallel computing.

To complete the picture, it is worthwhile to highlight a similarity between DG and FV methods. In many cases of interest Maxwell's equations can be cast in a particular form – that of a conservation law – which is the standard field of application of FV methods. These are particularly interesting thanks to their ability to capture discontinuous solutions which may occur in non-linear advection problems. DG methods share this property with FV ones since a first-order FV method can be viewed as a DG scheme in which a zeroth-order interpolation is adopted in all elements. Hence, a higher-order DG scheme can be seen as an enhancement to a classical FV one [DDL$^+$13].

Thanks to all these strengths, throughout the last two decades the DGTD method has become a credible competitor to FDTD and FEM methods for optical simulation, to the point of earning its right in the world of TCAD [ANS].

So far we have limited the attention to optical simulation – numerical approaches to optoelectronics deserve further discussion.
An extensive literature exists on the numerical treatment of the PDD system (in the transient and static cases) for linear and non-linear diffusion, with and without recombination terms.

A fundamental way of solving the static PDD system is the so-called Sharfetter-Gummel scheme, on which notable papers proposing FEM formulations [BMP89], [BMM$^+$05] are based. The scheme has been recently considered in combination with a DG method as well [CB20b].

[2]The resulting numerical solution is two-valued at interfaces between elements, and neighboring elements need to be coupled via suitable boundary conditions – a crucial point that will be discussed in detail throughout this work.

The quasi-static PDD system has been largely studied in the FV framework (see [CHF07], [BC12a], [BC12b], [BCCH19] and references therein). One fundamental principle in these works is to develop schemes that preserve some important physical properties, such as positiveness of carrier concentration and/or asymptotic stability, which in turn confer desirable properties to the discrete system. Furthermore, notable one-dimensional DGTD works exist, e.g. [LS04], [LS07], [LS10a]. It is worthwhile to also mention a recent DGTD approach that addresses a Poisson-Drift hydrodynamic model (see [GDI$^+$22] and [Blo70]).

The literature on the MDD system is comparably less rich. Purely mathematical studies exist (e.g. [Jer96]), along with a few works based on the FDTD method (e.g. [MPG$^+$14], [KYI09]), which are references in the field. Only recently, a DGTD approach has been considered [CB20a]. In fact, the literature on optoelectronic devices mostly relies on commercial solvers, which adopt the modular approach outlined in Section 1.1 and hence weakly couple electromagnetic wave propagation and carrier dynamics.

1.3 The project

The present doctoral project stems from joint research interests of the following academic and industrial entities.

The Atlantis team (Inria centre at Université Côte d'Azur) gathers applied mathematicians and computational scientists working on innovative numerical methods for nanophotonics. In this context, throughout previous research projects (e.g. [Viq15], [Sch18], [Gob20]), the team has developed DIOGENeS [3] – a suite of high-order time- and frequency-domain discontinuous Galerkin solvers, adapted to high-performance computing systems, for Maxwell's equations coupled to local and non-local dispersion models on unstructured tetrahedral meshes.

Silvaco, Inc. is a leading Electronic Design Automation (EDA) and Intellectual Property (IP) provider headquartered in Santa Clara, California and having sites in North America, Europe and Asia. Their portfolio includes software tools for process and device simulation, supporting integrated circuits (e.g. analog/mixed-signal, memories), power electronics, error reliability and photonic devices (e.g. solar cells). Semiconductor simulation is based on Atlas [Sil], a collection of two- and three-dimensional solvers for electronic devices and electromagnetic wave propagation, with a broad range of physical models. In particular, in the DD approximation, optoelectronic simulation relies on a module solving the PDD system with the Box Integration Method (BIM) and a Finite-Difference Time-Domain (FDTD) solver for Maxwell's equations.

The growing demand for optical simulation of semiconductor devices in both academic and industrial contexts has paved a common ground – at Atlantis, it has fostered the idea of expanding the know-how in DGTD methods to the field of semiconductor simulation and develop a multi-physics solver; at Silvaco, it has triggered the curiosity to investigate the potential of such methods in full-wave optoelectronic simulations.

Choosing a semiconductor model has entailed a trade-off between accuracy, theoretical complexity and computational load. As its long and great success suggests, the DD model represents an optimal synthesis of such constraints:

[3] https://diogenes.inria.fr/

- it is much less complex than the BTE, both theoretically and numerically – it allows deterministic resolution;
- it allows to refer to the repertoire of numerical methods for classical advection-diffusion-reaction problems;
- it provides an initial insight into charge carrier dynamics, but also lays the foundations for more accurate models – higher-order moments from the BTE lead to coupling further equations to the basic DD ones.

As to designating a key application for the developed solver, there has been convergence on Photo-Conductive Antennas (PCAs) for THz radiation generation. The choice is motivated in the next section. Here, we preliminarily remark two key points that will be apparent:

- PCAs require by definition the resolution of Maxwell's equations along with the DD ones;
- from a numerical standpoint, they demand to resolve strong variations in space and time of the electromagnetic field and charge distribution.

1.4 The selected application: THz radiation generation

The first time-domain spectroscopy and imaging systems were commercialized in the early 2000s and since then, the terahertz band ($0.1 - 10$ THz) of the electromagnetic spectrum has drawn increasing attention. THz emission was achieved in such systems by exciting ultrafast PCAs with sub-picosecond optical laser pulses. In the subsequent years several alternative technologies have been developed, but THz-PCAs have kept the lead thanks to the benefits of wide and continuous bandwidth, high signal-to-noise ratio, room temperature operation, and accurate magnitude and phase measurement. THz radiation is particularly attractive because it is nondestructive – the scanned samples are not altered. As a consequence, it is nowadays possible to enumerate a broad range of applications of THz-PCAs. In the industry of integrated circuits, for example, they allow fault localization and infra-package inspection. In physics, they find use in material characterization (e.g. conductivity and dielectric response). In biomedical science, they can be employed to analyze crystalline pharmaceuticals and assess margins of cancerous tissue. Moreover, considerable attention in the scientific community has been devoted to THz transmitters and receivers for wireless communication [Bur16],[KMN21].

A basic THz PCA and its operation are depicted in Figure 1.1. A semiconductor layer is contacted by two electrodes (e.g. gold), which are used to apply an electrostatic field. Ultrashort laser pulses (tens to hundreds of femtoseconds) impinge on the gap between the two electrodes and illuminate the semiconductor. A lens, typically made of silicon, collects and collimates the generated THz radiation. If the incident photon energy is greater than the bandgap, electron-hole pairs are formed in the proximity of the surface and are accelerated in opposite directions by the applied electrostatic field, yielding a *photo-current*.
As the light pulse fades in, the conductivity of the semiconductor increases proportionally to electron and hole concentrations, yielding a rise in the amplitude of the photo-current. Conversely, as the pulse fades out, the photo-generation rate drastically decreases and recombination takes the lead – photo-generated carriers rapidly recombine, the conductivity decreases consequently, and the

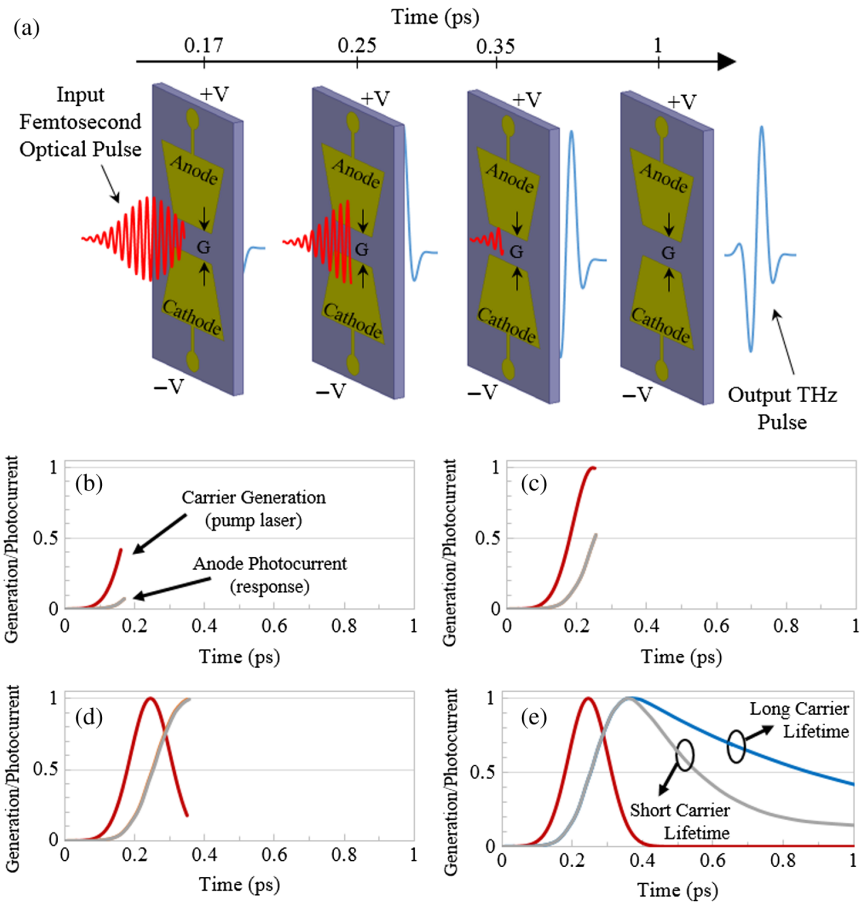


Figure 1.1: Illustration of the operation principle of a classical PCA [BES17]

photo-current as well. In essence, throughout the process, a pulsed photo-current is observed. It is to this pulse that the emission of a pulsed wave in the THz regime is ascribed.

Two main goals are identified in PCA design: maximizing the amplitude of the photo-current pulse and minimizing its fall time. The first aims at boosting the radiated field and is typically obtained by increasing the applied electrostatic field. The latter aims at ensuring that most of the radiated power actually is in the THz range – the fall of the photo-current closes the shape and hence ultimately determines its spectrum.

The fall of the photo-current is determined by recombination and in particular by carrier lifetime, a crucial physical quantity in this setup: the shorter the lifetime, the greater the emitted THz power [LGK+17].

The most common material for the illuminated layer is Low-Temperature-grown Gallium Arsenide (LT-GaAs). The reason lies in the high density of crystal defects, which yields a lifetime of the order of 0.3 ps [MPG+14]. Given that the bandgap of GaAs is 1.42 eV, the designated excitation wavelength (photon energy) is typically about 800 nm (1.55 eV) [Bur16].

An important trade-off exists in the preparation of the material. On the one hand, doping is necessary to increase the concentration of photo-induced electron-hole pairs (hence the photo-current and the emitted THz pulse). Nevertheless, higher conductivity also means higher optical

density, i.e. sharper air-semiconductor refractive mismatch, which implies higher field reflection at the air-semiconductor interface and hence less incident photons in the material.

The main pitfall of the described device is the low optical-to-THz conversion efficiency. Increasing the bias voltage is not sufficient and is anyhow limited by breakdown effects. A more viable option has been identified in enhancing light transmission into the semiconductor, to increase the probability of light absorption. The most widespread solution consists in introducing metal nanostructures on the gap between the electrodes – the resulting diffraction yields a strong near-field distribution in the underlying material and therefore improves power matching between the source and the device. It has been observed that such technological enhancement increases efficiency by a factor of 1.5 to 2 [LGK+17]. An example of the resulting structure is shown in Figure 1.2.

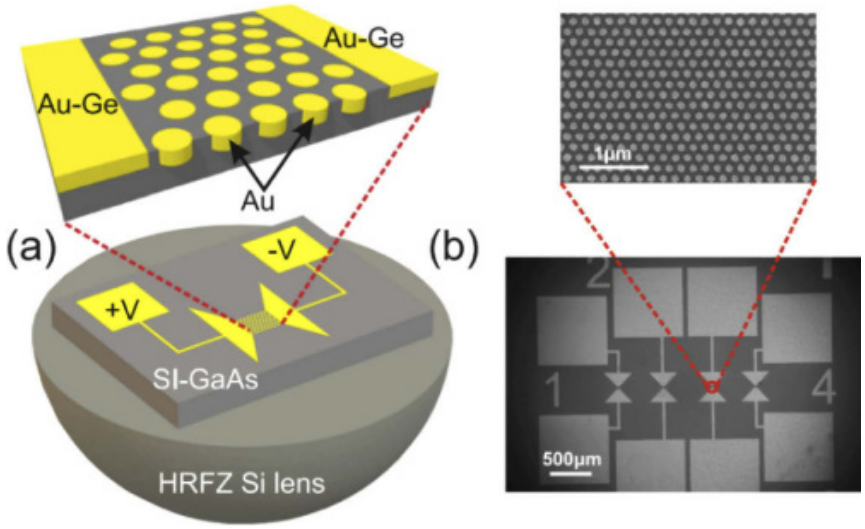


Figure 1.2: An enhanced PCA with an array of metallic nanoparticles in the gap between the two electrodes [BGP19].

PCA simulation has been the object of several different approaches. In the simplest ones, lumped-parameter models have been proposed; for example, in [KHSB13] a PCA is represented by an equivalent circuit, while in [YBG+15] a circuit replaces the semiconductor and interacts with a DGTD solver for Maxwell's equations. Attempts to solve the MDD system are found in the FDTD framework [KYI09],[MPG+14],[ZHA+15] and they all rely on a predictive (field-independent) model of photo-generation, which yields good agreement with experimental data for the classical antenna geometry.
When nanostructures are introduced, on the other hand, field-dependent photo-generation models need to be considered. However, in such cases the numerical method itself becomes anyhow inadequate – the aforementioned near-field distribution demands a very fine mesh to be correctly resolved, thereby dramatically increasing the computational load, as pointed out in Section 1.2.

All the above approaches are based on presuppositions legitimated by considering a simple, basic PCA geometry. In modern devices, however, the presence of nanostructures is systematic and entails considerably higher numerical complexity.

The task has been tackled with a two-solver FEM approach [BES17], [BPK$^+$19],[Bat21], which consists in a preliminary, purely electromagnetic simulation that calculates the photo-generation profile; the latter is then used in a subsequent PDD simulation that describes carrier dynamics. In [Bat21], actually, a third solver is used as well, to finally compute the THz field emitted by the calculated photo-current. This kind of multi-solver approach has an evident pitfall – the optical field is unaffected by carrier dynamics, hence screening cannot be described [KD16].

More recently, further DGTD approaches have been proposed. In [GDI$^+$22], Poisson's equation is coupled to a drift-only hydrodynamic model. In [CB20a] the MDD model is solved to simulate classical and nanoparticle-enhanced PCAs, and similarly in [CB21a], where dispersion is considered as well, and is proved to ensure a more accurate description of screening.

1.5 Outline

In this dissertation we illustrate, from conception to implementation, a DGTD solver for the MDD system enhanced with dispersion equations for semiconductors and metals. The solver has been implemented in two dimensions and allows to simulate both basic and complex PCAs, thanks to a field-dependent description of light absorption.

With respect to the existing literature on MDD simulation, several topics are explored, including:

- the applicability of criteria for determining the maximum mesh size that are commonly used in the DD literature;
- the advantages of locally hp-adaptive discretization;
- the calculation of the steady state of a PCA in Silvaco Atlas and the process of importing it into an unstructured mesh for the subsequent transient DGTD simulation;
- the role of steady state variables in the numerical DG formulation and their impact on the behavior of a PCA during the transient simulation;
- a detailed elaboration on how the Lorentz dispersion model from classical electrodynamics can describe light absorption and hence photo-gemeration;
- numerical orders of convergence, obtained in suitable test cases, based on explicit Runge-Kutta time integration schemes;
- the relation between the maximum stable time step and the mesh size in advection-diffusion problems – a generalization of the classical CFL condition is discussed;
- the possibility of paving the way for stability studies, by considering a priori energy bounds, or by following classical references on DG methods for elliptic problems.

In detail, our approach is organized as follows.

- Chapter 2 formally states the MDD model from a mathematical standpoint, in bounded and unbounded domains. The physical meaning of its actors is detailed – transport parameters,

such as mobility and diffusivity, are introduced along with the reaction terms describing recombination and generation processes. Special attention is devoted to the photo-generation model, which plays a pivotal role in the coupling of the electromagnetic field with carrier dynamics, and which is best described by introducing of a Lorentz dispersion model. Then, the topic of device biasing and steady state calculation is analyzed, with a presentation of the PDD model and the role of its solution in the subsequent (transient) MDD simulation. Drude and Lorentz models are finally introduced in the picture to model dispersion in metals and semiconductors, respectively. The chapter is concluded by setting the basis for a preliminary, a priori energy bound and recalling some existing mathematical results on the MDD model that link it to the PDD one.

- Chapter 3 lays the foundations for the DG formulation of the MDD model. The starting point is a local weak formulation, which paves the way for the crucial subject of defining the discontinuity of the numerical solution on the skeleton of the mesh. To this purpose, the advective and the diffusive part of the equations are separately analyzed, the former allowing a link to the FV framework. Space discretization is completed with the semi-discrete (matrix form) of the equations. The analysis proceeds with the derivation of a global weak formulation along the lines of a prominent reference on DG methods for elliptic problems. The subject of (explicit) time integration is finally contemplated, with a survey on a suitable condition for the maximum time step in the concurrent presence of advection and diffusion.

- Chapter 4 starts by deriving the two-dimensional (TE) version of the MDD model and proceeds through an itinerary of steps that aim at verifying the code from its elementary building blocks (source-free Maxwell's equations, simple advection, the heat equation) to the whole system, thanks to manufactured solutions. Along the way, particular attention is dedicated to the potential difficulty of tracking an exact solution while assigning physically meaningful values to parameters. Simulation results for three fundamental test cases are discussed, with convergence analyses.

- Chapter 5 tackles two essential aspects of simulation of real devices, in which no exact solution can be tracked: a) the problem of domain truncation and field injection; b) the crucial subject of providing a steady state solution to the MDD solver. The latter is handled as follows: each device is simulated in the TCAD environment of Silvaco Atlas, then the output of the electrostatic simulation, defined on a Cartesian mesh, is extracted and imported into the unstructured mesh used in the DGTD simulation. Finally, a selection of relevant PCAs is introduced and relevant simulation results are shown, with a detailed elaboration on the choice of an appropriate numerical setting.

- Chapter 6 draws a summary of the work with some background on additional routes that have been explored and proposes future research directions.

Chapter 2

Continuous model

This chapter presents the equations, the initial and the boundary conditions constituting the MDD model. A formulation for interior problems and exterior problems is considered – the former is mostly relevant to purely mathematical studies and numerical tests, whereas the second is here intended as the target scenario for actual device simulation.
Each function appearing in the equations is described. Designating a photo-generation function, as will be seen, is a task of special relevance as it links electromagnetic wave propagation to light absorption.
Another fundamental topic regards the static quantities that describe the steady state of the semiconductor prior to illumination. They obey a separate set of equations and boundary conditions, which will be detailed and linked to the subsequent transient simulation.
The scenario of an exterior problem is finally generalized to include Drude and Lorentz models describing dispersion in metals (e.g. the electrodes) and the semiconductor (for light absorption).
The chapter is concluded by further mathematical aspects on the MDD model, e.g. its asymptotic behavior and its link to the PDD one. Moreover the foundations for a priori energy bounds are sought.

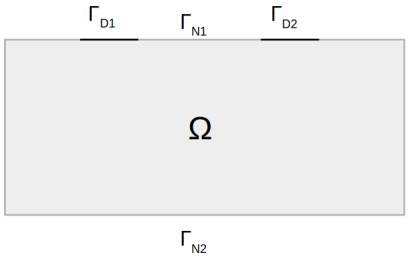


Figure 2.1: A two-dimensional bounded semiconductor domain Ω with $\partial\Omega = \Gamma_D \cup \Gamma_N$, $\Gamma_D = \Gamma_{D1} \cup \Gamma_{D2}$ representing metal contacts ideally connected to a voltage source, $\Gamma_N = \Gamma_{N1} \cup \Gamma_{N2}$ representing electrically insulating boundaries.

2.1 Interior problem

Let $\Omega \subset \mathbb{R}^3$ be an open convex polygonal domain filled with a semiconductor. Let $\partial\Omega = \Gamma_D \cup \Gamma_N$ denote the boundary of Ω, composed of the disjoint subsets Γ_D (semiconductor-PEC interfaces) and Γ_N (semiconductor-dielectric interfaces, electrically insulating boundaries).
Consistently with its physical meaning, Γ_D is assumed to have finitely many connected components. An exemplary two-dimensional geometry is proposed in Figure 2.1.

The considered mathematical model consists of Maxwell's equations coupled with charge continuity equations via drift and diffusion currents:

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_n + \mathbf{J}_p, \\
q\frac{\partial n}{\partial t} &= \nabla \cdot \mathbf{J}_n + qG - qR, \\
q\frac{\partial p}{\partial t} &= -\nabla \cdot \mathbf{J}_p + qG - qR, \\
\mathbf{J}_n &= \mathbf{J}_{n,drift} + \mathbf{J}_{n,diff} = qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha}, \\
\mathbf{J}_p &= \mathbf{J}_{p,drift} + \mathbf{J}_{p,diff} = qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta}, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p, \\
&\text{in } \Omega \times [0, +\infty[.
\end{aligned} \tag{2.1}$$

The unknowns of (2.1) are

$$\begin{aligned}
\mathbf{E}, \mathbf{H} &: \bar{\Omega} \times [0, +\infty[\to \mathbb{R}^3, \\
n, p &: \bar{\Omega} \times [0, +\infty[\to \mathbb{R}, \\
\boldsymbol{\alpha}, \boldsymbol{\beta} &: \bar{\Omega} \times [0, +\infty[\to \mathbb{R}^3,
\end{aligned}$$

and the following functions are assigned:

$$\begin{aligned}
\mu_n, D_n, \mu_p, D_p &: \mathbf{U} \in \mathbb{R}^3 \to \mathbb{R}, \\
G &: \bar{\Omega} \times [0, +\infty[\to \mathbb{R}, \\
R &: (u, v) \in \mathbb{R}^2 \to \mathbb{R}, \\
\mu, \epsilon &: \bar{\Omega} \to \mathbb{R}.
\end{aligned}$$

Their physical meaning is outlined below.

- $\mathbf{E}, \mathbf{H}$ represent the total electric and magnetic field respectively (more details in Section 2.4).
- n (p) is the electron (hole) concentration in the semiconductor, whose gradient is represented by the auxiliary vector field $\boldsymbol{\alpha}$ ($\boldsymbol{\beta}$).
- μ_n (μ_p) and D_n (D_p) are electron (hole) mobility and diffusivity respectively. They are generally non-linear functions of the electric field.

- G and R are non-linear functions modeling optical generation and net thermal recombination of electron-hole pairs.
- $\mu = \mu_0\mu_r$ denotes magnetic permeability (equal to the product of vacuum and relative permeabilities) and $\epsilon = \epsilon_0\epsilon_r$ denotes electric permittivity (equal to the product of vacuum and relative permittivities); μ_r and ϵ_r are generally space-varying.

The models used for mobility, generation and recombination are detailed in Section 2.3.

The problem is mathematically closed with the initial and boundary conditions

$$\begin{aligned}
\mathbf{E}(x,y,z,0) &= \mathbf{E}_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\
\mathbf{H}(x,y,z,0) &= \mathbf{H}_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\
n(x,y,z,0) &= n_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\
p(x,y,z,0) &= p_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\
\\
n(x,y,z,t) &= n_D(x,y,z) && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\
p(x,y,z,t) &= p_D(x,y,z) && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{E}(x,y,z,t) &= 0 && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\
\\
\hat{\mathbf{n}} \cdot \mathbf{J}_n = \hat{\mathbf{n}} \cdot \Big(n\mu_n \mathbf{E}(x,y,z,t) + D_n \boldsymbol{\alpha}(x,y,z,t) \Big) &= 0 && (x,y,z,t) \in \Gamma_N \times [0,+\infty[, \\
\hat{\mathbf{n}} \cdot \mathbf{J}_p = \hat{\mathbf{n}} \cdot \Big(p\mu_p \mathbf{E}(x,y,z,t) - D_p \boldsymbol{\beta}(x,y,z,t) \Big) &= 0 && (x,y,z,t) \in \Gamma_N \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{H}(x,y,z,t) &= \hat{\mathbf{n}} \times \mathbf{H}_N(x,y,z,t) && (x,y,z,t) \in \Gamma_N \times [0,+\infty[,
\end{aligned} \tag{2.2}$$

where $\hat{\mathbf{n}}$ is the outward unit normal vector to $\partial\Omega$ and $\mathbf{E}_0, \mathbf{H}_0, n_0, p_0, n_D, p_D, \mathbf{H}_N$ are given functions.

2.1.1 Interpreting initial and boundary conditions

The reader will notice that (2.1) features only two of Maxwell's equations. For the other two to be implied, thus omitted, the initial data cannot be arbitrary. Also, the right-hand side of the boundary conditions is chosen in compliance with the underlying physics. Assignation is done per the following criteria.

- The initial data must satisfy a few compatibility conditions. The first are

$$\begin{aligned}
\nabla \cdot \epsilon \mathbf{E_0} &= q(p_0 - n_0 + C) && \text{in } \Omega, \\
\nabla \cdot \mu \mathbf{H_0} &= 0 && \text{in } \Omega,
\end{aligned} \tag{2.3}$$

  where $C : \bar{\Omega} \to \mathbb{R}$ is the net doping. These ensure that for all $t > 0$ Gauss laws for electricity and magnetism are implied by the first two Maxwell's equations, thus not need to be explicitly included in the model.

- Γ_D represents semiconductor-metal interfaces.

  The condition on the electric field stems from assuming the metal to be a perfect electric conductor (no boundary condition would be imposed otherwise).

  As to the semiconductor, the nature of the interfaces must be specified. The hypothesis here is that these are ideal Ohmic contacts, meaning that the contact resistance and the depth of the space charge region are negligible.
  The latter condition implies that the space charge region collapses on the metallurgical junction, and becomes a mesh (boundary) face, thereby justifying that boundary conditions are imposed on Γ_D.
  Without resistance no power is dissipated, hence the contact can be described as being at thermodynamic equilibrium even when currents are flowing; as a consequence, carrier concentrations are pinned to their thermodynamic-equilibrium values [MK02] and hence verify the *law of mass action*

$$n(x,y,z,t)\ p(x,y,z,t) = n_i^2 \qquad (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \tag{2.4}$$

  where n_i denotes the intrinsic concentration.[1]

  It is worthwhile to mention an alternative way to justify the last condition. For a general metal-semiconductor interface one would typically enforce the Neumann boundary conditions [Nan91]

$$\begin{aligned} \hat{\mathbf{n}} \cdot \mathbf{J}_n(x,y,z,t) &= -q\ \mathfrak{v}_n\ \Delta n(x,y,z,t) && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\ \hat{\mathbf{n}} \cdot \mathbf{J}_p(x,y,z,t) &= q\ \mathfrak{v}_p\ \Delta p(x,y,z,t) && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \end{aligned} \tag{2.5}$$

  where $\Delta n, \Delta p$ denote the deviation of carrier concentrations from thermodynamic-equilibrium values, and $\mathfrak{v}_n, \mathfrak{v}_p$ are the surface recombination velocities of electrons and holes respectively. Even if dimension-wise these are velocities, they are not related to an actual velocity but rather to the rate at which excess carriers recombine at the interface [MK02]. The condition $\mathfrak{v}_n = \infty = \mathfrak{v}_p$ is an alternative definition of ideal Ohmic contact and implies $\Delta n = 0 = \Delta p$ in order to have finite current.

  If the dopant distribution C slowly varies with space – it is the case in all applications of our interest – the net charge can be assumed to be zero outside the space charge region [MK02]. In case of negligible depth, the edge of the region coincides with the metallurgical junction, which is located on Γ_D. Hence it is imposed that

$$p(x,y,z,t) - n(x,y,z,t) + C(x,y,z) = 0 \qquad (x,y,z,t) \in \Gamma_D \times [0,+\infty[. \tag{2.6}$$

  Substituting (2.6) into (2.4) defines the right-hand sides as

$$n_D = \frac{C + \sqrt{C^2 + 4n_i^2}}{2} \qquad p_D = \frac{-C + \sqrt{C^2 + 4n_i^2}}{2}. \tag{2.7}$$

  Note that (2.7) are time-independent as motivated by the hypothesis of thermodynamic equilibrium. They actually are space-independent as well – in practice, contacts are fabricated on uniformly doped regions, hence C is constant on each segment of Γ_D.

[1] This is defined as the free electron concentration that exists in an undoped semiconductor at thermodynamic equilibrium. It generally varies with temperature and doping – common models can be found for example in [Sil]. In the present context it is assumed to be constant.

- Γ_N represents semiconductor-dielectric interfaces, which are supposed to be electrically insulating i.e. to prevent electrons and holes from flowing outside the semiconductor. Mathematically, this translates to $\mathbf{J}_n$ and $\mathbf{J}_p$ being tangent to Γ_N.
  The hypothesis applies to a steady state as well, hence it is required that

$$\hat{\mathbf{n}} \cdot \left(n_0 \mu_n \mathbf{E_0} + D_n \nabla n_0 \right) = 0 = \hat{\mathbf{n}} \cdot \left(p_0 \mu_p \mathbf{E_0} - D_p \nabla p_0 \right) \quad \text{on} \quad \Gamma_N. \tag{2.8}$$

  Assigning $\mathbf{H}_N$ is required mathematically but it is cumbersome from a physical standpoint, as such a vector field should either represent a current density or the tangential component of the magnetic field in a surrounding volume in which the problem were not solved.
  In fact, the interior problem itself is just a mathematical tool – a stepping stone to a scattering problem, as anticipated in Section 2.1.
  It is licit to consider it so, because: a) (2.1) is valid in both scenarios; b) when switching between the two, only the electromagnetic part of (2.2) changes; c) the changes are irrelevant to the core task – developing a DG scheme for Maxwell's equations coupled with nonlinear advection-diffusion-reaction.
  In the end, $\hat{\mathbf{n}} \times \mathbf{H}|_{\Gamma_N} = \hat{\mathbf{n}} \times \mathbf{H}_N$ is meant to mimic (and prepare the ground for) a transmission condition to be imposed in an exterior problem.

2.2 Exterior problem

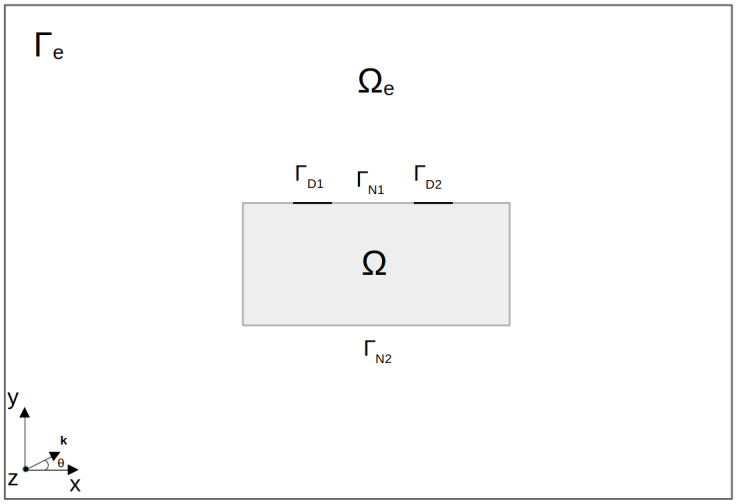


Figure 2.2: A two-dimensional representation of the original semiconductor domain Ω immersed in a bounded dielectric domain Ω_e and subject to an incident field – typically a plane wave of vector $\mathbf{k}$ forming an angle θ with the x axis.

We now let the bounded semiconductor domain Ω be immersed in Ω_e, an ideally unbounded dielectric which is artificially truncated for computational purposes (Figure 2.2). The dielectric is

assumed to be linear, isotropic, non-dispersive in time and homogeneous in space. The pair $(\mathbf{E}^i, \mathbf{H}^i)$ represents a plane wave, possibly pulsed in both space and time, whose spectrum is centered at an optical frequency; it impinges on the device and generates free charge carriers. This is the incident field and is solution to source-free Maxwell's equations in Ω_e in the absence of the scatterer. Linearity of such equations allows a scattered-incident field decomposition in $\Omega_e \setminus \Omega$, while inside Ω the incident field does not solve Maxwell's equations. In the end we write

$$(\mathbf{E}, \mathbf{H}) = \begin{cases} (\mathbf{E}^s, \mathbf{H}^s) + (\mathbf{E}^i, \mathbf{H}^i) & \text{in } \Omega_e \setminus \Omega \\ (\mathbf{E}^t, \mathbf{H}^t) & \text{in } \Omega \end{cases}, \tag{2.9}$$

with the scattered field $(\mathbf{E}^s, \mathbf{H}^s)$ and the internal field $(\mathbf{E}^t, \mathbf{H}^t)$ to be calculated numerically.
The equations in (2.1) remain valid, but outside Ω the model consists in source-free Maxwell's equations

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\ \nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t}, \\ &\text{in } \Omega_e \setminus \Omega \ \times [0, +\infty[, \end{aligned} \tag{2.10}$$

and $\mathbf{E}, \mathbf{H}$, μ, ϵ are defined in $\bar{\Omega}_e$ rather than just in $\bar{\Omega}$.
The initial and boundary conditions for the semiconductor equations remain the same, whereas a change in those for $\mathbf{E}, \mathbf{H}$ is required:

$$\begin{aligned} \mathbf{E}(x,y,z,0) &= \mathbf{E}_0(x,y,z) && (x,y,z) \in \bar{\Omega}_e, \\ \mathbf{H}(x,y,z,0) &= \mathbf{H}_0(x,y,z) && (x,y,z) \in \bar{\Omega}_e, \\ \hat{\mathbf{n}} \times \mathbf{E} + \sqrt{\frac{\mu}{\epsilon}} \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{H}) &= \hat{\mathbf{n}} \times \mathbf{E}^i + \sqrt{\frac{\mu}{\epsilon}} \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{H}^i) && \text{on } \Gamma_e \times [0, +\infty[, \\ n(x,y,z,0) &= n_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\ p(x,y,z,0) &= p_0(x,y,z) && (x,y,z) \in \bar{\Omega}, \\ n(x,y,z,t) &= n_D(x,y,z) && (x,y,z) \in \Gamma_D \times [0, +\infty[, \\ p(x,y,z,t) &= p_D(x,y,z) && (x,y,z) \in \Gamma_D \times [0, +\infty[, \\ \hat{\mathbf{n}} \times \mathbf{E}(x,y,z,t) &= 0 && (x,y,z,t) \in \Gamma_D \times [0, +\infty[, \\ \hat{\mathbf{n}} \cdot \mathbf{J}_n = \hat{\mathbf{n}} \cdot \Big(n\mu_n \mathbf{E}(x,y,z,t) + D_n \boldsymbol{\alpha}(x,y,z,t) \Big) &= 0 && (x,y,z,t) \in \Gamma_N \times [0, +\infty[, \\ \hat{\mathbf{n}} \cdot \mathbf{J}_p = \hat{\mathbf{n}} \cdot \Big(p\mu_p \mathbf{E}(x,y,z,t) - D_p \boldsymbol{\beta}(x,y,z,t) \Big) &= 0 && (x,y,z,t) \in \Gamma_N \times [0, +\infty[. \end{aligned} \tag{2.11}$$

The first novelty is the absence of $\hat{\mathbf{n}} \times \mathbf{H}|_{\Gamma_N}$ – transmission conditions are implicitly required on Γ_N, namely the continuity of $\hat{\mathbf{n}} \times \mathbf{E}$ and $\hat{\mathbf{n}} \times \mathbf{H}$. The second is the Absorbing Boundary Condition (ABC) on Γ_e, which is meant to approximate the notable Silver-Müller radiation condition and enforce propagation of $(\mathbf{E}^s, \mathbf{H}^s)$ towards infinity in all directions with no reflection, as it would be

if $\Omega_e = \mathbb{R}^3$. Bearing in mind (2.9) the ABC can be equivalently stated as

$$\begin{aligned} \hat{\mathbf{n}} \times \left(\mathbf{E}^s + \sqrt{\frac{\mu}{\epsilon}} \hat{\mathbf{n}} \times \mathbf{H}^s \right) = \mathbf{0} \qquad & \text{on } \Gamma_e \times [0, +\infty[, \\ (\mathbf{E}^s, \mathbf{H}^s) = (\mathbf{E} - \mathbf{E}^i, \mathbf{H} - \mathbf{H}^i), & \end{aligned} \tag{2.12}$$

and requires the scattered field to behave (locally to the observation point on Γ_e) as a plane wave impinging normally to the boundary. While the plane wave behavior can be achieved to a fair degree by placing Γ_e far enough from $\partial\Omega$, the angle of incidence cannot be controlled and represents an inherent inefficacy to the ABC – a topic that will be further discussed in Section 5.1.

2.3 Key functions in the Drift-Diffusion model

We present here in detail the reaction terms (recombination, optical generation) and the transport coefficients (mobility, diffusivity).

2.3.1 Recombination

Recombination is the process by which a free electron fills a hole: the free electron disappears (it becomes a bound electron) and the hole too (the vacancy is filled).
Generation is the reverse process – a bound electron acquires enough energy for a transition from the valence band to the conduction band, it thus becomes free and leaves a hole "behind".
Two mechanisms are typically modeled, as discussed below.

Generation-recombination due to lattice vibrations, commonly modeled by the prominent Shockley-Read-Hall model [SR52], which describes it as a two-step process thanks to lattice defects (called *traps*). The idea is that such imperfections create an allowed energy level in the bandgap, and offer the possibility of band-to-trap and trap-to-band transitions. Such intermediate passages require smaller amounts of energy than direct interband ones, hence they happen at a higher rate.

The Auger effect, which is qualitatively described in [Dem10] as follows.[2] An excited electron must release its excess energy in order to decay into an available lower-energy state. The decay is generally radiative (photon emission), but it can also happen that the energy is directly transferred to another electron; and if it's higher than the binding energy of the second electron, auto-ionization happens.

The semiconductor picture is more complex, with four types of carrier-carrier interaction to be considered [Sel84]:

(a) Electron capture – an electron recombines and the excess energy is transferred to another conduction electron;

(b) Hole capture – an electron recombines and the excess energy is transferred to a hole, which moves away from the valence band edge;

(c) Electron emission – a free electron is generated and the excitation energy comes from a high energetic electron in the conduction band;

[2] Named after P.V. Auger who discovered it in 1923 [Aug23].

(d) Hole emission – a free electron is generated and the excitation energy comes from a high energetic hole in the valence band.

As a result, the net thermal recombination rate R is described by

$$R(n,p) = \frac{np - n_i^2}{\tau_n(p_1 + p) + \tau_p(n_1 + n)} + (np - n_i^2)(C_{An}n + C_{Ap}p). \tag{2.13}$$

with the first addend being the SRH model and the second accounting for the Auger effect. By net, it is meant that $R > 0$ reduces electron and hole concentrations (recombination), $R < 0$ pulls concentrations up (generation). The parameters appearing in (2.13) are described below.[3]

- τ_n (τ_p) is the electron (hole) lifetime, that is the average time it takes to electrons (holes) to recombine.

- n_1 (p_1) is the electron (hole) concentration that would be established at thermodynamic equilibrium if the (extrinsic) Fermi level E_F matched the trap one E_t [VG06].[4]
  In formulae,

$$n_1 = n_i \exp\left(\frac{E_F - E_i}{k_B T}\right)\Big|_{E_F = E_t}, \qquad p_1 = n_i \exp\left(\frac{E_i - E_F}{k_B T}\right)\Big|_{E_F = E_t}. \tag{2.15}$$

  In the expressions n_i is the intrinsic concentration, E_i is the intrinsic Fermi level, k_B is the Boltzmann constant and T is the temperature of the lattice.

  Note that unless otherwise expressly stated, in the present work it is assumed that carriers are in thermal equilibrium with the lattice, at a common uniform temperature $T = 300$ K.

- C_{An} (C_{Ap}) is the Auger capture coefficient for electrons (holes) and accounts for the rate at which events of type (a) (type (b)) happen. It is measured in $\mathrm{m}^6\ \mathrm{s}^{-1}$.

  Note that (2.17) does contemplate events of type (c) and (d) but the corresponding coefficients do not appear explicitly because they are expressed in terms of C_{An} and C_{Ap} [Sel84].

The function R models non-radiative recombination-generation processes working towards establishing the thermodynamic condition $np = n_i^2$ (the law of mass action) and finds use in both transient and stationary problems. To our purposes, however, it is convenient to separate R into a stationary and a dynamic part as follows.
Electron concentration can be thought of as the sum of a static term and a transient deviation: $n = n_{DC} + n_{AC}$. Similarly, $p = p_{DC} + p_{AC}$. The concept will be detailed in Section 2.4, where separate sets of equations for the static and dynamic components will be sought. The key point here is that n_{AC}, p_{AC} are null at equilibrium (by definition), sway away from it when optical generation

[3] A practical example of thermal generation will be shown in Section 5.3.

[4] Textbooks on semiconductor devices (e.g. [MK02]) usually introduce the Fermi level E_F along with the *Fermi-Dirac distribution function*, giving the probability that an allowed one-electron energy level E is occupied, which has the form

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)}, \tag{2.14}$$

and predicts that at $T = 0$K all the levels below E_F are occupied and all those above E_F are unoccupied.

of electron-hole pairs takes place, to finally collapse back to zero once illumination is over; and to describe such behavior in dedicated equations it is desirable to consider the decomposition

$$R = R' + R'', \tag{2.16}$$

with

$$\begin{aligned} R'(n_{DC}, p_{DC}) &= R(n_{DC}, p_{DC}) \\ &= \frac{n_{DC}\ p_{DC} - n_i^2}{\tau_n(p_1 + p_{DC}) + \tau_p(n_1 + n_{DC})} + (n_{DC}\ p_{DC} - n_i^2)\ (C_{An}\ n_{DC} + C_{Ap}\ p_{DC}), \end{aligned} \tag{2.17}$$

modeling steady-state recombination, which works towards bringing n_{DC}, p_{DC} to the thermodynamic equilibrium values, and

$$\begin{aligned} R''(n_{AC}, p_{AC}) &= R(n_{AC}, p_{AC})|_{n_i=0} \\ &= \frac{n_{AC}\ p_{AC}}{\tau_n(p_1 + p_{AC}) + \tau_p(n_1 + n_{AC})} + n_{AC}\ p_{AC}\ (C_{An}\ n_{AC} + C_{Ap}\ p_{AC}), \end{aligned} \tag{2.18}$$

specifically operating on transient deviations with respect to the steady-state and working towards a null equilibrium condition.

2.3.2 Optical generation

G represents the rate at which electron-hole pairs are generated by photon absorption, i.e. the number of absorbed photons per unit volume and time.

2.3.2.1 Predictive approach

For normal incidence on the semiconductor surface a common simple expression is [RV04],[ST07]

$$G = \Phi_0\ \alpha\ \eta_o\ \eta_i\ \exp(-\alpha d), \tag{2.19}$$

where:

- $\Phi_0(\mathrm{m}^{-2}\ \mathrm{s}^{-1})$ is the incident photon flux (number of incident photons per unit surface) per unit time;
- $\alpha(\mathrm{m}^{-1})$ is the absorption coefficient and is frequency-dependent;
- $\eta_o = 1 - \mathfrak{R}$ is the *optical quantum efficiency* and accounts for photon loss by light reflection at the surface; $\mathfrak{R}$ denotes Fresnel's reflection coefficient;
- η_i is the *internal quantum efficiency* – the fraction of transmitted photons that is actually absorbed;
- d is the distance from the surface – modeling exponential attenuation of the light flux towards the bulk.[5]

[5] As pointed out in [RV04], η_o can be quite low for a semiconductor-air interface, e.g. 0.7 in the case of silicon. This is why photo-detectors are usually treated with anti-reflective coating – in essence an impedance matching layer.

This model is used in FDTD papers (e.g. [MPG+14],[KYI09]) to solve (2.1) with focus on PCAs similar to the one in Figure 2.3. The incident photon flux is described by a modulated irradiance - reproducing the Gaussian character in space and time of the narrow-band laser - divided by the photon energy corresponding to the carrier frequency ν_0:

$$\Phi_0 = \Phi_0(x,y,z,t) \quad = \frac{I_0}{h\nu_0} \exp\left(-\left(\frac{x-x_0}{\sigma_x}\right)^2 - \left(\frac{y-y_0}{\sigma_y}\right)^2 - \left(\frac{t-t_0-\frac{z-z_0}{v_{sc}}}{\sigma_t}\right)^2 \right). \tag{2.20}$$

Here, $z = z_0$ locates the semiconductor surface and (x_0, y_0, z_0) is the gap center. Normal incidence is assumed and wave transmission (photon injection) along z is predicted. In essence, incident photons are transmitted at the interface just like in (2.19), then travel away from it at the speed of light $v_{sc} := \sqrt{1/(\epsilon\mu)}$.
Predict is the keyword here because irradiance is another word for the time-averaged Poynting vector [Hec12], hence calculating its actual value inside the semiconductor is incompatible with the goal of solving Maxwell's and semiconductor equations at the same time.

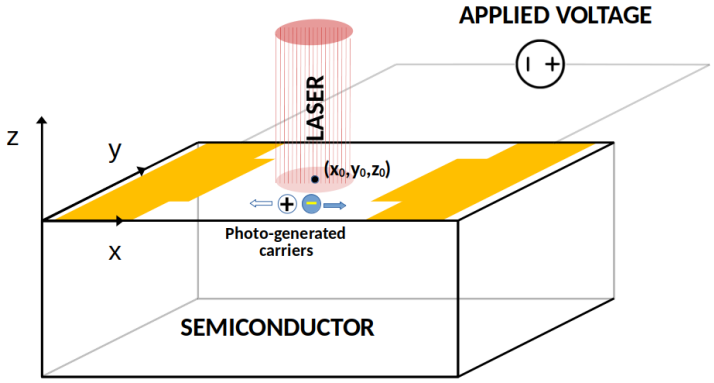


Figure 2.3: An illustration of the photo-generation process considered in [MPG+14] and [KYI09]. The center of the laser spot matches that of the gap, namely (x_0, y_0, z_0). The spot is much smaller than the gap, so that diffraction by the electrodes can be neglected and the transmitted photon flux can be approximated as $\eta_o\Phi_0$ (with Φ_0 given by (2.20)). Below the semiconductor surface ($z = z_0$), photo-generated electrons and holes are accelerated in opposite directions thanks to the applied voltage.

It is important to remark that experimental results validate (2.20) for geometries featuring a planar and sufficiently wide gap, i.e. in which the laser spot is well contained in the gap so that diffraction by metallic components is negligible. However, great interest has been devoted by the scientific community to geometry optimization. As highlighted in [LGK+17], in attempts to boost photo-generation (thus optical-to-THz conversion efficiency) major benefit has been drawn from the integration of plasmonic structures in the gap. The underlying principle is that electromagnetic radiation penetrating sub-wavelength metallic nanoparticles induces collective oscillations in the electron gas, which yield near-field enhancement [Mai07]. In all the investigated geometries it is pivotal to intercept the incident wavefront with sub-wavelength-sized metallic obstacles. Hence, predicting the field distribution becomes unfeasible and photo-generation enhancement can only be observed if the actual irradiance inside the semiconductor is calculated.

2.3.2.2 State of the art of commercial software

The aforementioned incompatibility is typically overcome in commercial software by having several separate modules:

- one for wave propagation which solves Maxwell's equations;
- one for charge carrier dynamics that solves either of the two following systems, which are detailed in Section 2.4:
  (a) the static version of (2.1);
  (b) an electro-quasi-static version of (2.1) in which time-varying semiconductor equations are coupled with Poisson's equation for a time-varying electric potential.

It is the case of the software Silvaco Atlas [Sil], where a preliminary simulation computes the *dark* steady state of an optoelectronic device, then the following loop takes place until a given stop condition is fulfilled.

1. The calculated $\mathbf{J}_n$ and $\mathbf{J}_p$ are sources to Maxwell's equations in the Luminous FDTD solver.
2. Luminous simulates electromagnetic scattering by solving Maxwell's equations until a steady state is reached (the energy difference between two consecutive iterations is monitored).
3. The calculated optical field is passed back to the charge transport module (of type (a)) which computes a new steady state of n and p with photo-generation in the equations.

The generation function is

$$\begin{aligned} G(x,y,z) &= \alpha \ \frac{S(x,y,z,\tilde{t})}{h\nu_0}, \\ \alpha &= 4\pi k(\nu_0), \\ S(x,y,z,\tilde{t}) &= \frac{||\mathbf{E}(x,y,z,\tilde{t})||^2}{2\zeta}, \end{aligned} \tag{2.21}$$

$\tilde{t}$ being the discrete time at which the optical steady state is reached (and a new semiconductor simulation starts), ν_0 the carrier frequency of the (narrow-band) incident field, ζ the characteristic impedance, and k the imaginary part of the refractive index.
It is interesting to analyze how (2.21) compares to (2.19). Beside $\eta_i = 1$ (all injected photons are absorbed) and $d = \infty$ (unlimited light penetration depth), the idea is: compute the photon flux inside the semiconductor (= Injected power density / $h\nu_0$) from the optical simulation rather than predict how the incident photon flux is transmitted across the surface ($\Phi_0 \ \eta_o \exp(-\alpha d)$).
The injected power density appearing in (2.21) is a function of the steady-state optical field and its expression recalls the time-averaged Poynting vector of a plane wave.

Anther relevant model is found in [Bur16] for narrow-band simulation of THz PCAs in the COMSOL environment. A frequency-domain solver for Maxwell's equations calculates the sinusoidal steady state at the carrier frequency ν_0; then, a quasi-static solver of type (b) determines electron and hole dynamics and the evolution of the time-varying electric potential, with

$$G(x,y,z,t) = \alpha(\nu_0) \ \frac{P(x,y,z,\nu_0)}{h\nu_0} \ \exp\left(4 \ ln(0.5) \ \left(\frac{t-t_0}{\sigma_t}\right)^2\right), \tag{2.22}$$

where α has the same expression as (2.21) and P is the (sinusoidal) steady-state injected power density. Again, P is a Poynting-like function of the optical electric field, but the expression is much less simple than (2.21) and omitted for brevity.

It is worthwhile to notice that (2.22) brings three key concepts together: power loss (α); active power associated to the computed electromagnetic field distribution inside the semiconductor (P); and the instantaneous, pulsed character of optical excitation (the Gaussian function).

An important limitation of (2.20) and (2.22) with respect to (2.21) should be highlighted. The total drift current is related to the electric field by Ohm's law

$$\mathbf{J}_{drift} = (qn\mu_n + qp\mu_p)\mathbf{E}, \tag{2.23}$$

with $qn\mu_n + qp\mu_p$ being the conductivity of the semiconductor. As the excitation pulse rises, n and p increase and so does the conductivity; as a result, the field penetration depth is reduced (and photon injection with it). The first two models fail to capture the mechanism because they fix G prior to the semiconductor simulation, whereas the feedback loop between Silvaco modules seems a more rigorous option.

2.3.2.3 Rigorous approach

The quest for real-time photo-generation should start from finding time-domain counterparts to the frequency-domain quantities encountered so far – time-domain absorption should be the keyword. In quantum optics, absorption is defined in terms of the probability of electron transition between two energy levels, which is calculated relying on perturbation theory [GAF10]. Such a course is pursued in [BPW$^+$22] to study valence-to-conduction-band transitions in semiconductors. The authors elaborate on the view of the photon absorption process as a transfer of energy to the electron throughout the interaction time window, in which the wave function evolves from an initial state v (of the valence band) to a final state c (of the conduction band). This leads to establishing a fascinating bridge to classical electrodynamics, which identifies the instantaneous absorbed power density with the work done by the electric field on the moving electron, with the definition of

$$P_{v\to c}(x,y,z,t) = \mathbf{J}_{v\to c}(x,y,z,t)\cdot\mathbf{E}(x,y,z,t) \tag{2.24}$$

as the power density absorbed at time t during the transition from state v to state c. $\mathbf{J}_{v\to c}$ denotes the current density operator associated to such a transition; its detailed expression is quite complex and intentionally omitted to keep the focus on the search of a counterpart to (2.24) that suit our non-quantum model.

The collective displacement of bound electrons is commonly described semi-classically by the Lorentz model - a damped harmonic oscillator - and subsequently expressed in terms of a current density. Following [Gra16], electron motion in the presence of an electric field is described by Newton's law as

$$m\frac{\partial^2\mathbf{x}}{\partial t^2} = q\mathbf{E} - m\gamma\frac{\partial\mathbf{x}}{\partial t} - K\mathbf{x} \iff \frac{\partial^2\mathbf{x}}{\partial t^2} + \gamma\frac{\partial\mathbf{x}}{\partial t} + \omega_r^2\mathbf{x} = \frac{q}{m}\mathbf{E}, \tag{2.25}$$

where $\mathbf{E} = \mathbf{E}(x,y,z,t)$, $\mathbf{x} = (x,y,z,t)$ is the (time-varying) position vector, q is the elementary charge, m the effective mass, $\omega_r := \sqrt{K/m}$ the resonance (angular) frequency, K the constant in Hooke's law and γ^{-1} the average time interval between two collisions (the *relaxation time*).

By associating a dipole moment to each atom and letting N denote the number of dipoles per unit volume, it is possible to define the polarization density vector field

$$\mathbf{P}(x,y,z,t) = qN\mathbf{x}(x,y,z,t). \tag{2.26}$$

The Fourier transform of the last two equations yields [6]

$$\begin{aligned} \overline{\mathbf{x}}(x,y,z,\omega) &= \frac{1}{\omega_r^2 - \omega^2 + i\omega\gamma}\frac{q}{m}\overline{\mathbf{E}}(x,y,z,\omega), \\ \overline{\mathbf{P}}(x,y,z,\omega) &= qN\overline{\mathbf{x}}(x,y,z,\omega), \end{aligned} \tag{2.27}$$

and leads to the electric constitutive relations

$$\begin{aligned} \overline{\mathbf{P}}(x,y,z,\omega) &= \epsilon_0 \ \frac{\omega_p^2}{\omega_r^2 - \omega^2 + i\omega\gamma} \ \overline{\mathbf{E}}(x,y,z,\omega), \\ \overline{\mathbf{D}}(x,y,z,\omega) &= \epsilon(\omega) \ \overline{\mathbf{E}}(x,y,z,\omega) = \epsilon_0\epsilon_\infty\overline{\mathbf{E}}(x,y,z,\omega) + \overline{\mathbf{P}}(x,y,z,\omega), \\ \epsilon(\omega) &= \epsilon_0\left(\frac{\omega_p^2}{\omega_r^2 - \omega^2 + i\omega\gamma} + \epsilon_\infty\right) = \epsilon'(\omega) - i\epsilon''(\omega), \end{aligned} \tag{2.28}$$

where $\overline{\mathbf{D}}$ denotes the Fourier transforms of the electric displacement field, $\omega_p = \sqrt{\dfrac{q^2N}{m\epsilon_0}}$ is called the *plasma angular frequency*; ϵ_∞ is introduced to model the contribution of free electrons to polarization.[7]

The imaginary part of the complex permittivity accounts for the energy lost in the process of establishing a net dipole moment density in the material [KCK06]. The resonance condition $\omega = \omega_r$ defines an absorption peak, i.e. a peak in interband transition probability.

Substituting (2.26) into the equation of motion or inverse-transforming the **P**-**E** relation from (2.28) yields the set of ordinary differential equations

$$\begin{aligned} \frac{\partial\mathbf{P}}{\partial t} &= \mathbf{J}_l, \\ \frac{\partial\mathbf{J}_l}{\partial t} &= \epsilon_0\omega_p^2\mathbf{E} - \gamma\mathbf{J}_l - \omega_r^2\mathbf{P}, \end{aligned} \tag{2.29}$$

which introduces the polarization current density $\mathbf{J}_l$ (to be added to $\mathbf{J}_n$ and $\mathbf{J}_p$ in the Ampère-Maxwell law in order to couple (2.29) to Maxwell's equations).[8] The classical counterpart to (2.24) is then

$$P_{abs}(x,y,z,t) = \mathbf{J}_l(x,y,z,t)\cdot\mathbf{E}(x,y,z,t), \tag{2.30}$$

and expresses that the power ceded by the electromagnetic wave to the medium corresponds to the work done by the electric field throughout the displacement of the charges constituting the polarization current [Gra16].

[6] The Fourier transform of a signal $x(t)$ is here defined as $\overline{x}(\omega) = \mathfrak{F}[x](\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t}dt$.

[7] The ∞ subscript can be justified following [Mai07], which introduces it to account for residual polarization due to free electrons in noble metals excited at $\omega > \omega_p$. The parameter is used to fit experimental data of semiconductors as well [SWM17].

[8] The l subscript is introduced to distinguish polarization current from drift and diffusion ones. No subscript is attached to **P** as there is no risk of ambiguity.

The fact of deriving (2.29) from a harmonic oscillator should suggest to expect that some of the power absorbed by the system at time t be stored in the form of energy (to sustain oscillations). The point can be better understood in the monochromatic case, following a classical argument from linear circuit theory [CDK87]. Let us omit space dependence for brevity and let $\mathbf{E} = \mathbf{E}_0 \cos(\omega t + \Phi_E)$. Linearity of the dispersion model implies $\mathbf{J}_l = \mathbf{J}_0 \cos(\omega t + \Phi_J)$ and $\mathbf{P} = \mathbf{P}_0 \cos(\omega t + \Phi_P)$ for some $\mathbf{J}_0, \mathbf{P}_0, \Phi_J, \Phi_P$. The resulting instantaneous absorbed power (per unit volume) is

$$P_{abs}(t) = \mathbf{E}_0 \cdot \mathbf{J}_0 \frac{\cos(\Phi_E - \Phi_J) + \cos(2\omega t + \Phi_E + \Phi_J)}{2}. \tag{2.31}$$

In the sinusoidal regime, a quantity of fundamental interest is the time average of the instantaneous power delivered to a system in a period $T = 2\pi/\omega$, which represents the power dissipated in a period and is known as the *active power*. Applying this definition to the absorbed power (2.31) yields

$$\hat{P} := \frac{1}{T} \int_0^T P_{abs}(t)\ dt = \frac{\mathbf{E}_0 \cdot \mathbf{J}_0}{2} \cos(\Phi_E - \Phi_J), \tag{2.32}$$

which identifies the active absorbed power with the static component of $P_{abs}(t)$. This should be expected as the time-varying component of $P_{abs}(t)$ has a time-varying sign as well, hence cannot describe such an irreversible process as dissipation.

The salient point is: our quest for real-time optical absorption requires isolating the portion of (2.30) that models dissipation. This is straightforward when an explicit expression like (2.31) is available but a more general procedure should be conceived. The natural way forward is to seek, based on (2.29), a power (density) balance of the form

$$P_{abs}(t) = P_{diss}(t) + \frac{\partial W(t)}{\partial t}, \tag{2.33}$$

where P_{diss} denotes the instantaneous dissipated power per unit volume and W is the energy per unit volume stored in the oscillator. The key is to multiply the second equation of (2.29) by $\mathbf{J}_l$,

$$\mathbf{J}_l \cdot \frac{\partial \mathbf{J}_l}{\partial t} = \epsilon_0 \omega_p^2 \mathbf{J}_l \cdot \mathbf{E} - \gamma \mathbf{J}_l \cdot \mathbf{J}_l - \omega_r^2 \mathbf{J}_l \cdot \mathbf{P}. \tag{2.34}$$

then, using the first equation of (2.29), (2.34) is recast as

$$\begin{aligned} \epsilon_0 \omega_p^2\ P_{abs} &= \mathbf{J}_l \cdot \frac{\partial \mathbf{J}_l}{\partial t} + \gamma \mathbf{J}_l \cdot \mathbf{J}_l + \omega_r^2 \frac{\partial \mathbf{P}}{\partial t} \cdot \mathbf{P} \\ &= \frac{1}{2} \frac{\partial (\mathbf{J}_l \cdot \mathbf{J}_l)}{\partial t} + \gamma \mathbf{J}_l \cdot \mathbf{J}_l + \omega_r^2 \frac{1}{2} \frac{\partial (\mathbf{P} \cdot \mathbf{P})}{\partial t}, \end{aligned} \tag{2.35}$$

and finally

$$P_{abs} = \frac{\gamma}{\epsilon_0 \omega_p^2} \mathbf{J}_l \cdot \mathbf{J}_l + \frac{\partial}{\partial t} \frac{1}{2\epsilon_0 \omega_p^2} \Big[\mathbf{J}_l \cdot \mathbf{J}_l + \omega_r^2 \mathbf{P} \cdot \mathbf{P} \Big]. \tag{2.36}$$

To fully appreciate (2.36) it is better to restate it in terms of the state variables of the harmonic oscillator – the position $\mathbf{x}$ and the velocity $\dfrac{\partial \mathbf{x}}{\partial t}$. From (2.26) and the first equation of (2.29),

$$\mathbf{J}_l = qN \frac{\partial \mathbf{x}}{\partial t}, \tag{2.37}$$

and in turn

$$\frac{\gamma}{\epsilon_0\omega_p^2}\mathbf{J}_l\cdot\mathbf{J}_l=\frac{\gamma}{\epsilon_0\dfrac{q^2N}{m\epsilon_0}}\mathbf{J}_l\cdot\mathbf{J}_l=\frac{\gamma m}{q}\frac{\partial\mathbf{x}}{\partial t}\cdot qN\frac{\partial\mathbf{x}}{\partial t}=\gamma mN\ \frac{\partial\mathbf{x}}{\partial t}\cdot\frac{\partial\mathbf{x}}{\partial t}. \tag{2.38}$$

Hence, the first term at the right-hand side of (2.36) is the scalar product of the frictional force per unit volume and the velocity, i.e. the instantaneous power (per unit volume) dissipated by the external force to impose oscillation against friction. Its sign is constant and non-negative as required.
Similarly,

$$\begin{aligned}\frac{1}{2\epsilon_0\omega_p^2}\mathbf{J}_l\cdot\mathbf{J}_l&=N\ \frac{1}{2}\ m\ \frac{\partial\mathbf{x}}{\partial t}\cdot\frac{\partial\mathbf{x}}{\partial t},\\ \frac{1}{2\epsilon_0\omega_p^2}\omega_r^2\mathbf{P}\cdot\mathbf{P}&=\frac{1}{2\epsilon_0\dfrac{q^2N}{m\epsilon_0}}\frac{K}{m}\ q^2N^2\ \mathbf{x}\cdot\mathbf{x}=N\ \frac{1}{2}\ K\ \mathbf{x}\cdot\mathbf{x},\end{aligned} \tag{2.39}$$

are the kinetic and potential energy (per unit volume) respectively.
It is thus concluded that (2.36) expresses the sought power (density) balance with

$$\begin{aligned}P_{diss}&=\frac{\gamma}{\epsilon_0\omega_p^2}\mathbf{J}_l\cdot\mathbf{J}_l,\\ W&=\frac{1}{2\epsilon_0\omega_p^2}\Big[\mathbf{J}_l\cdot\mathbf{J}_l+\omega_r^2\mathbf{P}\cdot\mathbf{P}\Big].\end{aligned} \tag{2.40}$$

and that the first term (P_{diss}) should be responsible for optical absorption.

For a consistency check, it is interesting to go back to the monochromatic case and calculate the active power. Time-averaging both sides of (2.36) and bearing in mind the periodicity of $\mathbf{J}_l$ and $\mathbf{P}$,

$$\begin{aligned}\hat{P}=\frac{1}{T}\int_0^T P_{abs}\ dt&=\frac{1}{T}\frac{\gamma}{\epsilon_0\omega_p^2}\int_0^T\mathbf{J}_l\cdot\mathbf{J}_l\ dt+\frac{1}{T}\frac{1}{2\epsilon_0\omega_p^2}\int_0^T d\Big[\mathbf{J}_l\cdot\mathbf{J}_l+\omega_r^2\mathbf{P}\cdot\mathbf{P}\Big]\\ &=\frac{1}{T}\frac{\gamma}{\epsilon_0\omega_p^2}\int_0^T\mathbf{J}_l\cdot\mathbf{J}_l\ dt+\frac{1}{2\epsilon_0\omega_p^2}\Big[\mathbf{J}_l\cdot\mathbf{J}_l+\omega_r^2\mathbf{P}\cdot\mathbf{P}\Big]\Big|_0^T\\ &=\frac{1}{T}\frac{\gamma}{\epsilon_0\omega_p^2}\int_0^T\mathbf{J}_l\cdot\mathbf{J}_l\ dt,\end{aligned} \tag{2.41}$$

which is coherent with the interpretation.

The foregoing developments suggest that a sensible photo-generation function complying with the requirement of real-time updates based on the computed electromagnetic field is

$$G(x,y,z,t)=\ \eta_i\ \frac{\gamma}{\epsilon_0\omega_p^2}\ \frac{||\mathbf{J}_l(x,y,z,t)||^2}{h\nu_0}. \tag{2.42}$$

The internal efficiency η_i has been recovered from (2.19) to admit that some resonating dipoles may not release any electron ($\eta_i<1$). The classical notion (2.30) is therefore supplemented with a probabilistic concept belonging to the quantum framework, to make the picture slightly more realistic.

On the other hand, there is no need for the exponential attenuation term anymore, because electromagnetic wave propagation in a dispersive medium, i.e. with a complex wave vector, is now explicitly described by the mathematical model. Moreover the model naturally captures electric-field screening (the electric field is instantaneously affected by the change in the carrier densities and hence in the material's conductivity).[9]
The function proposed in (2.42) is also considered in [CB21b], although with a more concise approach. Interestingly, in the reference it is shown how a Poynting-like approximation – rather than one based on a dispersion model – yields appreciably different results and overestimates the photo-generation rate.

An important concluding remark regards the practical use of (2.28-2.29). The model's parameters are tuned in order to fit experimental data, which should position ω_r in the proximity of the absorption peak of the material. However several peaks can exist in general, and a single oscillator may not provide enough degrees of freedom to achieve sufficient adherence to the reference curves. For a reference on tailoring semiconductor dispersion models to wide frequency ranges, the reader is suggested to look at [SWM17]. A general description of dispersion models coupled to Maxwell's equations in the finite-difference framework can be found in [TH05]. The topic is also discussed in [Viq15] with focus on DGTD methods applied to plasmonics.
Multiple oscillators are likely to be required in broad-band simulations. In narrow-band simulations, on the other hand, what matters most is adherence in the proximity of the carrier frequency. The scope of the present work is limited to the latter situation, and one oscillator will appear to be sufficient in practice. To retrieve $\epsilon_\infty, w_p, w_r, \gamma$, the fitting method from [Viq18] and available in DIOGENeS [Atl] has been used. The incorporation of (2.42) in (2.1) will be formally carried out in Section 2.4.

2.3.3 Mobility

Mobility ($\mathrm{m^2\ V^{-1}\ s^{-1}}$) relates drift velocity to the electric field: $\boldsymbol{v}_n = -\mu_n \mathbf{E}$, $\boldsymbol{v}_p = \mu_p \mathbf{E}$. It models the fact that carriers are accelerated in a crystal and therefore can lose momentum due to various phenomena, such as lattice vibrations (lattice scattering), attraction/rejection by ions of dopant (ionized impurity scattering), a rough interface between two materials (surface scattering).
Mobility has the following expression, which stems from the Drude model of free electrons [AM76]:

$$\mu_n = \frac{q\tau_{cn}}{m_n^*}, \qquad \mu_p = \frac{q\tau_{cp}}{m_p^*}, \tag{2.43}$$

where τ_{cn} (τ_{cp}) denotes the average time between collisions and m_n^* (m_p^*) the effective mass, which is a function of energy band curvature.

Different scattering mechanisms are described by dedicated, tailored models. It is common to assume them statistically independent, so that they concur per Matthiessen's rule [RA15],[AM76], which states

$$\frac{1}{\mu_n} = \frac{1}{\mu_{n,lattice}} + \frac{1}{\mu_{n,ion}} + \frac{1}{\mu_{n,surf}} \tag{2.44}$$

for electrons, then similarly for holes. The rationale is that the mechanism happening at the highest rate should dominate, i.e. be the one to limit drift the most.

[9] It may be worthwhile to note that (2.42) has the required dimension of $\mathrm{m^{-3}\ s^{-1}}$, as it is the ratio of volumic power density ($\mathrm{W\ m^{-3}}$) over energy (J), times a dimensionless quantity (η_i).

Evidently, mobility varies with temperature and doping. Its degradation at high electric field values is also of great relevance.
A quite exhaustive hierarchy of models is presented in [VG06]. In the present context, special attention is devoted to field-dependent mobility due to the very strong electrostatic field existing in photo-conductive antennas, and it is assumed that

$$\begin{aligned}\mu_n(E) &= \frac{\mu_{n0}}{\left(1+\left(\frac{E\ \mu_{n0}}{v_{n,sat}}\right)^{\beta_n}\right)^{\frac{1}{\beta_n}}},\\ \mu_p(E) &= \frac{\mu_{p0}}{\left(1+\left(\frac{E\ \mu_{p0}}{v_{p,sat}}\right)^{\beta_p}\right)^{\frac{1}{\beta_p}}},\end{aligned} \tag{2.45}$$

which is known as the *Caughey-Thomas model* [CT67]. The only difference between the two equations lies in the values assigned to the parameters, thus it suffices to analyze the first one. For low values of $E := ||\mathbf{E}||$, μ_n can be approximated by the low-field value μ_{n0} and $\boldsymbol{v}_n = -\mu_{n0}\mathbf{E}$. As E increases, the velocity reaches a saturation amplitude $v_{n,sat}$; at this point, further increasing E yields nonlinear effects. The fitting parameters β_n, β_p are positive and depend on the material.

2.3.4 Diffusivity

Diffusivity ($\mathrm{m}^2\ \mathrm{s}^{-1}$), also referred to as the *diffusion constant* or the *diffusion coefficient*, appears by virtue of Fick's law

$$\mathbf{J}_{n,diff} = D_n \nabla n, \qquad \mathbf{J}_{p,diff} = -D_p \nabla p, \tag{2.46}$$

which is readily recognized in (2.1) and models the fact that particles diffuse from regions of higher concentration to regions of lower concentration. The signs reflect the convention of $\mathbf{J}_p$ ($\mathbf{J}_n$) being directed as (opposite to) the flow.

As discussed in [AM76], electron and hole currents must vanish at thermodynamic equilibrium, that is

$$\mathbf{J}_{n,drift} + \mathbf{J}_{n,diff} = \mathbf{0} = \mathbf{J}_{p,drift} + \mathbf{J}_{p,diff}, \tag{2.47}$$

and it can be verified that this implies the *Einstein relations*

$$D_n = V_T \mu_n, \qquad D_p = V_T \mu_p, \tag{2.48}$$

where $V_T = \frac{k_B T}{q}$ is the thermal voltage, q is the elementary electric charge, k_B is the Boltzmann constant, T is the absolute temperature.

In the applications of our interest, doping is not high enough to violate the hypothesis, but it is worthwhile to mention that the presented form of (2.48) only holds if the semiconductor is non-degenerate. Otherwise, correction terms depending on carrier densities become necessary [Kan18].[10]

The Einstein relations are still valid when thermodynamic equilibrium is weakly perturbed [BSL+07]. At high fields, it becomes necessary to upgrade (and complexify) the $D - \mathbf{E}$ relation in order to account for hot-carrier diffusion [KCK06]. Several models based on different hypotheses

[10] A semiconductor is said to be non-degenerate if the Fermi level E_F is more than $3k_BT$ away from both the conduction and the valence band edges [VG06]. A fundamental concept in semiconductor physics is that N-type (P-type) doping shifts E_F towards the conduction (valence) band, thereby modifying the distance just mentioned.

on the Boltzmann transport equation have been derived in the literature [Sel84], [Jün09] but there seems to be no convergence on a particular one. A survey of textbooks on semiconductor device physics (e.g. [MK02], [SN06], [SB14], [RV04]) would show a tendency to consider high-field effects on drift (that is, velocity saturation as per (2.45)) while retaining (2.48). In [Hes00], for example, the issue is addressed; but it is also argued that adequate refinement of the drift-diffusion model to include hot-carrier diffusion might require a computational load comparable to Monte Carlo simulations of the Boltzmann transport equation itself.
All such reasons should justify the spotlight on several mobility models, while not on high-field diffusivity, in commercial software (e.g. [Sil], [COM]). A decision has been made in the present work to align with such trend, in accordance with the literature on photo-conductive antenna simulation (e.g. [MPG+14], [Bur16]).

2.4 Device biasing

A fundamental concept in semiconductor device theory is that of static (or DC) operating point – a static input is applied with the aim of enforcing a convenient steady state, an identified optimal setting in which to apply a time-varying (or AC) input. In the enforced steady state, the device is said to be *biased*.
We are mostly interested in devices that are biased electrically, by means of a voltage applied via metal contacts, and shall assume that no static magnetic field is applied.
Once the device is biased, the unknowns of our model are decomposed in static (DC) and time-varying (AC) parts:

$$\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \\ n \\ p \end{bmatrix} (x,y,z) = \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \\ n \\ p \end{bmatrix}_{DC} (x,y,z) + \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \\ n \\ p \end{bmatrix}_{AC} (x,y,z,t), \tag{2.49}$$

and similarly the auxiliary variables $\boldsymbol{\alpha}, \boldsymbol{\beta}$.
We shall assume the DC quantities assigned or numerically calculated by a separate solver. The interest here lies in understanding their physical meaning and exploiting the consequences of (2.49) on the MDD model.[11]

2.4.1 The static Poisson-Drift-Diffusion model

$\mathbf{E}_{DC}, n_{DC}, p_{DC}$ constitute the solution to the electrostatic problem

[11] In the context of this section the dispersion model introduced earlier does not play a relevant role, thus we find it convenient to omit $\mathbf{P}$, $\mathbf{J}_l$ in (2.49). Their static and dynamic components will be introduced in Section 2.5.

$$
\begin{aligned}
\nabla \times \mathbf{E} &= \mathbf{0}, \\
\nabla \cdot \epsilon \mathbf{E} &= 0, \\
&\text{in } \Omega_e \setminus \Omega, \\
\\
\nabla \times \mathbf{E} &= \mathbf{0}, \\
\nabla \cdot \epsilon \mathbf{E} &= q(p - n + C), \\
\nabla \cdot \mathbf{J}_n - qR' &= 0, \\
-\nabla \cdot \mathbf{J}_p - qR' &= 0, \\
\mathbf{J}_n &= -qn\mu_n \nabla \phi + qD_n \nabla n, \\
\mathbf{J}_p &= -qp\mu_p \nabla \phi - qD_p \nabla p, \\
&\text{in } \Omega,
\end{aligned}
\tag{2.50}
$$

whose resolution usually relies on the introduction of an electrostatic potential $\phi : \bar{\Omega}_e \to \mathbb{R}$ such that $\mathbf{E} = -\nabla\phi$, yielding

$$
\begin{aligned}
\nabla \cdot (\epsilon \nabla \phi) &= 0, \\
&\text{in } \Omega_e \setminus \Omega, \\
\\
\nabla \cdot (\epsilon \nabla \phi) &= -q(p - n + C), \\
\nabla \cdot \mathbf{J}_n - qR' &= 0, \\
-\nabla \cdot \mathbf{J}_p - qR' &= 0, \\
\mathbf{J}_n &= -qn\mu_n \nabla \phi + qD_n \nabla n, \\
\mathbf{J}_p &= -qp\mu_p \nabla \phi - qD_p \nabla p, \\
&\text{in } \Omega.
\end{aligned}
\tag{2.51}
$$

The boundary conditions read

$$
\begin{aligned}
n &= n_D && \text{on } \Gamma_D, \\
p &= p_D && \text{on } \Gamma_D, \\
\phi &= V_{bias} + \phi_{bi} && \text{on } \Gamma_D, \\
\hat{\mathbf{n}} \cdot \mathbf{J}_n &= 0 && \text{on } \Gamma_N, \\
\hat{\mathbf{n}} \cdot \mathbf{J}_p &= 0 && \text{on } \Gamma_N, \\
\hat{\mathbf{n}} \cdot \nabla \phi &= 0 && \text{on } \partial\Omega_e,
\end{aligned}
\tag{2.52}
$$

with n_D, p_D given by (2.7), V_{bias} being the applied voltage and $\phi_{bi} = V_T \log(n_D)$ the *built-in potential* – the potential barrier that spontaneously arises at the metal-semiconductor interfaces [Mar86]. The homogeneous Neumann condition on $\partial\Omega_e$ is meant to approximate the asymptotic behavior of ϕ and forces electrostatic field lines to be tangent to the boundary of the computational domain, consistently with the absence of electric charges at infinity.
Note that in the resolution of (2.51-2.52) one also enforces the continuity of the normal component of the electric displacement field $\hat{\mathbf{n}} \cdot \mathbf{D} = -\hat{\mathbf{n}} \cdot \epsilon \, \nabla\phi$ on Γ_N.

The currents calculated in (2.51) are expected to produce a stationary magnetic field $\mathbf{H}_{DC}$

according to

$$\begin{aligned} \nabla \times \mathbf{H} &= \mathbf{0}, \\ \nabla \cdot \mu \mathbf{H} &= 0, \\ \text{in } \Omega_e \setminus \Omega, & \\ \nabla \times \mathbf{H} &= \mathbf{J}_n + \mathbf{J}_p, \\ \nabla \cdot \mu \mathbf{H} &= 0, \\ \text{in } \Omega, & \end{aligned} \tag{2.53}$$

with $\hat{\mathbf{n}} \times \mathbf{H}$ to be assigned on the boundary of the computational domain, however it is apparent that $\mathbf{H}_{DC}$ plays no role in (2.51-2.52): in practice, (2.53) does not need to be solved.

The absence of G in (2.51) deserves a remark. It was outlined in Section 2.3.2.2 that (2.51) is solved *with* G in Silvaco Atlas, but this is a particular implementation choice that allows to couple the optical solver with the semiconductor one. The general rule is to employ the static PDD system in cases where no light is involved.

2.4.2 The quasi-static Poisson-Drift-Diffusion model

An alternative route to $\mathbf{E}_{DC}, n_{DC}, p_{DC}$ is to see these as the steady state asymptotically reached by a modified version of (2.1) based on an electro-quasi-static approximation [VG06], sometimes referred to as the *van Roosbroeck system*.[12] The difference with respect to (2.51) is that the electric field obeys the laws of electrostatics but is time-varying, hence derived from a time-varying potential $\phi : \bar{\Omega}_e \times [0, +\infty[\to \mathbb{R}$. The resulting system is

$$\begin{aligned} \nabla \cdot (\epsilon \nabla \phi) &= 0, \\ & \text{in } \Omega_e \setminus \Omega \times [0, \infty[, \\ \nabla \cdot (\epsilon \nabla \phi) &= -q(p - n + C), \\ q\frac{\partial n}{\partial t} &= \nabla \cdot \mathbf{J}_n - qR, \\ q\frac{\partial p}{\partial t} &= -\nabla \cdot \mathbf{J}_p - qR, \\ \mathbf{J}_n &= -qn\mu_n \nabla \phi + qD_n \nabla n, \\ \mathbf{J}_p &= -qp\mu_p \nabla \phi - qD_p \nabla p, \\ & \text{in } \Omega \times [0, +\infty[. \end{aligned} \tag{2.54}$$

The boundary conditions have the same form as (2.52) except that now $V_{bias} = V_{bias}(t)$. Calculating the steady state by means of (2.54) could be motivated by an interest in monitoring the transient behavior between thermodynamic equilibrium and the desired DC operating point. In this case, (2.51) would have to be solved for $V_{bias} = 0$ V, to provide the initial conditions for the transient simulation. This first step is common to (2.51) and (2.54), then resolution strategies diverge:

- (2.51) is recast as a single non-linear equation in the unknown ϕ, and solved with an iterative method for increasing values of V_{bias};

[12] After W. W. van Roosbroeck, who first formulated it in 1950. The original paper is here referred to as [vR50].

- (2.54) is solved in the time domain with a time-varying V_{bias}.

It should be noted that initial and boundary conditions depend on the particular application and, more importantly, that there exist several situations in which a transient response itself is of primary interest (rather than the calculation of a subsequent steady state). One example has been outlined in Section 2.3, with reference to [Bur16], in which (2.54) is solved to observe THz emission resulting from a a pre-defined photo-generation function.
More generally speaking, the electro-quasi-static approximation of Maxwell's equations has proven to be relevant in nanophotonics, in the study of plasmon resonances in nanoparticles [May13].[13]

2.4.3 DC and AC components in the Maxwell-Drift-Diffusion model

The introduction of (2.49) in the full-wave model yields two types of terms: time-varying (containing AC quantities only or AC quantities multiplied by DC quantities) and purely static (containing DC quantities only). Two key advantages stem.

- Purely static terms can be simplified because they solve a self-consistent set of equations describing the pre-calculated DC operating point. They don't need to be coded in the transient solver.

- In applications of our interest typically $||\mathbf{E}_{DC}|| >> ||\mathbf{E}_{AC}||$. This allows $f(||\mathbf{E}||) \approx f(||\mathbf{E}_{DC}||)$ for $f = \mu_n, \mu_p, D_n, D_p$, thereby removing a major source of non-linearity and computational complexity.
  To keep the notation compact field dependence of mobility and diffusivity will be omitted in the formulae, unless strictly necessary.

The resulting equations are presented below, with the purely stationary terms in opaque red (their time derivatives are omitted).

[13] As the lowest-order term in a series of models derived with a perturbation technique. It should be noted that the reference also highlights the importance of higher-order corrections, hence of a full-wave approach.

$$
\begin{aligned}
\nabla \times \mathbf{E}_{AC} + \nabla \times \mathbf{E}_{DC} &= -\mu \frac{\partial \mathbf{H}_{AC}}{\partial t}, \\
\nabla \times \mathbf{H}_{AC} + \nabla \times \mathbf{H}_{DC} &= \epsilon \frac{\partial \mathbf{E}_{AC}}{\partial t}, \\
&\text{in } \Omega_e \setminus \Omega \times [0, +\infty[, \\
\\
\nabla \times \mathbf{E}_{AC} + \nabla \times \mathbf{E}_{DC} &= -\mu \frac{\partial \mathbf{H}_{AC}}{\partial t}, \\
\nabla \times \mathbf{H}_{AC} + \nabla \times \mathbf{H}_{DC} &= \epsilon \frac{\partial \mathbf{E}_{AC}}{\partial t} + q\ n_{AC}\ \mu_n(\mathbf{E}_{AC} + \mathbf{E}_{DC}) + qD_n \boldsymbol{\alpha}_{AC} \\
&\quad + q\ p_{AC}\ \mu_p(\mathbf{E}_{AC} + \mathbf{E}_{DC}) - qD_p \boldsymbol{\beta}_{AC} \\
&\quad + q\ n_{DC}\ \mu_n \mathbf{E}_{AC} + q\ p_{DC}\ \mu_p \mathbf{E}_{AC} \\
&\quad + q\ n_{DC}\ \mu_n \mathbf{E}_{DC} + qD_n \nabla n_{DC} + q\ p_{DC}\ \mu_p \mathbf{E}_{DC} - qD_p \nabla p_{DC}, \\
\frac{\partial n_{AC}}{\partial t} &= \nabla \cdot (n_{AC}\ \mu_n(\mathbf{E}_{AC} + \mathbf{E}_{DC})) + \nabla \cdot (D_n \boldsymbol{\alpha}_{AC}) + G - R''(n_{AC}, p_{AC}) \\
&\quad + \nabla \cdot (n_{DC}\ \mu_n \mathbf{E}_{AC}) \\
&\quad + \nabla \cdot (n_{DC}\ \mu_n \mathbf{E}_{DC}) + \nabla \cdot (D_n \nabla n_{DC}) - R'(n_{DC}, p_{DC}), \\
\frac{\partial p_{AC}}{\partial t} &= -\nabla \cdot (p_{AC}\ \mu_p(\mathbf{E}_{AC} + \mathbf{E}_{DC})) + \nabla \cdot (D_p \boldsymbol{\beta}_{AC}) + G - R''(n_{AC}, p_{AC}) \\
&\quad - \nabla \cdot (p_{DC}\ \mu_p \mathbf{E}_{AC}) \\
&\quad - \nabla \cdot (p_{DC}\ \mu_p \mathbf{E}_{DC}) + \nabla \cdot (D_p \nabla p_{DC}) - R'(n_{DC}, p_{DC}), \\
\boldsymbol{\alpha}_{AC} &= \nabla n_{AC}, \\
\boldsymbol{\beta}_{AC} &= \nabla p_{AC}, \\
&\text{in } \Omega \times [0, +\infty[.
\end{aligned}
\tag{2.55}
$$

A similar reduction stems from injecting (2.49) in initial and boundary conditions, which become

$$
\begin{aligned}
n_{AC}(x, y, z, 0) &= 0 && (x, y, z) \in \bar{\Omega}, \\
p_{AC}(x, y, z, 0) &= 0 && (x, y, z) \in \bar{\Omega}, \\
n_{AC}(x, y, z, t) &= 0 && (x, y, z, t) \in \Gamma_D \times [0, +\infty[, \\
p_{AC}(x, y, z, t) &= 0 && (x, y, z, t) \in \Gamma_D \times [0, +\infty[, \\
\hat{\mathbf{n}} \cdot \mathbf{J}_{n,AC} &= 0 && (x, y, z, t) \in \Gamma_N \times [0, +\infty[, \\
\hat{\mathbf{n}} \cdot \mathbf{J}_{p,AC} &= 0 && (x, y, z, t) \in \Gamma_N \times [0, +\infty[,
\end{aligned}
\tag{2.56}
$$

for the semiconductor, where

$$
\begin{aligned}
\mathbf{J}_{n,AC} &= n_{AC}\ \mu_n\ (\mathbf{E}_{AC} + \mathbf{E}_{DC}) + D_n \boldsymbol{\alpha}_{AC} + n_{DC}\ \mu_n \mathbf{E}_{AC}, \\
\mathbf{J}_{p,AC} &= p_{AC}\ \mu_p\ (\mathbf{E}_{AC} + \mathbf{E}_{DC}) - D_p \boldsymbol{\beta}_{AC} + p_{DC}\ \mu_p \mathbf{E}_{AC}.
\end{aligned}
\tag{2.57}
$$

For the electromagnetic field, depending on whether the problem is interior or exterior, one enforces

$$
\begin{aligned}
\mathbf{E}_{AC}(x,y,z,0) &= \mathbf{0} && (x,y,z) \in \bar{\Omega}, \\
\mathbf{H}_{AC}(x,y,z,0) &= \mathbf{0} && (x,y,z) \in \bar{\Omega}, \\
\hat{\mathbf{n}} \times \mathbf{E}_{AC}(x,y,z,t) &= 0 && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{H}_{AC}(x,y,z,t) &= \hat{\mathbf{n}} \times \mathbf{H}_{N,AC}(x,y,z,t) && (x,y,z,t) \in \Gamma_N \times [0,+\infty[,
\end{aligned}
\tag{2.58}
$$

or

$$
\begin{aligned}
\mathbf{E}_{AC}(x,y,z,0) &= \mathbf{E}^i_{AC}(x,y,z,0) && (x,y,z) \in \bar{\Omega}_e, \\
\mathbf{H}_{AC}(x,y,z,0) &= \mathbf{H}^i_{AC}(x,y,z,0) && (x,y,z) \in \bar{\Omega}_e, \\
\hat{\mathbf{n}} \times \mathbf{E}_{AC}(x,y,z,t) &= 0 && (x,y,z,t) \in \Gamma_D \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{E} + \hat{\mathbf{n}} \times \sqrt{\frac{\mu}{\epsilon}}(\hat{\mathbf{n}} \times \mathbf{H}) &= \hat{\mathbf{n}} \times \mathbf{E}^i + \hat{\mathbf{n}} \times \sqrt{\frac{\mu}{\epsilon}}(\hat{\mathbf{n}} \times \mathbf{H}^i) && \text{on } \Gamma_e \times [0,+\infty[.
\end{aligned}
\tag{2.59}
$$

Such conditions are motivated as follows.

- n_{AC}, p_{AC} represent the excess of electron-hole pairs with respect to the initial stationary distributions n_{DC}, p_{DC}. Such excess is assumed to be photo-generated by pulsed illumination at some $t > 0$. It is therefore null at $t = 0$ and (asymptotically) collapses back to zero, as light fades, under the effect of recombination (2.18). As a result, initial conditions on n_{AC}, p_{AC} are homogeneous.
- Since (2.7) are static, n_{AC}, p_{AC} are zero on Γ_D.
- By the same argument, the conditions on Γ_N are unchanged, consistently with the requirement that charge carriers not leave the semiconductor.
- As just anticipated, it is assumed that the semiconductor is hit by the incident field at $t > 0$ so that the scattered field $(\mathbf{E}^s_{AC}, \mathbf{H}^s_{AC})$ and the total one inside the semiconductor $(\mathbf{E}^t_{AC}, \mathbf{H}^t_{AC})$ are initially null.
- Assuming perfect electric conductors, the tangential component of the time-varying electric field is null on Γ_D.
- The tangential component of the time-varying magnetic field is imposed on Γ_N for an interior problem (as shown in practice in Section 4.2). For a scattering one, this is implicitly replaced by transmission conditions, and absorbing boundary conditions are enforced on $\mathbf{E}_{AC}, \mathbf{H}_{AC}$, to which the decomposition (2.9) applies.

The ultimate target of the developed solver is the purely time-varying version of (2.1), that is (2.55) after simplification of the opaque terms. Nevertheless, for the sake of clarity and readability, the decomposition (2.49) will not be unfolded in the forthcoming sections unless otherwise stated – with no loss of generality, the fact that the equations contain terms to be simplified will remain implied.

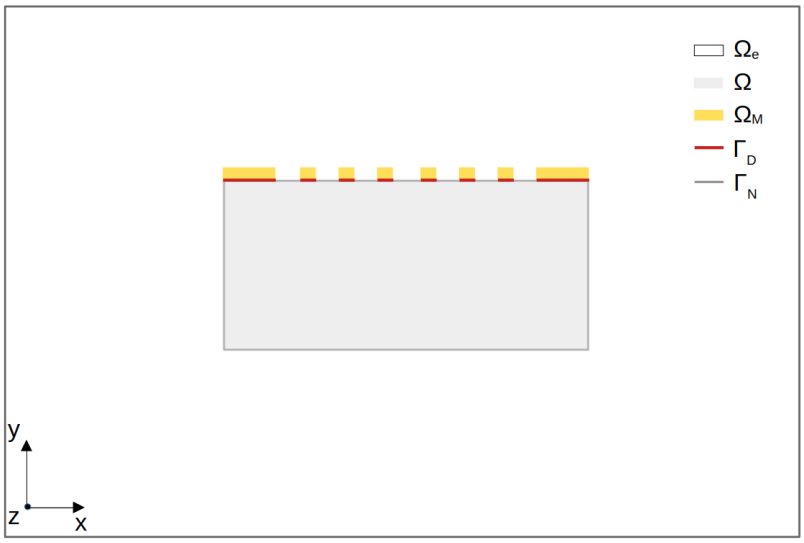


Figure 2.4: The semiconductor domain Ω now contacting a collection of dispersive metal regions Ω_M via Ohmic interfaces $\Gamma_D = \partial\Omega_M \cap \partial\Omega$.

2.5 Dispersion

It was pointed out in Section 2.3 that full coupling in (2.1) is achieved when G is calculated at the same time as the electromagnetic field, ideally via a dispersion model for the semiconductor. Aside from that, simulation of plasmon-enhanced devices demands to lift the original setting in which Γ_D is a collection of perfectly electrically conducting faces, and let metals be imperfect conductors of finite-depth, typically modeled as dispersive dielectrics.

The new setting is depicted in Figure 2.4. Electrodes and in-between electrically floating plasmonic *isles* are represented by finite-depth (yellow) regions constituting Ω_M. Γ_D, marked in red, collects the interfaces: $\Gamma_D = \partial\Omega_M \cap \partial\Omega$. These remain Ohmic contacts, hence Dirichlet boundary conditions on n and p are not affected; to the electromagnetic field, on the other hand, they become ordinary internal mesh faces on which to enforce continuity of tangential components.

Regardless of the specific material, the usual approach to dispersion consists in defining a complex permittivity based on Drude and Drude-Lorentz models, let it be a Padé approximant of experimental data, and write a set of time-domain auxiliary differential equations to be solved along with Maxwell's ones. The interested reader is referred to [Viq15] for more details – here, the final results are stated.

As anticipated in Section 2.3, one Lorentz pole is sufficient to our purposes. Then, if it is assumed that oscillations of the electron gas inside the metallic regions do not trigger interband transitions, a simple one-pole Drude model is acceptable as well.

The Lorentz dispersion model for the semiconductor is precisely (2.28-2.29) from Section 2.3. The Drude model is analogous, except that: the high-frequency relative dielectric constant ϵ_∞ generally has a different value; $\omega_r = 0$, meaning no interband transitions; the plasma frequency

and the reciprocal of the average time between collisions are denoted by ω_d and γ_d respectively. For $(x, y, z) \in \Omega_M$ the constitutive relations in the frequency domain read

$$\begin{aligned}
\overline{\mathbf{D}}(x, y, z, \omega) &= \epsilon(\omega)\ \overline{\mathbf{E}}(x, y, z, \omega) = \epsilon_0 \epsilon_\infty \mathbf{E}(x, y, z, \omega) + \overline{\mathbf{P}}(x, y, z, \omega), \\
\overline{\mathbf{P}}(x, y, z, \omega) &= \epsilon_0\ \frac{\omega_d^2}{-\omega^2 + i\omega\gamma_d}\ \overline{\mathbf{E}}(x, y, z, \omega), \\
\epsilon(\omega) &= \epsilon_0 \left(\frac{\omega_d^2}{-\omega^2 + i\omega\gamma_d} + \epsilon_\infty \right) = \epsilon'(\omega) - i\epsilon''(\omega),
\end{aligned} \tag{2.60}$$

and their time-domain counterpart is

$$\begin{aligned}
\frac{\partial \mathbf{P}}{\partial t} &= \mathbf{J}_d, \\
\frac{\partial \mathbf{J}_d}{\partial t} &= \epsilon_0 \omega_d^2 \mathbf{E} - \gamma_d \mathbf{J}_d.
\end{aligned} \tag{2.61}$$

In fact, $\mathbf{P}$ is not essential and $\mathbf{J}_d$ can be designated as the only relevant unknown.

The enhanced MDD system modeling scattering from a plasmon-enhanced semiconductor device with real-time photo-generation is assembled as follows:

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t}, \\
&\qquad \text{in } \Omega_e \setminus \Omega \times [0, +\infty[, \\
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon_0 \epsilon_\infty \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_d, \\
\frac{\partial \mathbf{J}_d}{\partial t} &= \epsilon_0 \omega_d^2 \mathbf{E} - \gamma_d \mathbf{J}_d, \\
&\qquad \text{in } \Omega_M \times [0, +\infty[, \\
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon_0 \epsilon_\infty \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_n + \mathbf{J}_p + \mathbf{J}_l, \\
q \frac{\partial n}{\partial t} &= \nabla \cdot \mathbf{J}_n + qG - qR, \\
q \frac{\partial p}{\partial t} &= -\nabla \cdot \mathbf{J}_p + qG - qR, \\
\frac{\partial \mathbf{P}}{\partial t} &= \mathbf{J}_l, \\
\frac{\partial \mathbf{J}_l}{\partial t} &= \epsilon_0 \omega_p^2 \mathbf{E} - \gamma \mathbf{J}_l - \omega_r^2 \mathbf{P}, \\
&\qquad \text{in } \Omega \times [0, +\infty[,
\end{aligned} \tag{2.62}$$

where $G = G(\mathbf{J}_l)$ as per (2.42). Recalling that $\mu = \mu_0 \mu_r$, $\epsilon = \epsilon_0 \epsilon_r$, we point out that in applications of our interest, typically: $\mu_r = 1$ and ϵ_r (ϵ_∞ in case of dispersion) is material-wise constant.

The system (2.62) is closed with the initial and boundary conditions (2.11), after remotion of $\hat{\mathbf{n}} \times \mathbf{E}|_{\Gamma_D} = \mathbf{0}$, plus the following initial conditions:[14]

$$\begin{aligned} \mathbf{P}(x,y,z,0) &= \epsilon_0 \, \frac{\omega_p^2}{\omega_r^2} \, \mathbf{E}_{DC} && (x,y,z) \in \bar{\Omega}, \\ \mathbf{J}_l(x,y,z,0) &= \mathbf{0} && (x,y,z) \in \bar{\Omega}, \\ \mathbf{J}_d(x,y,z,0) &= \mathbf{0} && (x,y,z) \in \bar{\Omega}_M. \end{aligned} \tag{2.63}$$

The interpretation of (2.63) goes as follows. Drude and Lorentz unknowns can be decomposed as

$$\begin{aligned} \mathbf{P} &= \mathbf{P}_{DC} + \mathbf{P}_{AC}, \\ \mathbf{J}_l &= \mathbf{J}_{l,DC} + \mathbf{J}_{l,AC}, \\ \mathbf{J}_d &= \mathbf{J}_{d,DC} + \mathbf{J}_{d,AC}, \end{aligned} \tag{2.64}$$

with $\mathbf{J}_{l,DC} = \mathbf{J}_{d,DC} = \mathbf{0}$ (polarization current is time-varying by definition) and $\mathbf{P}_{DC} = \epsilon_0 \, \frac{\omega_p^2}{\omega_r^2} \, \mathbf{E}_{DC}$ derived from (2.29) to model the dielectric response of the semiconductor to the electrostatic field. Then, the AC part of (2.64) is required to be null at $t = 0$, according to the presupposition made in (2.59) that the device is not illuminated until some $t > 0$.

It should be remarked that discretizing the auxiliary differential equations is straightforward thanks to the absence of spatial derivatives. Hence, it is of no major interest to consider dispersion in the process of conceiving a suitable semi-discrete counterpart of (2.1) and the phenomenon will not be included until the task is accomplished.

2.6 Towards a priori energy bounds for the interior problem

In this section we propose a preliminary study on the solutions to the MDD model formulated in an interior problem, as a potential stepping stone to a future stability analysis on the developed DGTD scheme. This topic is closely related to the asymptotic behavior of the continuous MDD model and to the stationary state to which the latter should tend, which is described by a static interior PDD model. We start by deriving an a priori estimate for the continuous MDD system. The estimated quantity is the relative energy with respect to the DC operating point, i.e. an equilibrium state that is generally different from thermodynamic equilibrium and to which the semiconductor is expected to relax after the illumination window. We shall call such a state (n_e, p_e, φ_e) and momentarily lift the heavier DC subscript. Prior to dwelling on energy estimation, it is convenient to highlight some key points concerning the static PDD model – to which (n_e, p_e, φ_e) is a solution – and its link to the MDD one.[15]

2.6.1 Preliminary observations on the static PDD model

The static PDD model was presented in Section 2.4 as the set of equations (2.50) or (2.51), with the boundary conditions (2.52).

[14] Having considered dispersion in metals, Γ_D has become to the electromagnetic field a collection of internal edges, on which to impose the continuity of $\hat{\mathbf{n}} \times \mathbf{E}$ and $\hat{\mathbf{n}} \times \mathbf{H}$ rather than boundary conditions.

[15] As will be pointed out all along the computations, allowing (n_e, p_e, φ_e) to be the thermodynamic equilibrium state would simplify the analysis.

In that context we referred to the exterior problem set in Figure 2.2, with the computational domain Ω_e strictly containing the semiconductor Ω. Moreover, $\partial\Omega = \Gamma_D \cup \Gamma_N$ was assumed to be composed of metallic (Γ_D) and electrically insulating (Γ_N) disjoint segments.
Dealing with a theoretically infinite (and artificially truncated) domain was driven by the interest in physically meaningful simulations – the model was meant to describe the steady state of a photo-conductive device before and after illumination. In terms of boundary conditions, this led to

- the assignation of $\phi|_{\Gamma_D}$ according to an applied voltage,
- an approximate asymptotic condition, e.g. $\hat{\mathbf{n}} \cdot \nabla\phi|_{\partial\Omega_e} = 0$, on the artificial boundary ($\hat{\mathbf{n}}$ denoting the outward unit normal vector to $\partial\Omega_e$).

From a mathematical standpoint, the interest mostly lies in the coupling between the equations of electrostatics (or electrodynamics) and those describing charge transport in semiconductors; and since the coupling only occurs in Ω, it is common to restrict the attention to interior problems, even if – as pointed out in Section 2.1.1 – these may not have a rigorous physical meaning.

Switching to an interior problem entails a change in the boundary conditions on ϕ, namely

- the assignation of $\phi|_{\Gamma_D}$,
- the assignation of $\epsilon\hat{\mathbf{n}}\cdot\nabla\phi|_{\Gamma_N}$ ($\hat{\mathbf{n}}$ denoting this time the outward unit normal vector to $\partial\Omega_N$).[16]

As to the semiconductor, the conditions (2.52) are retained. The interior PDD model then reads

$$\begin{aligned}
\nabla\cdot(\epsilon\nabla\phi) &= 0, \\
&\text{in } \Omega_e \setminus \Omega, \\
\nabla\cdot(\epsilon\nabla\phi) &= -q(p-n+C), \\
\nabla\cdot\mathbf{J}_n - qR' &= 0, \\
-\nabla\cdot\mathbf{J}_p - qR' &= 0, \\
\mathbf{J}_n &= -qn\mu_n\nabla\phi + qD_n\nabla n, \\
\mathbf{J}_p &= -qp\mu_p\nabla\phi - qD_p\nabla p, \\
&\text{in } \Omega.
\end{aligned} \tag{2.65}$$

$$\begin{aligned}
n &= n_D && \text{on } \Gamma_D, \\
p &= p_D && \text{on } \Gamma_D, \\
\phi &= V_{bias} + \phi_{bi} && \text{on } \Gamma_D, \\
\hat{\mathbf{n}}\cdot\mathbf{J}_n &= 0 && \text{on } \Gamma_N, \\
\hat{\mathbf{n}}\cdot\mathbf{J}_p &= 0 && \text{on } \Gamma_N, \\
\hat{\mathbf{n}}\cdot\nabla\phi &= \sigma_0 && \text{on } \Gamma_N,
\end{aligned}$$

with n_D, p_D given by (2.7), $\phi_{bi} = V_T \log(n_D)$, and V_{bias} and σ_0 to be assigned.

The aforementioned state (n_e, p_e, φ_e) is by definition a solution to (2.65). Its existence of has been studied in [Mar86] in the case of $\sigma_0 = 0$; the detailed derivation is not reported, but it is proven

[16] Note that when switching to an exterior problem this condition translates to the continuity of $\epsilon\hat{\mathbf{n}}\cdot\nabla\phi$ across Γ_N.

that the system admits at least one solution, under suitable regularity hypotheses on the data. The uniqueness of the solution is obtained if the applied voltage V_{bias} is close to zero (i.e. (n_e, p_e, φ_e) is close to thermal equilibrium). The strategy of the existence proof is based on a transformation of the unknowns (from n, p to so-called *Slotboom variables*) which makes the divergence equations self-adjoint.

2.6.2 Well-posedness and asymptotic behavior of the MDD model

Let us consider the MDD model for an interior problem, composed of the equations (2.1) and the initial and boundary conditions (2.2).
Studies on well posedness and asymptotic behavior have been conducted by Jochmann [Joc96, Joc98, Joc99, Joc01, Joc05], who has established a link between the MDD and the PDD model in the following setting.

H1: The initial data from (2.2) are such that $n_0 \geq 0$, $p_0 \geq 0$ and $n_0 \in L^\infty(\Omega)$, $p_0 \in L^\infty(\Omega)$ and $\mathbf{E}_0, \mathbf{H}_0 \in L^2(\Omega)$.

H2: The boundary data are such that $n_D \in W^{1,\infty}(\partial\Omega), p_D \in W^{1,\infty}(\partial\Omega)$, and $\mathbf{H}_N \in W^{1,2}_{loc}(\mathbb{R}, L^2(\Omega))$ with $\mathrm{curl}(\mathbf{H}_N) \in L^2_{loc}(\mathbb{R}, L^2(\Omega))$.

H3: A pivotal role in the connection is played by

$$\epsilon \hat{\mathbf{n}} \cdot \mathbf{E}|_{\Gamma_N} = \sigma. \tag{2.66}$$

Indeed, it is assumed that there exists $\sigma \in W^{1,2}_{loc}(\mathbb{R}, L^2(\partial\Omega) \cap L^\infty(\mathbb{R}), L^{p_0}(\partial\Omega))$, with $p_0 > 2$, such that

$$\int_\Omega \mathrm{curl}(\mathbf{H}_N) \cdot \nabla\phi = \frac{d}{dt} \int_{\Gamma_N} \sigma(t)\ \phi, \quad \forall \phi \in C^\infty_0(\mathbb{R}^3 \setminus \Gamma_D), \tag{2.67}$$

compatibly with (2.3).

H4: Boundary and initial data must comply with the compatibility conditions (2.8), (2.3). A compatibility condition is defined for σ as well, which must be such that $\epsilon \hat{\mathbf{n}} \cdot \mathbf{E}_0|_{\Gamma_N} = \sigma(0)$.

H5: Mobility and diffusivity are considered as $L^\infty(\Omega)$ positive functions.

H6: It is supposed that ν, $\frac{\varepsilon}{\mu_n}$ and $\frac{\varepsilon}{\mu_p}$ belong to $W^{1,\infty}(\Omega)$.

H7: The recombination term $R : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}, (x, n_1, n_2) \mapsto R(x, n_1, n_2)$ is supposed to be locally Lipschitz continuous with respect to $n = (n_1, n_2)$ and is such that $R(x, n) \geq 0$, if $n_1 n_2 \geq n_i^2$ and $R(x, n) \leq 0$ if $n_1 n_2 \leq n_i^2$.

H8: $G \equiv 0$.

One of the fundamental results obtained by Jochmann is summarized as follows.

Theorem 1 ([Joc99])**.** *The problem* (2.1) *with the initial and boundary conditions (2.2), supplemented by H1 to H8, admits a global weak solution* $(\mathbf{E}, \mathbf{H}, n, p)$*, with*

$$n, p \in L^2_{loc}([0, +\infty[, H^1(\Omega)) \cap L^\infty_{loc}([0, +\infty[, L^\infty(\Omega)) \cap \mathcal{C}([0, +\infty[, L^2(\Omega))$$

and

$$(\mathbf{E}, \mathbf{H}) \in \mathcal{C}([0, +\infty[, L^2(\Omega)).$$

The equations for electron and hole densities are to be understood in a weak sense as follows

$$\int_0^{+\infty} \int_\Omega n \partial_t \varphi dx dt = \int_0^{\infty} \int_\Omega [D_n \nabla n - \mu_n n \mathbf{E}] \nabla \varphi dx dt + \int_0^{+\infty} \int_\Omega R(x, n, p) \varphi dx dt, \quad (2.68)$$

$$\int_0^{+\infty} \int_\Omega p \partial_t \varphi dx dt = \int_0^{\infty} \int_\Omega [D_p \nabla p + \mu_p p \mathbf{E}] \nabla \varphi dx dt + \int_0^{+\infty} \int_\Omega R(x, n, p) \varphi dx dt, \quad (2.69)$$

for all $\varphi \in \mathcal{C}_0^\infty(]0, +\infty[, Y)$*, where* Y *is the closure of* $\mathcal{C}_0^\infty(\mathbb{R}^3 \setminus \Gamma_D)$.[17]

Uniqueness is also proved in [Joc99] for solutions with the additional regularity

$$\mathbf{E} \in L^\infty_{loc}([0, \infty[, L^p(\Omega)), \quad (2.70)$$

for some $p > 3$ in 3D and $p > 2$ in 2D.

Theorem 2 ([Joc05])**.** *If* $(n, p, \mathbf{E}, \mathbf{H})$ *is a solution to the Maxwell-drift-diffusion system that has the regularity* $n, p \in L^\infty(0, T, L^\infty(\Omega))$ *and* $(\mathbf{E}, \mathbf{H} - \mathbf{H}_N) \in L^\infty(0, T, (L^2(\Omega))^6)$*, then*

$$ess \inf_{(0,T]\times\Omega} n(t, x) > 0$$

and

$$ess \inf_{(0,T]\times\Omega} p(t, x) > 0.$$

Moreover, in [Joc05] it is established that if H_N and $\sigma(t) - \sigma_0$ decay suitably as $t \to +\infty$, then p and n converge respectively to p_e and n_e and $\mathbf{E}$ converges to $\nabla \varphi_e$.

Finally, it should be pointed out that similar fundamental results have been obtained for the asymptotic behavior on the quasi-static PDD system (2.54) (see [GG96], or [Jün95] for more a general description of diffusion).

2.6.3 Energy principle

On the grounds of the above discussion we are finally at the stage of developing our energy analysis. We consider the following formal relative energy of the MDD system with respect to the equilibrium state (n_e, p_e, φ_e) (see also [Joc05] for a similar definition):

$$\mathcal{E}(t) = \nu^{-1} q \int_\Omega \int_{n_e(x)}^{n(t,x)} (\ln(s) - \ln(n_e(x))) \, ds dx + \nu^{-1} q \int_\Omega \int_{p_e(x)}^{p(t,x)} (\ln(s) - \ln(p_e(x))) \, ds dx + \frac{1}{2} \varepsilon \|\mathbf{E} - \nabla \varphi_e\|^2_{L^2(\Omega)} + \frac{1}{2} \mu \|\mathbf{H} - \mathbf{H}_N\|^2_{L^2(\Omega)}. \quad (2.71)$$

[17] The theorem is stated in [Joc99], but for the proof, the author redirects the reader to an earlier paper [Joc96] in which μ_n, μ_p, D_n, D_p are assumed to be constant. Then, for the general case in which these are space-varying, a second paper under review is cited – the published result can be found in [Joc01].

The theorem below proves that (2.71) is bounded in finite time and also provides a first, preliminary a priori estimate on the unknowns.[18]

Theorem 3. *Let us suppose that*

- Ω *is an open bounded domain;*
- D_n, D_p, μ_n, μ_p *are real positive constants;*
- *Hypotheses H1-H7 introduced for Theorem 1 hold, along with* $\mathbf{H}_N \in \mathcal{C}^0(0,T,H(curl,\Omega))$*;*
- *a non-negative* $G \in L^\infty(\Omega)$ *is introduced in the DD equations;*
- *The static PDD problem 2.65 admits a solution* (n_e, p_e, φ_e) *such that*
  $\varphi_e \in H^1(\Omega)$,
  $n_e, p_e \in L^1(\Omega)$,
  $n_e^{-1}, p_e^{-1} \in L^\infty(\Omega)$,
  $n_e^{-1}\mathbf{J}_n^e, p_e^{-1}\mathbf{J}_p^e \in L^\infty(\Omega)$*;*
- *the MDD system (now featuring G) admits a solution* $(\mathbf{E}, \mathbf{H}, n, p)$ *such that*
  $\mathbf{E}, \mathbf{H} \in \mathcal{C}^0(0,T,H(curl,\Omega)) \cap \mathcal{C}^1(0,T,L^2(\Omega))$,
  $n, p \in \mathcal{C}^1(0,T,L^2(\Omega)) \cap \mathcal{C}^0(0,T,H^1(\Omega)) \cap \mathcal{C}^0(0,T,L^\infty(\Omega))$ *and are strictly positive,*
  $n^{-\frac{1}{2}}\mathbf{J}_n,\ p^{-\frac{1}{2}}\mathbf{J}_p \in \mathcal{C}^0(0,T,L^2(\Omega))$*;*
- R *is such that* $R(x,n_1,n_2) \geq 0$*, if* $n_1 n_2 \geq n_e p_e$ *and* $R(x,n_1,n_2) \leq 0$ *if* $n_1 n_2 \leq n_e p_e$*.*

Then, one has for all $t \in [0,T]$*,*

$$\mathcal{E}(t) \leq -\frac{\alpha + \beta C_e}{\beta} + \left(\mathcal{E}(0) + \frac{\alpha + \beta C_e}{\beta}\right)\exp(\beta t), \tag{2.72}$$

with

$$\alpha = \frac{1}{2\varepsilon}\|curl\ \mathbf{H}_N\|^2_{\mathcal{C}^0(0,T,L^2(\Omega))} + 2\nu^{-1}q\|G\|_{L^\infty}|\Omega|,$$

$$\beta = q^{-2}\left(D_n^{-1}\left\|\frac{\mathbf{J}_n^e}{n_e}\right\|_{L^\infty} + D_p^{-1}\left\|\frac{\mathbf{J}_p^e}{p_e}\right\|_{L^\infty} + 1 + q^2\max(\|n_e^{-1}\|_{L^\infty}, \|p_e^{-1}\|_{L^\infty})\right),$$

$$C_e = \nu^{-1}q(e-1)(\|n_e\|_{L^1} + \|p_e\|_{L^1}).$$

As a result, the energy is bounded in finite time. Furthermore, one has that for all $t \in [0,T]$

$$\mathcal{E}(t) \geq \nu^{-1}q\|n\|_{L^1} + \nu^{-1}q\|p\|_{L^1} + \frac{1}{2}\varepsilon\|\mathbf{E} - \nabla\varphi_e\|^2_{L^2(\Omega)} + \frac{1}{2}\mu\|\mathbf{H} - \mathbf{H}_N\|^2_{L^2(\Omega)} - C_e. \tag{2.73}$$

[18] To the purpose of the theorem, and with particular attention to the forthcoming expression of C_e, we remark the difference between Euler's number e and the subscript e, e.g. in expressions like n_e, p_e, which is a short for *equilibrium*.

Proof. One has

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= \nu^{-1}q\langle\partial_t n, \ln(n) - \ln(n_e)\rangle + \nu^{-1}q\langle\partial_t p, \ln(p) - \ln(p_e)\rangle \\ &\quad + \langle\varepsilon\partial_t\mathbf{E}, \mathbf{E} - \nabla\varphi_e\rangle + \langle\mu\partial_t\mathbf{H}, \mathbf{H} - \mathbf{H}_N\rangle.\end{aligned}$$

Using the equations in the MDD model, one finds

$$\begin{aligned}\frac{d\mathcal{E}}{dt} =&\nu^{-1}q\langle\partial_t n, \ln(n) - \ln(n_e)\rangle + \nu^{-1}q\langle\partial_t p, \ln(p) - \ln(p_e)\rangle + \langle\text{curl }\mathbf{H}, \mathbf{E} - \nabla\varphi_e\rangle + \langle-\text{curl }\mathbf{E}, \mathbf{H} - \mathbf{H}_N\rangle \\ &- \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E} - \nabla\varphi_e\rangle \\ =&\nu^{-1}q\langle\partial_t n, \ln(n) - \ln(n_e)\rangle + \nu^{-1}q\langle\partial_t p, \ln(p) - \ln(p_e)\rangle + \langle\mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E} - \nabla\varphi_e\rangle_{\partial\Omega} + \langle\text{curl }\mathbf{E}, \mathbf{H}_N\rangle \\ &- \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E} - \nabla\varphi_e\rangle \\ =&\nu^{-1}\langle\nabla\cdot\mathbf{J}_n + qG - qR, \ln(n) - \ln(n_e)\rangle - \nu^{-1}\langle\nabla\cdot\mathbf{J}_p - qG + qR, \ln(p) - \ln(p_e)\rangle + \langle\mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E} - \nabla\varphi_e\rangle_{\partial\Omega} \\ &+ \langle\text{curl }\mathbf{E}, \mathbf{H}_N\rangle - \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E} - \nabla\varphi_e\rangle \\ =&\nu^{-1}\langle\nabla\cdot\mathbf{J}_n, \ln(n) - \ln(n_e)\rangle - \nu^{-1}\langle\nabla\cdot\mathbf{J}_p, \ln(p) - \ln(p_e)\rangle \\ &+ \nu^{-1}\langle qG - qR, \ln(n) - \ln(n_e)\rangle + \nu^{-1}\langle qG - qR, \ln(p) - \ln(p_e)\rangle + \langle\mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E} - \nabla\varphi_e\rangle_{\partial\Omega} \\ &+ \langle\text{curl }\mathbf{E}, \mathbf{H}_N\rangle - \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E} - \nabla\varphi_e\rangle \\ =&-\nu^{-1}\langle\mathbf{J}_n, \nabla(\ln(n) - \ln(n_e))\rangle + \nu^{-1}\langle\mathbf{J}_p, \nabla(\ln(p) - \ln(p_e))\rangle \\ &+ \nu^{-1}\langle qG - qR, \ln(n) - \ln(n_e)\rangle + \nu^{-1}\langle qG - qR, \ln(p) - \ln(p_e)\rangle \\ &+ \langle\text{curl }\mathbf{E}, \mathbf{H}_N\rangle - \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E} - \nabla\varphi_e\rangle + \langle\mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E} - \nabla\varphi_e\rangle_{\partial\Omega} \\ &+ \nu^{-1}\langle\mathbf{J}_n\cdot\hat{\mathbf{n}}, \ln(n) - \ln(n_e)\rangle_{\partial\Omega} - \nu^{-1}\langle\mathbf{J}_p\cdot\hat{\mathbf{n}}, \ln(p) - \ln(p_e)\rangle_{\partial\Omega} \\ =&\underbrace{-\nu^{-1}\langle\mathbf{J}_n, \nabla\ln(n)\rangle + \nu^{-1}\langle\mathbf{J}_p, \nabla\ln(p)\rangle - \langle\mathbf{J}_n + \mathbf{J}_p, \mathbf{E}\rangle}_{=:I(n,p,\mathbf{E})} \\ &+ \nu^{-1}\langle\mathbf{J}_n, \underbrace{\nabla(\ln(n_e + \nu\varphi_e)}_{=D_n^{-1}q^{-1}\dfrac{\mathbf{J}_p^e}{n_e}}\rangle - \nu^{-1}\langle\mathbf{J}_p, \underbrace{\nabla(\ln(p_e) - \nu\varphi_e)}_{=D_p^{-1}q^{-1}\dfrac{\mathbf{J}_n^e}{p_e}}\rangle \\ &+ \nu^{-1}\langle qG - qR, \ln(n) - \ln(n_e)\rangle + \nu^{-1}\langle qG - qR, \ln(p) - \ln(p_e)\rangle \\ &+ \langle\text{curl }\mathbf{E}, \mathbf{H}_N\rangle + \langle\mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E}\rangle + \nu^{-1}\langle\mathbf{J}_n\cdot\hat{\mathbf{n}}, \ln(n) - \ln(n_e)\rangle_{\partial\Omega} - \nu^{-1}\langle\mathbf{J}_p\cdot\hat{\mathbf{n}}, \ln(p) - \ln(p_e)\rangle_{\partial\Omega}.\end{aligned}$$

But

$$\begin{aligned}\mathbf{J}_n &= qn\mu_n\mathbf{E} + qD_n\nabla n, \\ \mathbf{J}_p &= qp\mu_p\mathbf{E} - qD_p\nabla p.\end{aligned}$$

Therefore, with $n > 0$ and $p > 0$,

$$\begin{aligned}\frac{\mathbf{J}_n}{n} &= q\mu_n \mathbf{E} + qD_n \nabla \ln(n),\\ \frac{\mathbf{J}_p}{p} &= q\mu_p \mathbf{E} - qD_p \nabla \ln(p).\end{aligned}$$

This gives

$$\begin{aligned}\nabla \ln(n) &= q^{-1} D_n^{-1} \frac{\mathbf{J}_n}{n} - \nu \mathbf{E},\\ \nabla \ln(p) &= \nu \mathbf{E} - q^{-1} D_p^{-1} \frac{\mathbf{J}_p}{p}.\end{aligned}$$

Thus

$$\begin{aligned}I(n,p,\mathbf{E}) &= -\nu^{-1}\langle \mathbf{J}_n, q^{-1} D_n^{-1} \frac{\mathbf{J}_n}{n} - \nu \mathbf{E}\rangle\\ &\quad +\nu^{-1}\langle \mathbf{J}_p, \nu\mathbf{E} - q^{-1} D_p^{-1} \frac{\mathbf{J}_p}{p}\rangle - \langle \mathbf{J}_n + \mathbf{J}_p, \mathbf{E}\rangle.\end{aligned}$$

We obtain

$$\begin{aligned}I(n,p,\mathbf{E}) &= -\nu^{-1}\langle \mathbf{J}_n, q^{-1} D_n^{-1} \frac{\mathbf{J}_n}{n} - \nu \mathbf{E}\rangle\\ &\quad +\nu^{-1}\langle \mathbf{J}_p, \nu\mathbf{E} - q^{-1} D_p^{-1} \frac{\mathbf{J}_p}{p}\rangle - \langle \mathbf{J}_n + \mathbf{J}_p, \mathbf{E}\rangle.\end{aligned}$$

As a consequence

$$I(n,p,\mathbf{E}) = \underbrace{-\nu^{-1}q^{-1}\langle \mathbf{J}_n, D_n^{-1}\frac{\mathbf{J}_n}{n}\rangle - \nu^{-1}q^{-1}\langle \mathbf{J}_p, D_p^{-1}\frac{\mathbf{J}_p}{p}\rangle}_{-\nu^{-1}q^{-1}\|(D_n n)^{-\frac{1}{2}}\mathbf{J}_n\|_{L^2}^2 - \nu^{-1}q^{-1}\|(D_p p)^{-\frac{1}{2}}\mathbf{J}_p\|_{L^2}^2} \leq 0.$$

To summarize, we arrive at

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= I(n,p,\mathbf{E}) + \nu^{-1}q\langle G - R, \ln(n) - \ln(n_e)\rangle + \nu^{-1}q\langle G - R, \ln(p) - \ln(p_e)\rangle\\ &\quad +\langle \mathrm{curl}\ \mathbf{E}, \mathbf{H}_N\rangle + \langle \mathbf{H}\times\hat{\mathbf{n}}, \mathbf{E} - \nabla\varphi_e\rangle\\ &\quad +\nu^{-1}\langle \mathbf{J}_n\cdot\hat{\mathbf{n}}, \ln(n) - \ln(n_e)\rangle_{\partial\Omega} - \nu^{-1}\langle \mathbf{J}_p\cdot\hat{\mathbf{n}}, \ln(p) - \ln(p_e)\rangle_{\partial\Omega}\\ &\quad -\nu^{-1}D_n^{-1}q^{-1}\langle \mathbf{J}_n, \frac{\mathbf{J}_n^e}{n_e}\rangle - \nu^{-1}D_p^{-1}q^{-1}\langle \mathbf{J}_p, \frac{\mathbf{J}_p^e}{p_e}\rangle.\end{aligned}$$

Let us now take into account the boundary conditions and reintroduce $\nabla\varphi_e$.

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= I(n,p,\mathbf{E}) + \nu^{-1}q\langle G - R, \ln(n) - \ln(n_e)\rangle + \nu^{-1}q\langle G - R, \ln(p) - \ln(p_e)\rangle\\ &\quad +\langle \mathrm{curl}\ (\mathbf{E} - \nabla\varphi_e), \mathbf{H}_N\rangle + \langle \mathbf{H}_N\times\hat{\mathbf{n}}, \mathbf{E}\rangle_{\Gamma_N} - \langle \mathbf{H}\times\hat{\mathbf{n}}, \nabla\varphi_e\rangle_{\partial\Omega}\\ &\quad -\nu^{-1}D_n^{-1}q^{-1}\langle \mathbf{J}_n, \frac{\mathbf{J}_n^e}{n_e}\rangle - \nu^{-1}D_p^{-1}q^{-1}\langle \mathbf{J}_p, \frac{\mathbf{J}_p^e}{p_e}\rangle.\end{aligned}$$

Having supposed that $n_e = n_D$ is constant on each connected component of Γ_D, $\nabla\varphi_e \times \hat{\mathbf{n}} = 0$ on Γ_D. [19] Then, using the boundary conditions we obtain that

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= I(n,p,\mathbf{E}) + \nu^{-1}q\langle G-R, \ln(np)-\ln(n_e p_e)\rangle \\ &\quad +\langle \mathbf{E}-\nabla\varphi_e, \text{curl } \mathbf{H}_N\rangle \\ &\quad -\nu^{-1}D_n^{-1}q^{-1}\langle \mathbf{J}_n, \frac{\mathbf{J}_n^e}{n_e}\rangle - \nu^{-1}D_p^{-1}q^{-1}\langle \mathbf{J}_p, \frac{\mathbf{J}_p^e}{p_e}\rangle.\end{aligned}$$

We now observe that

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &= \underbrace{I(n,p,\mathbf{E})}_{\leq 0} + \nu^{-1}q\langle G, \ln(np)-\ln(n_e p_e)\rangle - \underbrace{\nu^{-1}q\langle R, \ln(np)-\ln(n_e p_e)\rangle}_{\geq 0} \\ &\quad +\langle \mathbf{E}-\nabla\varphi_e, \text{curl } \mathbf{H}_N\rangle - \nu^{-1}D_n^{-1}q^{-1}\langle \mathbf{J}_n, \frac{\mathbf{J}_n^e}{n_e}\rangle - \nu^{-1}D_p^{-1}q^{-1}\langle \mathbf{J}_p, \frac{\mathbf{J}_p^e}{p_e}\rangle.\end{aligned}$$

Since we suppose that D_n (D_p) and μ_n (μ_p) are constant functions, the following bound holds:

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &\leq -\nu^{-1}q^{-1}(D_n^{-1}\|n^{-\frac{1}{2}}\mathbf{J}_n\|_{L^2}^2 + D_p^{-1}\|p^{-\frac{1}{2}}\mathbf{J}_p\|_{L^2}^2) + \nu^{-1}q\langle G, \ln(np)-\ln(n_e p_e)\rangle \\ &\quad +\|\mathbf{E}-\nabla\varphi_e\|_{L^2}\|\text{curl } \mathbf{H}_N\|_{L^2} + \nu^{-1}D_n^{-1}q^{-1}\|\mathbf{J}_n\|_{L^1}\left\|\frac{\mathbf{J}_n^e}{n_e}\right\|_{L^\infty} + \nu^{-1}D_p^{-1}q^{-1}\|\mathbf{J}_p\|_{L^1}\left\|\frac{\mathbf{J}_p^e}{p_e}\right\|_{L^\infty}.\end{aligned}$$

We now use the Cauchy-Schwarz inequality to write $||\mathbf{J}_n||_{L^1} = ||n^{\frac{1}{2}}n^{-\frac{1}{2}}\mathbf{J}_n||_{L^1} \leq ||n^{\frac{1}{2}}||_{L^2}||n^{-\frac{1}{2}}\mathbf{J}_n||_{L^2}$ and similarly for $\mathbf{J}_p$ obtaining

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &\leq -\nu^{-1}q^{-1}(D_n^{-1}\|n^{-\frac{1}{2}}\mathbf{J}_n\|_{L^2}^2 + D_p^{-1}\|p^{-\frac{1}{2}}\mathbf{J}_p\|_{L^2}^2) + \nu^{-1}q\langle G, \ln(np)-\ln(n_e p_e)\rangle \\ &\quad +\|\mathbf{E}-\nabla\varphi_e\|_{L^2}\|\text{curl } \mathbf{H}_N\|_{L^2} + \nu^{-1}D_n^{-1}q^{-1}\|n^{\frac{1}{2}}\|_{L^2}\|n^{-\frac{1}{2}}\mathbf{J}_n\|_{L^2}\left\|\frac{\mathbf{J}_n^e}{n_e}\right\|_{L^\infty} \\ &\quad +\nu^{-1}D_p^{-1}q^{-1}\|p^{\frac{1}{2}}\|_{L^2}\|p^{-\frac{1}{2}}\mathbf{J}_p\|_{L^2}\left\|\frac{\mathbf{J}_p^e}{p_e}\right\|_{L^\infty}.\end{aligned}$$

Then, applying Young's inequality we get

$$\begin{aligned}\frac{d\mathcal{E}}{dt} &\leq \nu^{-1}q^{-1}D_n^{-1}\left(\frac{1}{2}-1\right)\|n^{-\frac{1}{2}}\mathbf{J}_n\|_{L^2}^2 + \nu^{-1}q^{-1}D_p^{-1}\left(\frac{1}{2}-1\right)\|p^{-\frac{1}{2}}\mathbf{J}_p\|_{L^2}^2 \\ &\quad +\nu^{-1}q\langle G, \ln(np)-\ln(n_e p_e)\rangle + \|\mathbf{E}-\nabla\varphi_e\|_{L^2}\|\text{curl } \mathbf{H}_N\|_{L^2} \\ &\quad +\frac{1}{2}\nu^{-1}q^{-1}D_n^{-1}\|n\|_{L^1}\left\|\frac{\mathbf{J}_n^e}{n_e}\right\|_{L^\infty}^2 \\ &\quad +\frac{1}{2}\nu^{-1}q^{-1}D_p^{-1}\|p\|_{L^1}\left\|\frac{\mathbf{J}_p^e}{p_e}\right\|_{L^\infty}^2.\end{aligned}$$

[19] We recall that n_D was defined in (2.7) as a function of the net doping concentration C, which is constant on each connected component of Γ_D (contacts are fabricated on uniformly doped regions). Then, $\varphi_e = V_{bias} + \varphi_{bi}$ on Γ_D, where $\varphi_{bi} = V_T \log(n_D)$. Hence, φ_e too is constant along each connected component of Γ_D.

Furthermore, one can find lower bound for (2.71) as follows.[20] One has, for a.a. $(t,x) \in [0,T]\times\Omega$

$$\int_{n_e(x)}^{n(t,x)} \ln(s) - \ln(n_e(x))ds \geq n(t,x) - n_e(x)(e-1).$$

We obtain a similar result for p. One has, for a.a. $(t,x) \in [0,T]\times\Omega$

$$\int_{p_e(x)}^{p(t,x)} \ln(s) - \ln(p_e(x))ds \geq p(t,x) - p_e(x)(e-1).$$

To summarize, we obtain that

$$\mathcal{E}(t) \geq \nu^{-1}q\|n\|_{L^1}+\nu^{-1}q\|p\|_{L^1}-\nu^{-1}q(e-1)(\|n_e\|_{L^1}+\|p_e\|_{L^1})+\frac{1}{2}\varepsilon\|\mathbf{E}-\nabla\varphi_e\|^2_{L^2(\Omega)}+\frac{1}{2}\mu\|\mathbf{H}-\mathbf{H}_N\|^2_{L^2(\Omega)}.$$

This means that in particular, if we denote by C_e the quantity $\nu^{-1}q(e-1)(\|n_e\|_{L^1}+\|p_e\|_{L^1})$, we obtain

$$\nu^{-1}q\|n\|_{L^1} + \nu^{-1}q\|p\|_{L^1} + \frac{1}{2}\varepsilon\|\mathbf{E}-\nabla\varphi_e\|^2_{L^2} \leq \mathcal{E}(t) + C_e.$$

Thus

$$\begin{aligned}\frac{d\mathcal{E}}{dt} \leq\ & \underbrace{-\frac{1}{2}\nu^{-1}D_n^{-1}q^{-1}\|n^{-\frac{1}{2}}\mathbf{J}_n\|^2_{L^2} - \frac{1}{2}\nu^{-1}q^{-1}D_p^{-1}\|p^{-\frac{1}{2}}\mathbf{J}_p\|^2_{L^2}}_{=:\tilde{I}\leq 0} + \nu^{-1}q\langle G, \ln(np) - \ln(n_ep_e)\rangle \\ & + q^{-2}\left(D_n^{-1}\left\|\frac{J_n^e}{n_e}\right\|_{L^\infty} + D_p^{-1}\left\|\frac{J_p^e}{p_e}\right\|_{L^\infty} + 1\right)(\mathcal{E}(t)+C_e) + \frac{1}{2\varepsilon}\|\text{curl }\mathbf{H}_N\|^2_{L^2}.\end{aligned}$$

Since $G \geq 0$ does not depend on the unknowns and is in L^∞,

$$\begin{aligned}\nu^{-1}q\langle G, \ln(np) - \ln(n_ep_e)\rangle &= \nu^{-1}q\int_\Omega G(\ln(np) - \ln(n_ep_e)), \\ &\leq \nu^{-1}q\int_\Omega G(|\frac{n-n_e}{n_e} + \frac{p-p_e}{p_e}|), \\ &\leq \nu^{-1}q\|G\|_{L^\infty}(2|\Omega| + \|n\|_{L^1}\|n_e^{-1}\|_{L^\infty} + \|p\|_{L^1}\|p_e^{-1}\|_{L^\infty}), \\ &\leq 2\nu^{-1}q\|G\|_{L^\infty}|\Omega| + \|G\|_{L^\infty}\max(\|n_e^{-1}\|_{L^\infty}, \|p_e^{-1}\|_{L^\infty})(\mathcal{E}+C_e),\end{aligned}$$

hence

[20] It can be proven that, for every $c_0 \in \mathbb{R}$ and $a \in \mathbb{R}_+$, the function

$$g_a\ :\ \xi \in \mathbb{R}_+ \to \int_a^\xi [ln(s) - ln(a)]ds - c_0\xi \in \mathbb{R}, \tag{2.74}$$

verifies

$$g_a(\xi) \geq a(1-e^{c_0}), \qquad \forall \xi \in \mathbb{R}_+. \tag{2.75}$$

In the proof of the theorem, the above inequality is applied to the triples $(\xi, a, c_0) = (n(t,x), n_e(x), 1),\ (p(t,x), p_e(x), 1)$ for $(t,x) \in [0,T]\times\Omega$.

$$\frac{d\mathcal{E}}{dt} \leq \underbrace{\frac{1}{2\varepsilon}\|\mathrm{curl}\ \mathbf{H}_N\|_{L^2}^2 + 2\nu^{-1}q\|G\|_{L^\infty}|\Omega|}_{\alpha}$$
$$+ \underbrace{q^{-2}\left(D_n^{-1}\left\|\frac{J_n^e}{n_e}\right\|_{L^\infty} + D_p^{-1}\left\|\frac{J_p^e}{p_e}\right\|_{L^\infty} + 1 + q^2\|G\|_{L^\infty}\max(\|n_e^{-1}\|_{L^\infty}, \|p_e^{-1}\|_{L^\infty})\right)}_{\beta}(\mathcal{E} + C_e)\,.$$

Then a Grönwall inequality yields for all $t \in [0, T]$,

$$\mathcal{E}(t) \leq -\frac{\alpha + \beta C_e}{\beta} + \left(\mathcal{E}(0) + \frac{\alpha + \beta C_e}{\beta}\right)\exp(\beta t),$$

and we conclude that the energy is bounded in finite time.

□

2.7 Summary of DD models

To conclude the chapter, we propose a brief summary of the models stated so far.

- Interior problems (Figure 2.1)

  – *Maxwell-Drift-Diffusion, without dispersion*

  Equations (2.1), initial/boundary conditions (2.2)

  – *Static Poisson-Drift-Diffusion*

  Equations and initial/boundary conditions (2.65)

- Exterior problems (Figure 2.2)

  – *Maxwell-Drift-Diffusion, without dispersion*

  Equations (2.1 and 2.10), initial/boundary conditions (2.11)

  – *Maxwell-Drift-Diffusion, with dispersion*

  Equations (2.62), initial/boundary conditions (2.11) (excepting $\hat{\mathbf{n}} \times \mathbf{E}|_{\Gamma_D}$) and (2.63)

  – *Static Poisson-Drift-Diffusion*

  Equations (2.51), boundary conditions (2.52)

  – *Quasi-static Poisson-Drift-Diffusion*

  Equations (2.54), boundary conditions (2.52) (with $V_{bias} = V_{bias}(t)$)

Chapter 3

Discrete model

In Section 1.2 we pondered several numerical frameworks and highlighted the DG one as an attractive and credible alternative to conventional approaches (FD, FV, FE).
DG methods were originally conceived for conservation laws, whereas in our applications we are demanded to handle diffusion. Their extension to elliptic problems has drawn special attention, due to the discontinuous nature of the numerical solution. A notable reference on this subject and its historical background can be found in [CKS00]; essentially, two main approaches can be identified, as described below.

1. Letting space derivatives of the unknown in the numerical formulation (as in continuous finite elements) but enforcing regularity by compensatory jumps. To that purpose, *interior penalty* terms are introduced in the discrete equation. A detailed example on Poisson's equation in one dimension can be found in [Riv08].

2. Introducing an auxiliary variable for the gradient of the unknown and therefore an auxiliary equation. The approach has been extensively explored, with a special focus on Laplace's equation – [ABCM02] is a prominent reference on the several schemes offered by the literature (some of which relying on interior penalties as well) and their mathematical properties.

In the present work, the second approach is followed and this section develops its application to the MDD system.
As it will appear clear over the course of this chapter, all the challenges can be understood and overcome by working in a bounded domain. Indeed: a) the coupling takes place inside the semiconductor; b) the major change when switching between an interior and an exterior problem is in initial and boundary conditions; c) such conditions play a minor role in the above-mentioned task; d) outside the semiconductor the model reduces to source-free Maxwell's equations.
The forthcoming discussion is organized as follows. In Section 3.1, the local DG formulation is formally stated in its integral form, and the key task of resolving the discontinuity of the numerical solution at interfaces between elements is addressed. To this purpose, the advective and diffusive part of the MDD model are studied separately; in particular, the advective part allows an interesting parallel with FV methods. The semi-discrete formulation is then stated in a matrix form. In Section 3.1.3, a global integral formulation is derived along the lines of a notable procedure described in [ABCM02], with the idea of investigating the possibility of a future theoretical stability study. The chapter is concluded outlining explicit time integration of the MDD model and identifying a suitable stability condition for our advection-diffusion context.

3.1 Semi-discrete DG formulation: local form

The starting point is a local weak formulation on a subset K of Ω for suitable test functions $\boldsymbol{\tau}, v$. We let $\hat{\mathbf{n}}$ denote the outward unit normal vector to ∂K.[1]

$$
\begin{aligned}
\int_K \mu \frac{\partial \mathbf{H}}{\partial t} \cdot \boldsymbol{\tau} &= -\int_K \mathbf{E} \cdot \nabla \times \boldsymbol{\tau} + \oint_{\partial K} \mathbf{E} \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau}, \\
\int_K \epsilon \frac{\partial \mathbf{E}}{\partial t} \cdot \boldsymbol{\tau} &= \int_K \mathbf{H} \cdot \nabla \times \boldsymbol{\tau} - \oint_{\partial K} \mathbf{H} \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau} \\
&- \int_K \big(q\mu_n n + q\mu_p p\big)\mathbf{E} \cdot \boldsymbol{\tau} - \int_K q D_n \boldsymbol{\alpha} \cdot \boldsymbol{\tau} + \int_K q D_p \boldsymbol{\beta} \cdot \boldsymbol{\tau}, \\
\int_K \frac{\partial n}{\partial t} v &= \oint_{\partial K} v n \mu_n \, \mathbf{E} \cdot \hat{\mathbf{n}} - \int_K n \, \mu_n \nabla v \cdot \mathbf{E} \\
&+ \oint_{\partial K} v D_n \, \boldsymbol{\alpha} \cdot \hat{\mathbf{n}} - \int_K D_n \nabla v \cdot \boldsymbol{\alpha} + \int_K (G - R) v, \\
\int_K \frac{\partial p}{\partial t} v &= -\oint_{\partial K} v p \mu_p \, \mathbf{E} \cdot \hat{\mathbf{n}} + \int_K p \, \mu_p \nabla v \cdot \mathbf{E} \\
&+ \oint_{\partial K} v D_p \, \boldsymbol{\beta} \cdot \hat{\mathbf{n}} - \int_K D_p \nabla v \cdot \boldsymbol{\beta} + \int_K (G - R) v, \\
\int_K \boldsymbol{\alpha} \cdot \boldsymbol{\tau} &= \oint_{\partial K} n \, \boldsymbol{\tau} \cdot \hat{\mathbf{n}} - \int_K n \nabla \cdot \boldsymbol{\tau}, \\
\int_K \boldsymbol{\beta} \cdot \boldsymbol{\tau} &= \oint_{\partial K} p \, \boldsymbol{\tau} \cdot \hat{\mathbf{n}} - \int_K p \nabla \cdot \boldsymbol{\tau}.
\end{aligned}
\tag{3.1}
$$

Next, let $h > 0$ and $\mathcal{T}_h$ be a discretization of Ω of maximum size h. We set

$$
\begin{aligned}
V_h &:= \{v \in L^2(\Omega) : v|_K \in \mathcal{P}_{\mathfrak{p}}(K), \quad \forall K \in \mathcal{T}_h\}, \\
\Sigma_h &:= \{\boldsymbol{\tau} \in [L^2(\Omega)]^3 : \boldsymbol{\tau}|_K \in [\mathcal{P}_{\mathfrak{p}}(K)]^3, \quad \forall K \in \mathcal{T}_h\},
\end{aligned}
\tag{3.2}
$$

where $\mathfrak{p} \in \mathbb{N}_0$ denotes the desired order for the local polynomial approximation in K and $\mathcal{P}_{\mathfrak{p}}(K)$ is the space of polynomial functions defined on K of order less or equal to $\mathfrak{p}$.[2][3]

The semi-discrete numerical formulation of (2.1) consists in finding $\mathbf{H}_h, \mathbf{E}_h, \boldsymbol{\alpha}_h, \boldsymbol{\beta}_h \in \Sigma_h$ and $n_h, p_h \in V_h$ such that for all $\boldsymbol{\tau} \in \Sigma_h$, $v \in V_h$ and $K \in \mathcal{T}_h$,

[1] It is convenient for our purposes to let all vector (scalar) test functions be denoted by $\boldsymbol{\tau}$ (v), but this is not a general rule – there may be contexts in which different vector (scalar) equations need to be projected on different test functions.

[2] A more rigorous notation for the order would be $\mathfrak{p}_K$, to stress one of the key advantages of the discontinuous Galerkin methods with respect to the continuous ones – the ability to be $\mathfrak{p}$-adaptive, i.e. to allow different values of $\mathfrak{p}$ in different elements. This aspect is implied here to keep the notation light.

[3] In the present work, space discretization is based on nodal Lagrange polynomials.

$$
\begin{aligned}
\int_K \mu \frac{\partial \mathbf{H}_h}{\partial t} \cdot \boldsymbol{\tau} &= -\int_K \mathbf{E}_h \cdot \nabla \times \boldsymbol{\tau} + \oint_{\partial K} \mathbf{E}_* \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau}, \\
\int_K \epsilon \frac{\partial \mathbf{E}_h}{\partial t} \cdot \boldsymbol{\tau} &= \int_K \mathbf{H}_h \cdot \nabla \times \boldsymbol{\tau} - \oint_{\partial K} \mathbf{H}_* \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau} - \int_K \big(q\mu_n n_h + q\mu_p p_h\big)\mathbf{E}_h \cdot \boldsymbol{\tau} \\
&\quad - \int_K q D_n \boldsymbol{\alpha}_h \cdot \boldsymbol{\tau} + \int_K q D_p \boldsymbol{\beta}_h \cdot \boldsymbol{\tau}, \\
\\
\int_K \frac{\partial n_h}{\partial t} v &= \oint_{\partial K} v\big(n\mu_n\ \mathbf{E}\big)_* \cdot \hat{\mathbf{n}} - \int_K n_h\ \mu_n \nabla v \cdot \mathbf{E}_h \\
&\quad + \oint_{\partial K} v\big(D_n\ \boldsymbol{\alpha}\big)_* \cdot \hat{\mathbf{n}} - \int_K D_n \nabla v \cdot \boldsymbol{\alpha} + \int_K (G-R)v, \\
\int_K \frac{\partial p_h}{\partial t} v &= -\oint_{\partial K} v\big(p\mu_p\ \mathbf{E}\big)_* \cdot \hat{\mathbf{n}} + \int_K p_h\ \mu_p \nabla v \cdot \mathbf{E}_h \\
&\quad + \oint_{\partial K} v\big(D_p\ \boldsymbol{\beta}\big)_* \cdot \hat{\mathbf{n}} - \int_K D_p \nabla v \cdot \boldsymbol{\beta} + \int_K (G-R)v, \\
\\
\int_K \boldsymbol{\alpha}_h \cdot \boldsymbol{\tau} &= \oint_{\partial K} n_*\ \boldsymbol{\tau} \cdot \hat{\mathbf{n}} - \int_K n_h \nabla \cdot \boldsymbol{\tau}, \\
\int_K \boldsymbol{\beta}_h \cdot \boldsymbol{\tau} &= \oint_{\partial K} p_*\ \boldsymbol{\tau} \cdot \hat{\mathbf{n}} - \int_K p_h \nabla \cdot \boldsymbol{\tau},
\end{aligned}
\tag{3.3}
$$

where the subscript $*$ attached to a quantity denotes its numerical approximation on ∂K, commonly named *numerical flux*.[4]

A formal definition of numerical flux on the skeleton of the mesh requires a proper functional setting:

$$
\begin{aligned}
H^l(\mathcal{T}_h) &:= \{w \in L^2(\Omega) : w|_K \in H^l(K), \quad \forall K \in \mathcal{T}_h\},\ l \in \mathbb{N}, \\
\Gamma &:= \bigcup_{K \in \mathcal{T}_h} \partial K, \\
\Gamma^0 &:= \Gamma \setminus \partial\Omega, \\
T(\Gamma) &:= \Pi_{K \in \mathcal{T}_h} L^2(\partial K).
\end{aligned}
\tag{3.4}
$$

We remark that:

- $V_h \subset H^l(\mathcal{T}_h)$, $\Sigma_h \subset [H^l(\mathcal{T}_h)]^3$ for any $l \in \mathbb{N}$;
- functions in $H^l(\mathcal{T}_h)$ have their traces in $T(\Gamma) := \Pi_{K \in \mathcal{T}_h} L^2(\partial K)$;
- for each function in $H^l(\mathcal{T}_h)$, two traces can be defined in $T(\Gamma)$, one per side of each face of the mesh; the two differ on Γ_0 and coincide on $\partial\Omega$;
- $L^2(\Gamma) = \{q \in T(\Gamma) : q|_e^+ = q|_e^-, \quad \forall e \ \text{ internal face}\}$.

[4]The expression is borrowed from finite volumes, in which an advection equation is integrated over a control volume, yielding a flux integral of the current. Due to the discontinuous (cell-wise constant) character of the numerical solution, the integral loses its meaning when switching to discrete quantities; thus, the physical flux is replaced by an appropriate *numerical* flux – a function of the two neighboring states that ensures that physical flow across cells is correctly predicted. This aspect will be illustrated in Section 3.1.1.2.

The numerical flux is then defined as:

- a linear mapping $H^1(\mathcal{T}_h) \to T(\Gamma)$ for each scalar unknown (e.g. n_h);
- a linear mapping $[H^1(\mathcal{T}_h)]^3 \to [T(\Gamma)]^3$ for each vector unknown (e.g. $\mathbf{E}_h$).

An important role in the numerical formulation is played by average and jump operators on the generic interior face shared between two mesh elements K_i and K_j. These are defined as

$$\begin{aligned}
\{\cdot\} : q \in T(\Gamma) \quad &\to \quad \frac{1}{2}(q_i + q_j) \quad &\in L^2(\Gamma), \\
\{\cdot\} : \boldsymbol{\varphi} \in [T(\Gamma)]^2 \quad &\to \quad \frac{1}{2}(\boldsymbol{\varphi}_i + \boldsymbol{\varphi}_j) \quad &\in [L^2(\Gamma)]^2, \\
[\![\cdot]\!] : q \in T(\Gamma) \quad &\to \quad q_i \hat{\mathbf{n}}_i + q_j \hat{\mathbf{n}}_j \quad &\in [L^2(\Gamma)]^2, \\
[\![\cdot]\!] : \boldsymbol{\varphi} \in T(\Gamma) \quad &\to \quad \boldsymbol{\varphi}_i \cdot \hat{\mathbf{n}}_i + \boldsymbol{\varphi}_j \cdot \hat{\mathbf{n}}_j \quad &\in L^2(\Gamma),
\end{aligned} \tag{3.5}$$

where $q_i := q|_{\partial K_i}$, $\boldsymbol{\varphi}_i := \boldsymbol{\varphi}|_{\partial K_i}$ and $\hat{\mathbf{n}}_i$ is the outward unit normal vector to ∂K_i. For a boundary face, (3.5) are re-used, except that $q_1 = q$, $\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}$, $q_2 = 0$, $\boldsymbol{\varphi}_2 = \mathbf{0}$.
The introduction of such operators is motivated by the following observations.
In the local formulation on K, the numerical flux u_* associated to an unknown u is the trace of the numerical approximation u_h on the interior side of ∂K.
To recover the global properties of the analytical solution while independently solving local problems, transmission conditions must be imposed on the numerical approximation across different elements of $\mathcal{T}_h$. To this purpose, u_* is usually chosen to be, on each internal face e, a function of both traces defined on e.
The interpretation is consistent with the definition of (global) numerical flux given earlier. Indeed, the global numerical flux is assembled by collecting all the local numerical fluxes; two traces per internal face are collected, yielding a function belonging in $T(\Gamma)$.

To proceed towards discretization, it is convenient to update the notation:

- $\mathcal{T} := \{K_i : i = 1, 2, ..., N_{\mathcal{T}}\}$;
- $u_i := u_h|_{K_i}$ for the generic unknown u;
  - $\diamond$ $\mathfrak{p}_i :=$ degree of the polynomial approximation sought for u_i;
  - $\diamond$ $d_i := \dim \mathcal{P}_{\mathfrak{p}_i}(K_i)$;
- $\Phi_i := \{\phi_{ik} \in \mathcal{P}_{p_i}(K_i) : k \in \{1, 2, ..., d_i\}\}$;
  - $\diamond$ $\boldsymbol{\phi}^{\boldsymbol{x}}_{\boldsymbol{ik}} := [\phi_{ik}\ 0\ 0]^T$, $\boldsymbol{\phi}^{\boldsymbol{y}}_{\boldsymbol{ik}} := [0\ \phi_{ik}\ 0]^T$, $\boldsymbol{\phi}^{\boldsymbol{z}}_{\boldsymbol{ik}} := [0\ 0\ \phi_{ik}]^T$;
- $a_{il} := K_i \cap K_l \qquad i, l = 1, 2, ..., N_{\mathcal{T}}$;
  - $\diamond$ $\hat{\mathbf{n}}_{il} :=$ unit normal vector to a_{il} pointing from K_i to K_l;
  - $\diamond$ $\mathcal{N}_i := \{l : K_i \cap K_l \neq \emptyset\}$ such that $\partial K_i = \bigcup_{l \in \mathcal{N}_i} a_{il}$.

The set Φ_i, whose elements are restrictions to K_i of functions in V_h, is used to define the vector test functions: $\boldsymbol{\tau}$ is chosen such that $\boldsymbol{\tau}|_{K_i} \in \{\boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ik}} \ : \ v = x, y, z; \quad k = 1, 2, ..., d_i\}$.
To simplify matters, the expression $\partial K_i = \bigcup_{l \in \mathcal{N}_i} a_{il}$ is also used for boundary elements – it suffices to assume that if $K_i \cap \partial\Omega \neq \emptyset$ then $K_i \in \mathcal{N}_i$, i.e. K_i is its own neighbor. So $a_{ii} = K_i \cap \partial\Omega$ and $\hat{\mathbf{n}}_{ii}$

is just the outward unit normal vector to a_{ii}.
Boundary conditions on u are imposed by replacing u_* on each boundary face with the required value. If no condition is required for u on a particular face then $u_* = u_i$. On the basis of the new notation the numerical formulation on the i-th element, $i \in \{1, ..., N_{\mathcal{T}}\}$, with the k-th local test function, $k \in \{1, ..., d_i\}$, and $v \in \{x, y, z\}$, reads:

$$
\begin{aligned}
\int_{K_i} \mu \frac{\partial \mathbf{H}_i}{\partial t} \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} &= -\int_{K_i} \mathbf{E}_i \cdot \nabla \times \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} + \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \mathbf{E}_* \times \hat{\mathbf{n}}_{il} \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}}, \\
\int_{K_i} \epsilon \frac{\partial \mathbf{E}_i}{\partial t} \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} &= \int_{K_i} \mathbf{H}_i \cdot \nabla \times \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} - \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \mathbf{H}_* \times \hat{\mathbf{n}}_{il} \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} \\
&- \int_{K_i} \big(q\mu_n n_i + q\mu_p p_i\big) \mathbf{E}_i \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} - \int_{K_i} q D_n \boldsymbol{\alpha}_i \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} + \int_{K_i} q D_p \boldsymbol{\beta}_i \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}}, \\
\int_{K_i} \frac{\partial n_i}{\partial t} \phi_{ik} &= \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik} \big(n\mu_n \ \mathbf{E}\big)_* \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} n_i \ \mu_n \nabla \phi_{ik} \cdot \mathbf{E}_i \\
&+ \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik} \big(D_n \ \boldsymbol{\alpha}\big)_* \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} D_n \nabla \phi_{ik} \cdot \boldsymbol{\alpha} + \int_{K_i} (G - R) \phi_{ik}, \\
\int_{K_i} \frac{\partial p_i}{\partial t} \phi_{ik} &= -\sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik} \big(p\mu_p \ \mathbf{E}\big)_* \cdot \hat{\mathbf{n}}_{il} + \int_{K_i} p_i \ \mu_p \nabla \phi_{ik} \cdot \mathbf{E}_i \\
&+ \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik} \big(D_p \ \boldsymbol{\beta}\big)_* \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} D_p \nabla \phi_{ik} \cdot \boldsymbol{\beta} + \int_{K_i} (G - R) \phi_{ik}, \\
\int_{K_i} \boldsymbol{\alpha}_i \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} &= \sum_{l \in \mathcal{N}_i} \int_{a_{il}} n_* \ \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} n_i \nabla \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}}, \\
\int_{K_i} \boldsymbol{\beta}_i \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} &= \sum_{l \in \mathcal{N}_i} \int_{a_{il}} p_* \ \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}} \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} p_i \nabla \cdot \boldsymbol{\phi}_{\boldsymbol{ik}}^{\boldsymbol{v}}.
\end{aligned}
\tag{3.6}
$$

3.1.1 Defining the numerical flux

We are finally at the point of defining the numerical flux on internal faces. Defining $\mathbf{J}_{n_*,drift} := q(n\mu_n\mathbf{E})_*$ and $\mathbf{J}_{n_*,diff} := q(D_n\boldsymbol{\alpha})_*$ (and similarly for holes), we identify two strategies, explored in Sections 3.1.1.2 and 3.1.1.3 respectively.

- Calculating $\mathbf{E}_*, \mathbf{H}_*, \mathbf{J}_{n_*,drift}, \mathbf{J}_{p_*,drift}$ in the no-diffusion limit (to this purpose, a preliminary analysis of the Maxwell-Drift system is proposed in Section 3.1.1.1); follow references on DG methods for diffusion to choose $(n_*, p_*, \mathbf{J}_{n_*,diff}, \mathbf{J}_{p_*,diff})$.[5]
- Separately selecting $(\mathbf{E}_*, \mathbf{H}_*)$, $(\mathbf{J}_{n_*,drift}, \mathbf{J}_{p_*,drift})$, $(n_*, p_*, \mathbf{J}_{n_*,diff}, \mathbf{J}_{p_*,diff})$ from references on DG methods for Maxwell's equations, electrostatics in semiconductors, conservation laws, diffusion.

[5]Note that n_* and p_* refer to the numerical flux in the auxiliary equations, whereas in the drift currents such quantities are not handled individually – expressions for the ensembles $n\mu_n\mathbf{E}$ and $p\mu_p\mathbf{E}$ are defined on the skeleton of the mesh.

3.1.1.1 The Maxwell-Drift system as a conservation law

A key aspect in the conception of DG schemes for Maxwell's equations is the possibility of recasting these in the form of a conservation law, which allows to handle the discontinuity of the numerical solution on mesh internal faces - inherent to *discontinuous* finite elements - with procedures peculiar to finite volumes. The point will be discussed in more detail in Section 3.1.1.2. Here, the goal is to seek a form of MDD system paving way for the application of such procedures (bearing in mind that these do not support the presence of diffusion, as highlighted in the calculations that follow).

To begin, vector equations and differential operators are broken down into their scalar components

$$\begin{aligned}
-\mu\partial_t H_x &= \partial_y E_z - \partial_z E_y,\\
-\mu\partial_t H_y &= \partial_z E_x - \partial_x E_z,\\
-\mu\partial_t H_z &= \partial_x E_y - \partial_y E_x,\\
\epsilon\partial_t E_x &= \partial_y H_z - \partial_z H_y - qn\mu_n E_x - qD_n\partial_x n - qp\mu_p E_x + qD_p\partial_x p,\\
\epsilon\partial_t E_y &= \partial_z H_x - \partial_x H_z - qn\mu_n E_y - qD_n\partial_y n - qp\mu_p E_y + qD_p\partial_y p,\\
\epsilon\partial_t E_z &= \partial_x H_y - \partial_y H_x - qn\mu_n E_z - qD_n\partial_z n - qp\mu_p E_z + qD_p\partial_z p,\\
\partial_t n &= \partial_x\Big(n\mu_n E_x + D_n\partial_x n\Big) + \partial_y\Big(n\mu_n E_y + D_n\partial_y n\Big) + \partial_z\Big(n\mu_n E_z + D_n\partial_z n\Big)\\
&\quad + G - R,\\
\partial_t p &= -\partial_x\Big(p\mu_p E_x - D_p\partial_x p\Big) - \partial_y\Big(p\mu_p E_y - D_p\partial_y p\Big) - \partial_z\Big(p\mu_p E_z - D_p\partial_z p\Big)\\
&\quad + G - R.
\end{aligned} \tag{3.7}$$

Equivalently,

$$\begin{aligned}
Q\begin{bmatrix} H_x\\ H_y\\ H_z\\ E_x\\ E_y\\ E_z\\ n\\ p\end{bmatrix}_t &= \begin{bmatrix} 0\\ -E_z\\ E_y\\ 0\\ -H_z\\ H_y\\ \mu_n nE_x + D_n\partial_x n\\ -\mu_p pE_x + D_p\partial_x p\end{bmatrix}_x + \begin{bmatrix} E_z\\ 0\\ -E_x\\ H_z\\ 0\\ -H_x\\ \mu_n nE_y + D_n\partial_y n\\ -\mu_p pE_y + D_p\partial_y p\end{bmatrix}_y + \begin{bmatrix} -E_y\\ E_x\\ 0\\ -H_y\\ H_x\\ 0\\ \mu_n nE_z + D_n\partial_z n\\ -\mu_p pE_z + D_p\partial_z p\end{bmatrix}_z\\
&\quad + q\begin{bmatrix} 0\\ 0\\ 0\\ -n\mu_n E_x - D_n\frac{\partial n}{\partial x} - p\mu_p E_x + D_p\frac{\partial p}{\partial x}\\ -n\mu_n E_y - D_n\frac{\partial n}{\partial y} - p\mu_p E_y + D_p\frac{\partial p}{\partial y}\\ -n\mu_n E_z - D_n\frac{\partial n}{\partial z} - p\mu_p E_z + D_p\frac{\partial p}{\partial z}\\ 0\\ 0\end{bmatrix} + \begin{bmatrix} 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ G-R\\ G-R\end{bmatrix},
\end{aligned} \tag{3.8}$$

where

$$Q = diag(-\mu, -\mu, -\mu, \epsilon, \epsilon, \epsilon, 1, 1),$$

and the symbol $\left[\,\cdot\,\right]_c$, $c = t, x, y, z$, denotes a partial derivative.

If we define the vectors

$$\mathbf{u} = \begin{bmatrix} H_x \\ H_y \\ H_z \\ E_x \\ E_y \\ E_z \\ n \\ p \end{bmatrix} \qquad \mathbf{f}^{(1)}(\mathbf{u}) = \begin{bmatrix} 0 \\ -E_z \\ E_y \\ 0 \\ -H_z \\ H_y \\ \mu_n n E_x + D_n \partial_x n \\ -\mu_p p E_x + D_p \partial_x p \end{bmatrix}$$

$$\mathbf{f}^{(2)}(\mathbf{u}) = \begin{bmatrix} E_z \\ 0 \\ -E_x \\ H_z \\ 0 \\ -H_x \\ \mu_n n E_y + D_n \partial_y n \\ -\mu_p p E_y + D_p \partial_y p \end{bmatrix} \qquad \mathbf{f}^{(3)}(\mathbf{u}) = \begin{bmatrix} -E_y \\ E_x \\ 0 \\ -H_y \\ H_x \\ 0 \\ \mu_n n E_z + D_n \partial_z n \\ -\mu_p p E_z + D_p \partial_z p \end{bmatrix} \tag{3.9}$$

$$\mathbf{g}(\mathbf{u}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -qn\mu_n E_x - qD_n \frac{\partial n}{\partial x} - qp\mu_p E_x + qD_p \frac{\partial p}{\partial x} \\ -qn\mu_n E_y - qD_n \frac{\partial n}{\partial y} - qp\mu_p E_y + qD_p \frac{\partial p}{\partial y} \\ -qn\mu_n E_z - qD_n \frac{\partial n}{\partial z} - qp\mu_p E_z + qD_p \frac{\partial p}{\partial z} \\ G - R \\ G - R \end{bmatrix}$$

the system can be rewritten as

$$\begin{aligned} Q\ \partial_t \mathbf{u} &= \partial_x \mathbf{f}^{(1)}(\mathbf{u}) + \partial_y \mathbf{f}^{(2)}(\mathbf{u}) + \partial_z \mathbf{f}^{(3)}(\mathbf{u}) + \mathbf{g}(\mathbf{u}) \\ &= \partial_{\mathbf{u}} \mathbf{f}^{(1)}(\mathbf{u}) \partial_x \mathbf{u} + \partial_{\mathbf{u}} \mathbf{f}^{(2)}(\mathbf{u}) \partial_y \mathbf{u} + \partial_{\mathbf{u}} \mathbf{f}^{(3)}(\mathbf{u}) \partial_z \mathbf{u} + \mathbf{g}(\mathbf{u}). \end{aligned} \tag{3.10}$$

Two limit cases are considered, with the idea of recovering and superposing the results when going back to the original picture.

1. No drift ($\mu_n = \mu_p = 0$), to be addressed by referring to DG schemes for diffusion, discussed in Section 3.1.1.3.

2. No diffusion ($D_n = D_p = 0$), for which $\mathbf{f}^{(1-3)}$ and $\mathbf{g}$ become purely advective and advective-reactive respectively. As a result,

$$Q\begin{bmatrix} H_x \\ H_y \\ H_z \\ E_x \\ E_y \\ E_z \\ n \\ p \end{bmatrix}_t = \begin{bmatrix} 0 \\ -E_z \\ E_y \\ 0 \\ -H_z \\ H_y \\ \mu_n n E_x \\ -\mu_p p E_x \end{bmatrix}_x + \begin{bmatrix} E_z \\ 0 \\ -E_x \\ H_z \\ 0 \\ -H_x \\ \mu_n n E_y \\ -\mu_p p E_y \end{bmatrix}_y + \begin{bmatrix} -E_y \\ E_x \\ 0 \\ -H_y \\ H_x \\ 0 \\ \mu_n n E_z \\ -\mu_p p E_z \end{bmatrix}_z + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -qn\mu_n E_x - qp\mu_p E_x \\ -qn\mu_n E_y - qp\mu_p E_y \\ -qn\mu_n E_z - qp\mu_p E_z \\ G - R \\ G - R \end{bmatrix}. \tag{3.11}$$

This diffusion-less version of (3.10) expresses the conservation of $\mathbf{u}$, and it is said to constitute a *hyperbolic* conservation law if the matrix

$$\mathcal{A} = -Q^{-1}\left(b_x \partial_{\mathbf{u}} \mathbf{f}^{(1)} + b_y \partial_{\mathbf{u}} \mathbf{f}^{(2)} + b_z \partial_{\mathbf{u}} \mathbf{f}^{(3)}\right) \tag{3.12}$$

is diagonalizable and has a real spectrum for all $\mathbf{u}$ and $\mathbf{b} = (b_x, b_y, b_z) \in \mathbb{R}^3$ such that $||\mathbf{b}|| = 1$ [GR91]. Given the Jacobian matrices

$$Q^{-1}\partial_{\mathbf{u}}\mathbf{f}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mu^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\epsilon^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \epsilon^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial \mu_n}{\partial E_x} n E_x + \mu_n n & \frac{\partial \mu_n}{\partial E_y} n E_x & \frac{\partial \mu_n}{\partial E_z} n E_x & \mu_n E_x & 0 \\ 0 & 0 & 0 & -\frac{\partial \mu_p}{\partial E_x} p E_x - \mu_p p & -\frac{\partial \mu_p}{\partial E_y} p E_x & -\frac{\partial \mu_p}{\partial E_z} p E_x & 0 & -\mu_p E_x \end{bmatrix}, \tag{3.13}$$

$$Q^{-1}\partial_{\mathbf{u}}\mathbf{f}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -\mu^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\epsilon^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial \mu_n}{\partial E_x} n E_y & \frac{\partial \mu_n}{\partial E_y} n E_y + \mu_n n & \frac{\partial \mu_n}{\partial E_z} n E_y & \mu_n E_y & 0 \\ 0 & 0 & 0 & -\frac{\partial \mu_p}{\partial E_x} p E_y & -\frac{\partial \mu_p}{\partial E_y} p E_y - \mu_p p & -\frac{\partial \mu_p}{\partial E_z} p E_y & 0 & -\mu_p E_y \end{bmatrix}, \tag{3.14}$$

$$Q^{-1}\partial_{\mathbf{u}}\mathbf{f}^{(3)} = \begin{bmatrix} 0 & 0 & 0 & 0 & \mu^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\epsilon^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \epsilon^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial \mu_n}{\partial E_x} n E_z & \frac{\partial \mu_n}{\partial E_y} n E_z & \frac{\partial \mu_n}{\partial E_z} n E_z + \mu_n n & \mu_n E_z & 0 \\ 0 & 0 & 0 & -\frac{\partial \mu_p}{\partial E_x} p E_z & -\frac{\partial \mu_p}{\partial E_y} p E_z & -\frac{\partial \mu_p}{\partial E_z} p E_z - \mu_p p & 0 & -\mu_p E_z \end{bmatrix}, \tag{3.15}$$

the linear combination (3.12) is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -\mu^{-1}b_z & \mu^{-1}b_y & 0 & 0 \\ 0 & 0 & 0 & \mu^{-1}b_z & 0 & -\mu^{-1}b_x & 0 & 0 \\ 0 & 0 & 0 & -\mu^{-1}b_y & \mu^{-1}b_x & 0 & 0 & 0 \\ 0 & \epsilon^{-1}b_z & -\epsilon^{-1}b_y & 0 & 0 & 0 & 0 & 0 \\ -\epsilon^{-1}b_z & 0 & \epsilon^{-1}b_x & 0 & 0 & 0 & 0 & 0 \\ \epsilon^{-1}b_y & -\epsilon^{-1}b_x & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{n,x}}{\partial E_x} & -\dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{n,y}}{\partial E_y} & -\dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{n,z}}{\partial E_z} & \mathbf{b}\cdot(-\mu_n\mathbf{E}) & 0 \\ 0 & 0 & 0 & \dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{p,x}}{\partial E_x} & \dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{p,y}}{\partial E_y} & \dfrac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_{p,z}}{\partial E_z} & 0 & \mathbf{b}\cdot(\mu_p\mathbf{E}) \end{bmatrix},$$

where we have introduced the compact notation

$$\begin{aligned} \tilde{\mathbf{J}}_n &:= \frac{\mathbf{J}_n}{q} = \frac{\mathbf{J}_{n,drift}}{q} = n\mu_n\mathbf{E}, \\ \frac{\partial\ \mathbf{b}\cdot\tilde{\mathbf{J}}_n}{\partial E_v} &= n\ \frac{\partial\mu_n}{\partial E_v}\ \mathbf{b}\cdot\mathbf{E} + n\mu_n b_v \qquad v = x, y, z, \end{aligned} \tag{3.16}$$

and similarly for $\tilde{\mathbf{J}}_p$.

The eigenvalues are functions of the coefficients, but for $||\mathbf{b}|| = 1$ they reduce to

$$\{0, 0, -\frac{1}{\sqrt{\epsilon\mu}}, -\frac{1}{\sqrt{\epsilon\mu}}, \frac{1}{\sqrt{\epsilon\mu}}, \frac{1}{\sqrt{\epsilon\mu}}, \mathbf{b}\cdot(-\mu_n\mathbf{E}), \mathbf{b}\cdot(\mu_p\mathbf{E})\}, \tag{3.17}$$

in which we recognize the speed of light in the semiconductor and electron and hole drift velocities $\boldsymbol{v}_n = -\mu_n\mathbf{E}$, $\boldsymbol{v}_p = \mu_p\mathbf{E}$. The eigenspaces are not reported due to their complex expressions, but Maple symbolic computations [Map21] would show that geometric and algebraic multiplicities of (3.17) match, and therefore that $(3.10)|_{D_n=D_p=0}$ is hyperbolic.

3.1.1.2 Calculating the advective flux

As discussed in Section 3.1.1.1, the MDD system in the no-diffusion limit is a hyperbolic conservation law of the form

$$\frac{\partial\mathbf{u}}{\partial t} + \nabla\cdot\mathbf{f}(\mathbf{u}) = \mathbf{0}. \tag{3.18}$$

In the framework of finite volume methods, the resolution of such an equation consists in seeking an element-wise constant function $\mathbf{u}_h$ whose level on the i-th element K_i is defined as

$$\mathbf{u}_i = \frac{1}{V_i}\int_{K_i}\mathbf{u},$$

where V_i denotes the volume of K_i. As a result, $\mathbf{u}_h$ is two-valued at each interface $a_{il} = K_i \cap K_l$ and therefore in need of being appropriately defined. This is a crucial task – it majorly affects the properties of the numerical scheme – and is usually based on the resolution of a special initial value

problem, called a *Riemann problem* (RP). A notable reference on the formal procedure is [Tor09], while here the goal is to review a few key points that are essential to our purposes.
First, by means of a suitable rotation of the coordinate system, the original equation is stated in terms of one spatial dimension in the direction of $\hat{\mathbf{n}}_{\mathbf{il}}$. If we call ξ the new coordinate, it can be shown that the resulting equation has the form

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathfrak{f}(\mathbf{u})}{\partial \xi} = \mathbf{0}, \tag{3.19}$$

for a suitable $\mathfrak{f}$ and with $\mathbf{u} = \mathbf{u}(\xi, t)$. Then, remembering that $\hat{\mathbf{n}}_{il}$ points from T_i to T_l in our conventions and assuming that a_{il} is crossed at $\xi = \xi_{il}$, $\mathbf{u}_h(\xi_{il})$ is defined as the solution to the RP

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} + \dfrac{\partial \mathfrak{f}(\mathbf{u})}{\partial \xi} = \mathbf{0}, \\ \mathbf{u} = \begin{cases} \mathbf{u}_i & \text{if} \quad \xi < \xi_{il}, \\ \mathbf{u}_l & \text{if} \quad \xi > \xi_{il}. \end{cases} \end{cases} \tag{3.20}$$

The calculation also leads to defining the numerical flux $\mathfrak{f}(\mathbf{u}(\xi_{il}))$ in a way that preserves the direction of propagation $\mathbf{u}$ and that for this reason earns it the appellation *upwind flux*.
As pointed out earlier in this section, the problem of numerical flux definition exists in DG methods as well and a similar procedure is employed, justifying our interest in studying $(3.10)|_{D_n=D_p=0}$, although an important difference needs to be remarked. Contrarily to (3.20), in the DG framework the state $\mathbf{u}_i$ ($\mathbf{u}_l$) is not constant and the initial condition is restated in terms of a left (right) limit:

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} + \dfrac{\partial \mathfrak{f}(\mathbf{u})}{\partial \xi} = \mathbf{0}, \\ \mathbf{u} = \begin{cases} \lim\limits_{\xi \to \xi_{il}^-} \mathbf{u}_i(\xi) & \text{if} \quad \xi < \xi_{il}, \\ \lim\limits_{\xi \to \xi_{il}^+} \mathbf{u}_l(\xi) & \text{if} \quad \xi > \xi_{il}. \end{cases} \end{cases} \tag{3.21}$$

The steps are illustrated in Figure 3.1. Further details on the topic can be found in [Koe11], [Viq15]; here, we wish to highlight the main difficulty in applying the procedure to the Maxwell-Drift system and calculating the upwind flux for $[\mathbf{E}, \mathbf{H}, n, p]_*$ – the physical flux in (3.10) is non-linear, as it contains field-concentration products. Indeed, if solving a linear RP is usually feasible, dealing with a nonlinear conservation law can be quite difficult, so that the general rule is to first linearize the physical flux, then exactly solve the approximated (linear) RP.

The most notable linearization is Roe's (see [GR91] for a general introduction, [HW02] for its application to DG methods). To present it, we set a suitable notation for the advective flow along $\hat{\mathbf{n}}_{il} = (\hat{n}_x, \hat{n}_y, \hat{n}_z)$, with

$$\begin{aligned} \mathfrak{f} &:= -Q^{-1}\left(\hat{n}_x \mathbf{f}^{(1)} + \hat{n}_y \mathbf{f}^{(2)} + \hat{n}_z \mathbf{f}^{(3)}\right) \in \mathbb{R}^d, \\ \mathcal{A} &:= \partial_{\mathbf{u}} \mathfrak{f} = -Q^{-1}\left(\hat{n}_x \partial_{\mathbf{u}} \mathbf{f}^{(1)} + \hat{n}_y \partial_u \mathbf{f}^{(2)} + \hat{n}_z \partial_{\mathbf{u}} \mathbf{f}^{(3)}\right) \in \mathbb{R}^{d,d}. \end{aligned} \tag{3.22}$$

The procedure prescribes to replace $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^{d,d}$ by $\tilde{\mathcal{A}} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d,d}$ evaluated at $(\mathbf{u}_i, \mathbf{u}_l)$, such that:

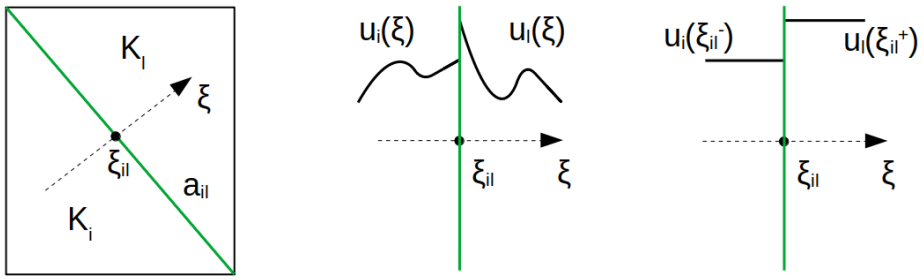


Figure 3.1: The steps of the definition of a RP in the DG framework. The numerical solution $\mathbf{u}_h$ is discontinuous on the interface a_{il} between two adjacent elements K_i and K_l. To define it, the conservation law is restated in one spatial variable ξ along the normal pointing from K_i to K_l, and closed with an initial condition that is discontinuous at the edge abscissa ξ_{il}. The left (right) value is the left (right) limit of the internal trace of $\mathbf{u}_i = \mathbf{u}_h|_{K_i}$ ($\mathbf{u}_l = \mathbf{u}_h|_{K_l}$).

1. $\mathfrak{f}(\mathbf{w}) - \mathfrak{f}(\mathbf{u}) = \tilde{\mathcal{A}}(\mathbf{u}, \mathbf{w})(\mathbf{w} - \mathbf{u}) \qquad \forall\ \mathbf{u}, \mathbf{w} \in \mathbb{R}^d$;
2. $\tilde{\mathcal{A}}$ is diagonalizable and its eigenvalues are real, i.e. the linearized conservation law is still hyperbolic;
3. $\tilde{\mathcal{A}}(\mathbf{u}, \mathbf{u}) = \mathcal{A}(\mathbf{u})$;

and $\tilde{\mathcal{A}}$ is called a *Roe's matrix*.
Some additional definitions on the way to the final expression:

$$
\begin{aligned}
&\Lambda = diag\Big(\lambda_1, \lambda_2, \ \dots\ , \lambda_d\Big), \text{ with } \lambda_r \ r\text{-th eigenvalue of } \tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l),\\
&\mathcal{U} = \text{matrix whose columns are the right eigenvectors of } \tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l), \text{ i.e. } \tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l) = \mathcal{U}\ \Lambda\ \mathcal{U}^{-1},\\
&\Lambda^+ = diag\Big(\ \max(\lambda_1, 0), \max(\lambda_2, 0), \ \dots\ , \max(\lambda_d, 0)\Big),\\
&\Lambda^- = diag\Big(\ \min(\lambda_1, 0), \min(\lambda_2, 0), \ \dots\ , \min(\lambda_d, 0)\Big),\\
&\quad \text{and} \quad |\tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l)| = \mathcal{U}\ (\Lambda^+ - \Lambda^-)\ \mathcal{U}^{-1}.
\end{aligned}
\tag{3.23}
$$

Finally, letting $\mathfrak{f}_i = \mathfrak{f}(\mathbf{u}_i)$, we can state Roe's numerical flux:

$$
\mathfrak{f}_* = \frac{\mathfrak{f}_i + \mathfrak{f}_l}{2} - |\tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l)|\ \frac{\mathbf{u}_l - \mathbf{u}_i}{2}. \tag{3.24}
$$

Well-established strategies exist to find a Roe matrix in ideal gas dynamics [GR91], whereas the task is not straightforward for an arbitrary problem. The simplest choice is $\tilde{\mathcal{A}}(\mathbf{u}_i, \mathbf{u}_l) = \mathcal{A}(\frac{\mathbf{u}_i + \mathbf{u}_l}{2})$, on which at least two notable schemes are based.

The first is the Lax-Wendroff scheme [GR91]

$$
\mathfrak{f}_* = \frac{\mathfrak{f}_i + \mathfrak{f}_l}{2} - \lambda^{-1}\mathcal{A}\left(\frac{\mathbf{u}_i + \mathbf{u}_l}{2}\right)\ \frac{\mathfrak{f}_l - \mathfrak{f}_i}{2}. \tag{3.25}
$$

The definition of λ is not immediate and deserves a brief digression on the scalar version of 3.18, that is

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0. \tag{3.26}$$

The Lax-Wendroff scheme is usually introduced alongside a more famous one, namely Lax-Friedrichs, which will be discussed in the next section. The point here is that, assuming space and time grids of sizes Δx and Δt, the two schemes share the definition

$$\lambda = \frac{\Delta x}{\Delta t}, \tag{3.27}$$

but also the stability requirement

$$\lambda^{-1} \max_u |f'(u)| \leq 1. \tag{3.28}$$

Hence, in practice, λ is typically identified with (and replaced by) $\max_u |f'(u)|$ when implementing the numerical flux. Following this philosophy, a generalization to multiple dimensions that is considered in DG schemes is

$$\lambda = \max\{|\lambda_{i1}|, ..., |\lambda_{id_i}|, |\lambda_{l1}|, ..., |\lambda_{ld_l}|\}, \tag{3.29}$$

with $\lambda_{ij} = j$-th eigenvalue of $\mathcal{A}(\mathbf{u}_i)$ [HW02].[6]

The second scheme, proposed in [GM01], reads

$$\mathfrak{f}_* = \frac{\mathfrak{f}_i + \mathfrak{f}_l}{2} - \tilde{\lambda}\frac{\mathbf{u}_l - \mathbf{u}_i}{2}, \tag{3.30}$$

where $\tilde{\lambda}$ is the maximum eigenvalue of $\mathcal{A}(\frac{\mathbf{u}_i+\mathbf{u}_l}{2})$.

To conclude the subsection we mention a further notable, yet simple strategy, called *flux splitting*, which exclusively applies to non-linear conservation laws in which $\mathfrak{f}$ depends on the unknowns but is homogeneous: $\mathfrak{f}(a\mathbf{u}) = a\ \mathfrak{f}(\mathbf{u}),\ \ \forall a \in \mathbb{R}$. The splitting consists in

$$\mathfrak{f}_* = \mathcal{A}^+(\mathbf{u}_i)\mathbf{u}_i + \mathcal{A}^-(\mathbf{u}_l)\mathbf{u}_l, \tag{3.31}$$

with $\mathcal{A}^{\pm}(\mathbf{u})$ defined along the lines of (3.23), that is

$$\begin{aligned}
\Lambda &= diag\Big(\lambda_1, \lambda_2,\ ...\ , \lambda_d\Big),\ \text{with } \lambda_r\ r\text{-th eigenvalue of } \mathcal{A}(\mathbf{u}),\\
\mathcal{U} &= \text{matrix whose columns are the right eigenvectors of } \mathcal{A}(\mathbf{u}), \text{ i.e. } \mathcal{A}(\mathbf{u}) = \mathcal{U}\ \Lambda\ \mathcal{U}^{-1},\\
\Lambda^+ &= diag\Big(\ \max(\lambda_1, 0), \max(\lambda_2, 0),\ ...\ , \max(\lambda_d, 0)\Big),\\
\Lambda^- &= diag\Big(\ \min(\lambda_1, 0), \min(\lambda_2, 0),\ ...\ , \min(\lambda_d, 0)\Big),\\
\mathcal{A}^{\pm}(\mathbf{u}) &= \mathcal{U}\ \Lambda^{\pm}\ \mathcal{U}^{-1}.
\end{aligned} \tag{3.32}$$

[6]The Lax-Friedrichs flux is not based on a Roe matrix, this is the reason why it is not reported here. For an explicit expression, and particularly in application to hole advection, the reader is referred to (3.37).
Moreover, it should be noted for the sake of completeness that (3.27) implies that space and time are uniformly discretized. However, considering a more general case would not change the conclusion (3.29).

The primary domain of application of flux splitting is ideal gas dynamics [GR91], while unfortunately (3.10) does not comply with the homogeneity requirement. Nevertheless, (3.31) offers an insight into the philosophy behind the numerical fluxes for advection presented in the next section, thanks to its *upwind* character. Indeed, with reference to the flow along $\hat{\mathbf{n}}_{il}$, i.e. across $a_{il} = K_i \cap K_l$ from K_i to K_l, it is observed that $\mathcal{A}^+$ ($\mathcal{A}^-$) contains positive (negative) velocities, thus represents outflow (inflow), and is associated to the internal (external) and thus upstream value of $\mathbf{u}$.

3.1.1.3 Composing the advective flux

Following references on DG methods for linear Maxwell's equations [HW02], [Koe11], [Viq15], we report the upwind flux obtained by solving a Riemann problem across a_{il}:

$$\mathbf{E}_* = \frac{Y_i\mathbf{E}_i + Y_l\mathbf{E}_l + \alpha\hat{\mathbf{n}}_{il} \times (\mathbf{H}_l - \mathbf{H}_i)}{Y_i + Y_l}, \qquad \mathbf{H}_* = \frac{Z_i\mathbf{H}_i + Z_l\mathbf{H}_l - \alpha\hat{\mathbf{n}}_{il} \times (\mathbf{E}_l - \mathbf{E}_i)}{Z_i + Z_l}, \tag{3.33}$$

where $Z_i = \sqrt{\epsilon_i/\mu_i}$ and $Y_i = 1/Z_i$ are the characteristic impedance and admittance of the medium in the element K_i. The *upwind parameter* $\alpha \in [0, 1]$ is used to control numerical dissipation – for $\alpha = 1$ (fully upwind flux) the Riemann problem is solved exactly, while taking $\alpha \in]0, 1[$ leads to approximate solutions; another notable choice is $\alpha = 0$ (centered flux), which introduces no dissipation. The resulting scheme is numerically stable and convergent, and enforces continuity of $\hat{\mathbf{n}}_{il} \times \mathbf{E}$ and $\hat{\mathbf{n}}_{il} \times \mathbf{H}$ by penalizing their jumps.

As to drift currents, we focus for example on the hole advection equation

$$\frac{\partial p}{\partial t} + \nabla \cdot (p\ \boldsymbol{v}_p) = 0, \tag{3.34}$$

and review the most interesting options identified in the literature, which are listed below.

- The Lax-Friedrichs flux (considered in [LS04] and [LS07] in one dimension)

$$(pv_p)_* = \frac{p_l v_{p,l} + p_i v_{p,i}}{2} - \lambda_p \frac{p_l - p_i}{2}, \tag{3.35}$$

  with

$$\lambda_p = \max_i(|v_{p,i}|), \tag{3.36}$$

  as anticipated in the previous subsection when discussing (3.25).[7] The maximum is global, i.e. over the whole mesh; in the case of an explicit time integration scheme, it is computed at each step.

  The extension of (3.35) to multi-dimensional DG methods, with focus on the flow across a_{il} along $\hat{\mathbf{n}}_{il}$, is [HW02]

$$(p\boldsymbol{v}_p)_* = \frac{p_l \boldsymbol{v}_{p,l} + p_i \boldsymbol{v}_{p,i}}{2} - \lambda_p \hat{\mathbf{n}}_{il} \frac{p_l - p_i}{2}, \tag{3.37}$$

  where

$$\lambda_p = \max(|\hat{\mathbf{n}}_{il} \cdot \boldsymbol{v}_p|). \tag{3.38}$$

[7] Note that here $u = p$ and f is a vector field. If we denote it as $\boldsymbol{f}(p) = p\boldsymbol{v}_p$, we notice that it is a homogeneous function of p, so that $\boldsymbol{f}'(p) = \boldsymbol{v}_p$.

A common variant is the *local* Lax-Friedrichs flux. In one dimension this is [CS89]

$$\lambda_p = \max(|v_{p,i}|, |v_{p,l}|). \tag{3.39}$$

For the extension to multiple dimensions, one possibility is to evaluate the peak velocity on each side of a_{il} and select the highest of the two values [D18]

$$\lambda_p = \max_{a_{il}}(|\hat{\mathbf{n}}_{il} \cdot \boldsymbol{v}_{p,i}|, |\hat{\mathbf{n}}_{il} \cdot \boldsymbol{v}_{p,l}|). \tag{3.40}$$

Alternatively one can evaluate λ_n and λ_p at each node pair on a_{il} [Brd12], but K_i and K_l will not be allowed to have different interpolation degrees – an important liberty, peculiar to DG methods, that we would rather not lose.

As remarked in [D18], in the context of nonlinear optics, the local version of the Lax-Friedrichs flux yields a less dissipative and thus more accurate scheme.

- The Lax-Wendroff flux

$$(p\boldsymbol{v}_p)_* = \frac{p_l\boldsymbol{v}_{p,l} + p_i\boldsymbol{v}_{p,i}}{2} - \lambda_p^{-1}\left(\frac{\boldsymbol{v}_{p,l} + \boldsymbol{v}_{p,i}}{2} \cdot \hat{\mathbf{n}}_{il}\right)\frac{p_l\boldsymbol{v}_{p,l} - p_i\boldsymbol{v}_{p,i}}{2}, \tag{3.41}$$

defined along the lines of (3.25) with λ_p given by (3.38).

- Flux splitting, proposed in one dimension in [LS10a], [CC94], [CT92].
It has been already pointed out that flux splitting is not applicable to $[\mathbf{H}, \mathbf{E}, n, p]$ due to drift current depending non-linearly on the conserved quantity. Here, on the other hand, we are isolating charge transport from Maxwell's equations with the aim of identifying a compatible numerical flux to be imported back into the coupled system, thus the conserved quantity consists in charge density and does not include $\mathbf{E}$.
As highlighted in Section 3.1.1.2, the principle of flux splitting is: if electric charge flows outside K_i, the interior trace is selected for the corresponding carrier density; conversely, if charge flows into K_i, the exterior trace is selected. So, on the i side of a_{il} the authors take

$$(\hat{\mathbf{n}}_{il} \cdot p\boldsymbol{v}_p)_* = \min(0, \hat{\mathbf{n}}_{il} \cdot \boldsymbol{v}_p)\ p_l + \max(0, \hat{\mathbf{n}}_{il} \cdot \boldsymbol{v}_p)\ p_i. \tag{3.42}$$

However, a compatibility issue arises: (3.42) requires $\boldsymbol{v}_p = \mu_p\mathbf{E}$ to have a single trace on a_{il}. As a matter of fact, the underlying electric field is approximated with continuous finite elements in the references.
At first glance one might be tempted to derive an expression of $\boldsymbol{v}_p$ in terms of $\mathbf{E}_*$ from (3.33). In fact, it should be noted that (3.42) coincides with the Lax-Friedrichs flux (3.37) for $\boldsymbol{v}_i = \boldsymbol{v}_l$. Hence, the latter has the same upwind character as (3.42) (plus the advantage of a more elegant and less complex implementation).

3.1.1.4 Handling diffusion

To deal with $\mathbf{J}_{n,diff}, \mathbf{J}_{p,diff}$ we follow [LS10b]. The idea is to define a "wind" direction $\boldsymbol{w}$ in the computational domain, so that an upwind and a downwind side with respect to $\boldsymbol{w}$ can be identified on each face a_{il} depending on the angle between $\hat{\mathbf{n}}_{il}$ and $\boldsymbol{w}$. Then, n_* is always taken on the upwind side and $\mathbf{J}_{n_*,diff}$ on the downwind one, or viceversa. Similarly for p_* and $\mathbf{J}_{p_*,diff}$. The criterion is

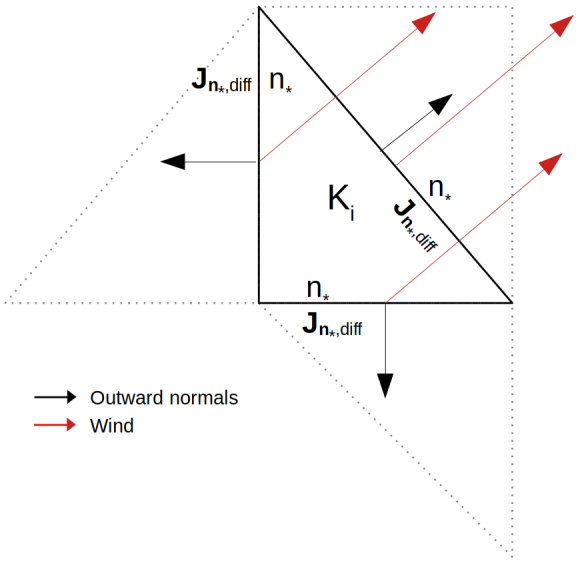


Figure 3.2: A two-dimensional illustration of alternate flux computation on the boundary of K_i. When $\hat{\mathbf{n}} \cdot \boldsymbol{w} < 0$ the downwind side for n and the upwind one for $\mathbf{J}_{n,diff}$ are selected. Conversely when $\hat{\mathbf{n}} \cdot \boldsymbol{w} > 0$. Here, for example, $\boldsymbol{w} = (1, 1)$ and the orientation of K_i is such that $n_* = n_i$ on the catheti, $\mathbf{J}_{n_*,diff} = \mathbf{J}_{n_l,diff}$ on the hypotenuse.

illustrated in Figure 3.2. In formulae, on the interface between elements K_i and K_l the alternate flux is [8]

$$\begin{aligned} \theta &= \frac{1}{2} + \frac{b}{2}\ sign(\hat{\mathbf{n}}_{il} \cdot \mathbf{w}), \\ b &\in \{-1, 0, 1\}, \\ n_* &= (1-\theta)\ n_i + \theta\ n_l, \\ \mathbf{J}_{n_*,diff} &= \theta\ \mathbf{J}_{n_i,diff} + (1-\theta)\ \mathbf{J}_{n_l,diff}, \\ p_* &= (1-\theta)\ p_i + \theta\ p_l, \\ \mathbf{J}_{p_*,diff} &= \theta\ \mathbf{J}_{p_i,diff} + (1-\theta)\ \mathbf{J}_{p_l,diff}. \end{aligned} \tag{3.43}$$

Note that b (thus θ) is fixed prior to the simulation. The setting $b = 1$ implies $\theta = 0$ for $\hat{\mathbf{n}} \cdot \boldsymbol{w} > 0$ and $\theta = 1$ for $\hat{\mathbf{n}} \cdot \boldsymbol{w} < 0$, so that n_* is on the upwind side and $\mathbf{J}_{n_*,diff}$ on the downwind one. The opposite scenario stems from $b = -1$. Another notable option is $b = 0$ (centered fluxes).

3.1.1.5 The final decision

For a system describing light propagation, (3.25) seems inconvenient – dividing the jump term by the characteristic speed (the speed of light) could make it so small to practically center the flux and therefore neglect flow direction.

[8] For simplicity and by virtue of the two equations being identical (up to a sign), the same parameter θ is used for both electrons and holes.

As to (3.30), the main drawback regards the difference in magnitude between characteristic speeds in the Maxwell-Drift system. It is expected that the speed of light lead and that $\tilde{\lambda} = \frac{1}{\sqrt{\epsilon\mu}} >> v_{drift}$, yielding an oversized jump term for the semiconductor part of the flux; in analogy with the Lax-Friedrichs flux presented in Section 3.1.1.3, we should expect such a large jump term to produce excessive dissipation.

In the end, consideration of a) the difficulty of finding a Roe matrix that would preserve light and particle flow at their different speeds, and b) the path traced in the literature on DGTD methods for semiconductor physics and Maxwell's equations, has led to the decision of composing the numerical flux and selecting:

- the upwind flux (3.33) for Maxwell's equations (in practice, we shall always assume $\alpha = 1$);
- the alternate flux (3.43) for carrier densities and diffusion currents;
- the Lax-Friedrichs flux (3.37), in its local version, for drift currents.

Before injecting the selected fluxes in the numerical formulation a few considerations on the drift currents are necessary, to discuss the extension of the Lax-Friedrichs flux to electron transport as well as the potential impact of an AC-DC decomposition.

3.1.1.6 The Lax-Friedrichs flux for bipolar transport

In terms of velocity, which defines the direction in which particles move, the drift terms of charge continuity equations have identical forms:

$$\begin{aligned} \frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{v}_n) &= \textit{diffusion and reaction terms}, \\ \frac{\partial p}{\partial t} + \nabla \cdot (p\boldsymbol{v}_p) &= \textit{diffusion and reaction terms}. \end{aligned} \tag{3.44}$$

Drift velocity is due to an electric field $\mathbf{E}$, and whether its nature be static, optical or a superposition of the two, negative (positive) charges will move opposite to (in the direction of) it. This is reproduced by the relations

$$\begin{aligned} \boldsymbol{v}_n &= -\mu_n \mathbf{E}, \\ \boldsymbol{v}_p &= \mu_p \mathbf{E}. \end{aligned} \tag{3.45}$$

Electric current conventionally has the same direction as the electric field, hence

$$\begin{aligned} \mathbf{J}_{n,drift} &= -qn\boldsymbol{v}_n = qn\mu_n \mathbf{E}, \\ \mathbf{J}_{p,drift} &= qp\boldsymbol{v}_p = qp\mu_p \mathbf{E}, \end{aligned} \tag{3.46}$$

which explains the sign difference

$$\begin{aligned} \frac{\partial n}{\partial t} - \nabla \cdot \mathbf{J}_{n,drift} &= \textit{diffusion and reaction terms}, \\ \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{J}_{p,drift} &= \textit{diffusion and reaction terms}. \end{aligned} \tag{3.47}$$

As a result, (3.37) should:

- be the same for n and p when stated in terms of $\boldsymbol{v}$;

- have different signs ahead of λ when stated in terms of $\mathbf{J}$ or $\mathbf{E}$; the upwind parameter λ is not affected as it depends on $|\boldsymbol{v}|$.

The model has been stated in terms of $\mathbf{J}$ and $\mathbf{E}$ so far and this shall be the rule in the forthcoming formulae – drift velocity will reappear in Section 4.2.

3.1.1.7 The impact of AC-DC decomposition on the numerical flux

It was observed in (2.55) that AC-DC decomposition yields cancellation of purely static terms and it can be noticed that the only significant change in the resulting system lies in drift currents. The hole one, for example, unfolds as:

$$\mathbf{J}_{p,drift} = q\ p_{AC}\ \mu_p(\mathbf{E}_{DC})\ (\mathbf{E}_{AC}+\mathbf{E}_{DC}) + q\ p_{DC}\ \mu_p(\mathbf{E}_{DC})\ \mathbf{E}_{AC} + q\ p_{DC}\ \mu_p(\mathbf{E}_{DC})\ \mathbf{E}_{DC}. \quad (3.48)$$

The expansion does not affect the way volume integrals are calculated. On the other hand, numerical flux definition requires some care. Letting the opaque terms out, (3.37) should translate to

$$\begin{aligned}\mathbf{J}_{p_*,drift} &= \frac{\mathbf{J}_{p_{l},drift}+\mathbf{J}_{p_{i},drift}}{2} - q\lambda_p\frac{p_l-p_i}{2}\hat{\mathbf{n}}_{il} \\ &= q\frac{\Big[p_{AC}\ \mu_p(\mathbf{E}_{DC})\ (\mathbf{E}_{AC}+\mathbf{E}_{DC})\Big]_l + \Big[p_{AC}\ \mu_p(\mathbf{E}_{DC})\ (\mathbf{E}_{AC}+\mathbf{E}_{DC})\Big]_i}{2} - q\lambda_p\frac{p_{AC,l}-p_{AC,i}}{2}\hat{\mathbf{n}}_{il} \\ &+ q\frac{\Big[p_{DC}\ \mu_p(\mathbf{E}_{DC})\ \mathbf{E}_{AC}\Big]_l + \Big[p_{DC}\ \mu_p(\mathbf{E}_{DC})\ \mathbf{E}_{AC}\Big]_i}{2} - q\lambda_p\frac{p_{DC,l}-p_{DC,i}}{2}\hat{\mathbf{n}}_{il},\end{aligned} \quad (3.49)$$

with the local upwind parameter (3.40) becoming

$$\lambda_p = \max_{a_{il}}\left(\Big|\mu_p(\mathbf{E}_{DC,i})\ \hat{\mathbf{n}}_{il}\cdot(\mathbf{E}_{AC,i}+\mathbf{E}_{DC,i})\Big|\ ,\ \Big|\mu_p(\mathbf{E}_{DC,l})\ \hat{\mathbf{n}}_{il}\cdot(\mathbf{E}_{AC,l}+\mathbf{E}_{DC,l})\Big|\right). \quad (3.50)$$

The last line features an average term and a penalization term, which demand some remarks.
The penalization term is opaque because it is not rigorous: p_{DC} does not represent a transported quantity and hence no flux at all should be associated to it; b) the upwind parameter should be calculated from the velocity appearing in the average term ($\mu_p\mathbf{E}_{AC}$) whereas it actually comes from the upwind flux of p_{AC} and is therefore based on the velocity $\mu_p(\mathbf{E}_{AC}+\mathbf{E}_{DC})$.
The average term is recognized as the centered flux for $\mathcal{J}_p := q\ p_{DC}\ \mu_p(\mathbf{E}_{DC})\ \mathbf{E}_{AC}$ and has its right in (3.50) because it is discontinuous – it contains the two-valued $\mathbf{E}_{AC}$. Then, since p_{DC} and $\mu_p(\mathbf{E}_{DC})$ are continuous in our analysis, the expression leads to the centered flux for $\mathbf{E}_{AC}$ itself, and we might be tempted to rather take $\mathbf{E}_{AC}$ from the upwind flux (3.33); in fact, handling the whole $\mathcal{J}_p$ as a discontinuous quantity leaves room for an interesting option: calculating $n_{DC}, p_{DC}, \mathbf{E}_{DC}$ with another discontinuous Galerkin solver.

The extension to $\mathbf{J}_{n,drift}$ requires a change in the sign ahead of the upwind parameter, giving

$$\begin{aligned}\mathbf{J}_{n_*,drift} &= \frac{\mathbf{J}_{n_{l,drift}} + \mathbf{J}_{n_{i,drift}}}{2} \;+\; q\lambda_n \frac{n_l - n_i}{2}\hat{\mathbf{n}}_{il} \\ &= q\frac{\Big[n_{AC}\ \mu_n(\mathbf{E}_{DC})\ (\mathbf{E}_{AC}+\mathbf{E}_{DC})\Big]_l + \Big[n_{AC}\ \mu_n(\mathbf{E}_{DC})\ (\mathbf{E}_{AC}+\mathbf{E}_{DC})\Big]_i}{2} \;+\; q\lambda_n \frac{n_{AC,l} - n_{AC,i}}{2}\hat{\mathbf{n}}_{il} \\ &+ q\frac{\Big[n_{DC}\ \mu_n(\mathbf{E}_{DC})\ \mathbf{E}_{AC}\Big]_l + \Big[n_{DC}\ \mu_n(\mathbf{E}_{DC})\ \mathbf{E}_{AC}\Big]_i}{2} \;+\; q\lambda_n \frac{n_{DC,l} - n_{DC,i}}{2}\hat{\mathbf{n}}_{il},\end{aligned} \tag{3.51}$$

with λ_n being $(3.50)|_{p=n}$.

In the end, drift current flux on ∂K_i should be intended as per (3.49) and (3.51). That said, unfolding such expressions will not be necessary in the numerical formulation itself, but will become essential when discussing its implementation (Section 4.2). Similarly, the fact that mobility and diffusivity depend on $\mathbf{E}_{DC}$ but not on $\mathbf{E}_{AC}$ will be kept implicit.

Before moving forward, a side remark on DC quantities is necessary.
Simplification of the purely static terms in (2.55) was legitimated by the fact of working in the continuous framework, in which these exactly solve the static PDD model (2.50). This led to our discrete model featuring $\mathbf{E}_{AC}$,$\mathbf{H}_{AC}$, n_{AC}, p_{AC} as the unknowns and $\mathbf{E}_{DC}$, n_{DC}, p_{DC} as (space-varying) coefficients. In fact, the latter also result from a numerical scheme, at least in all practical applications; thus, we should bear in mind that they carry a numerical error which drives them away from the exact solution of (2.50) and is inherited by the unknowns of the transient model (2.55).

3.1.2 Semi-discrete DG formulation: matrix form

With the integral formulation (3.6) and the developments of Section 3.1.1 we write a matrix form of the semi-discrete DG formulation. Starting with (3.33) we preliminarily note that

$$\begin{aligned}[Y_i\mathbf{E}_i - \alpha(\hat{\mathbf{n}}_{il}\times\mathbf{H}_i)]\times\hat{\mathbf{n}}_{il} &= Y_i\mathbf{E}_i\times\hat{\mathbf{n}}_{il} - \alpha[\mathbf{H}_i - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il}\cdot\mathbf{H}_i)], \\ [Y_l\mathbf{E}_l + \alpha(\hat{\mathbf{n}}_{il}\times\mathbf{H}_l)]\times\hat{\mathbf{n}}_{il} &= Y_l\mathbf{E}_l\times\hat{\mathbf{n}}_{il} + \alpha[\mathbf{H}_l - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il}\cdot\mathbf{H}_l)], \\ [Z_i\mathbf{H}_i + \alpha(\hat{\mathbf{n}}_{il}\times\mathbf{E}_i)]\times\hat{\mathbf{n}}_{il} &= Z_i\mathbf{H}_i\times\hat{\mathbf{n}}_{il} + \alpha[\mathbf{E}_i - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il}\cdot\mathbf{E}_i)], \\ [Z_l\mathbf{H}_l - \alpha(\hat{\mathbf{n}}_{il}\times\mathbf{E}_l)]\times\hat{\mathbf{n}}_{il} &= Z_l\mathbf{H}_l\times\hat{\mathbf{n}}_{il} - \alpha[\mathbf{E}_l - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il}\cdot\mathbf{E}_l)].\end{aligned} \tag{3.52}$$

Now, if we introduce the unit vector $\hat{\mathbf{v}} = \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ and the vectors

$$\mathbf{E}_{ij} = [E^x_{ij}\ E^y_{ij}\ E^x_{ij}]^T \qquad \mathbf{H}_{ij} = [H^x_{ij}\ H^y_{ij}\ H^x_{ij}]^T, \tag{3.53}$$

containing the j-th numerical component (with respect to the local basis of K_i) of the Cartesian

components of $\mathbf{E}$ and $\mathbf{H}$, we can express the integrals as

$$\begin{aligned}
\int_{a_{il}} \mathbf{E}_* \times \hat{\mathbf{n}}_{il} \cdot \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ik}} &= \sum_{j=1}^{d_i} \frac{Y_i \mathbf{E}_{ij} \times \hat{\mathbf{n}}_{il} - \alpha[\mathbf{H}_{ij} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{H}_{ij})]}{Y_i + Y_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{ij}\phi_{ik} \\
&\quad + \sum_{g=1}^{d_l} Y_l \frac{\mathbf{E}_{lg} \times \hat{\mathbf{n}}_{il} + \alpha[\mathbf{H}_{lg} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{H}_{lg}])}{Y_i + Y_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{lg}\phi_{ik}, \\
\int_{a_{il}} \mathbf{H}_* \times \hat{\mathbf{n}}_{il} \cdot \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ik}} &= \sum_{j=1}^{d_i} \frac{Z_i \mathbf{H}_{ij} \times \hat{\mathbf{n}}_{il} + \alpha[\mathbf{E}_{ij} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{ij}])}{Z_i + Z_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{ij}\phi_{ik} \\
&\quad + \sum_{g=1}^{d_l} Z_l \frac{\mathbf{H}_{lg} \times \hat{\mathbf{n}}_{il} - \alpha[\mathbf{E}_{lg} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{lg})]}{Z_i + Z_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{lg}\phi_{ik}.
\end{aligned} \tag{3.54}$$

We take advantage of $\hat{\mathbf{v}}$ to also recast the integrals

$$\begin{aligned}
\int_{K_i} \mathbf{E}_i \cdot \nabla \times \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ij}} &= -\sum_{j=1}^{d_i} \mathbf{E}_{ij} \cdot (\hat{\mathbf{v}} \times \int_{K_i} \phi_{ij} \nabla \phi_{ik}), \\
\int_{K_i} \mathbf{H}_i \cdot \nabla \times \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ij}} &= -\sum_{j=1}^{d_i} \mathbf{H}_{ij} \cdot (\hat{\mathbf{v}} \times \int_{K_i} \phi_{ij} \nabla \phi_{ik}).
\end{aligned} \tag{3.55}$$

Moving to diffusion, (3.43) yields the explicit expressions

$$\begin{aligned}
\int_{a_{il}} n_* \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ik}} \cdot \hat{\mathbf{n}}_{il} &= (1-\theta) \sum_{j=1}^{d_i} n_{ij} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{ij}\phi_{ik} + \theta \sum_{g=1}^{d_l} n_{lg} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{lg}\phi_{ik}, \\
\int_{a_{il}} p_* \boldsymbol{\phi}^{\boldsymbol{v}}_{\boldsymbol{ik}} \cdot \hat{\mathbf{n}}_{il} &= (1-\theta) \sum_{j=1}^{d_i} p_{ij} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{ij}\phi_{ik} + \theta \sum_{g=1}^{d_l} p_{lg} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{lg}\phi_{ik}, \\
\int_{a_{il}} \phi_{ik} \big(D_n \boldsymbol{\alpha}\big)_* \cdot \hat{\mathbf{n}}_{il} &= \int_{a_{il}} \phi_{ik} \big[\theta \big(D_n \boldsymbol{\alpha}\big)_i + (1-\theta) \big(D_n \boldsymbol{\alpha}\big)_l\big] \cdot \hat{\mathbf{n}}_{il}, \\
\int_{a_{il}} \phi_{ik} \big(D_p \boldsymbol{\beta}\big)_* \cdot \hat{\mathbf{n}}_{il} &= \int_{a_{il}} \phi_{ik} \big[\theta \big(D_p \boldsymbol{\beta}\big)_i + (1-\theta) \big(D_p \boldsymbol{\beta}\big)_l\big] \cdot \hat{\mathbf{n}}_{il}.
\end{aligned} \tag{3.56}$$

At this point we are ready to write the semi-discrete DG formulation. This time, to keep the notation compact, currents are expanded in separate equations.

$$
\begin{aligned}
\sum_{j=1}^{d_i} \frac{\partial \mathbf{H}_{ij}}{\partial t} \cdot \hat{\mathbf{v}} \int_{K_i} \mu\ \phi_{ij}\phi_{ik} &= \sum_{j=1}^{d_i} \mathbf{E}_{ij} \cdot \big(\hat{\mathbf{v}} \times \int_{K_i} \phi_{ij} \nabla \phi_{ik}\big) \\
&+ \sum_{l \in \mathcal{N}_i} \Big(\sum_{j=1}^{d_i} \frac{Y_i \mathbf{E}_{ij} \times \hat{\mathbf{n}}_{il} - \alpha[\mathbf{H}_{ij} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{H}_{ij}]}{Y_i + Y_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{ij}\phi_{ik} \\
&+ \sum_{g=1}^{d_l} \frac{Y_l \mathbf{E}_{lg} \times \hat{\mathbf{n}}_{il} + \alpha[\mathbf{H}_{lg} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{H}_{lg}]}{Y_i + Y_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{lg}\phi_{ik} \Big), \\
\sum_{j=1}^{d_i} \frac{\partial \mathbf{E}_{ij}}{\partial t} \cdot \hat{\mathbf{v}} \int_{K_i} \epsilon\ \phi_{ij}\phi_{ik} &= -\sum_{j=1}^{d_i} \mathbf{H}_{ij} \cdot \big(\hat{\mathbf{v}} \times \int_{K_i} \phi_{ij} \nabla \phi_{ik}\big) \\
&- \sum_{l \in \mathcal{N}_i} \Big(\sum_{j=1}^{d_i} \frac{Z_i \mathbf{H}_{ij} \times \hat{\mathbf{n}}_{il} + \alpha[\mathbf{E}_{ij} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{ij}]}{Z_i + Z_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{ij}\phi_{ik} \\
&+ \sum_{g=1}^{d_l} \frac{Z_l \mathbf{H}_{lg} \times \hat{\mathbf{n}}_{il} - \alpha[\mathbf{E}_{lg} - \hat{\mathbf{n}}_{il}(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{lg}]}{Y_i + Y_l} \cdot \hat{\mathbf{v}} \int_{a_{il}} \phi_{lg}\phi_{ik} \Big) \\
&- \int_{K_i} (\mathbf{J}_{n_i} + \mathbf{J}_{p_i}) \cdot \hat{\mathbf{v}}\ \phi_{ik},
\end{aligned}
\tag{3.57}
$$

$$
\begin{aligned}
q \sum_{j=1}^{d_i} \frac{\partial n_{ij}}{\partial t} \int_{K_i} \phi_{ij}\phi_{ik} &= \sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik}\ \mathbf{J}_{n_*} \cdot \hat{\mathbf{n}}_{il} - \int_{K_i} \mathbf{J}_{n_i} \cdot \nabla \phi_{ik} \\
&- q \int_{K_i} R(n_i, p_i)\ \phi_{ik} + q \sum_{j=1}^{d_i} G_{ij} \int_{K_i} \phi_{ij}\phi_{ik}, \\
q \sum_{j=1}^{d_i} \frac{\partial p_{ij}}{\partial t} \int_{K_i} \phi_{ij}\phi_{ik} &= -\sum_{l \in \mathcal{N}_i} \int_{a_{il}} \phi_{ik}\ \mathbf{J}_{p_*} \cdot \hat{\mathbf{n}}_{il} + \int_{K_i} \mathbf{J}_{p_i} \cdot \nabla \phi_{ik} \\
&- q \int_{K_i} R(n_i, p_i)\ \phi_{ik} + q \sum_{j=1}^{d_i} G_{ij} \int_{K_i} \phi_{ij}\phi_{ik},
\end{aligned}
\tag{3.58}
$$

$$
\begin{aligned}
\sum_{j=1}^{d_i} \boldsymbol{\alpha}_{ij} \cdot \hat{\mathbf{v}} \int_{K_i} \phi_{ij}\phi_{ik} &= \sum_{l \in \mathcal{N}_i} \Big((1-\theta) \sum_{j=1}^{d_i} n_{ij} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{ij}\phi_{ik} + \theta \sum_{g=1}^{d_l} n_{lg} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{lg}\phi_{ik} \Big) \\
&- \sum_{j=1}^{d_i} n_{ij}\ \hat{\mathbf{v}} \cdot \int_{K_i} \phi_{ij} \nabla \phi_{ik}, \\
\sum_{j=1}^{d_i} \boldsymbol{\beta}_{ij} \cdot \hat{\mathbf{v}} \int_{K_i} \phi_{ij}\phi_{ik} &= \sum_{l \in \mathcal{N}_i} \Big((1-\theta) \sum_{j=1}^{d_i} p_{ij} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{ij}\phi_{ik} + \theta \sum_{g=1}^{d_l} p_{lg} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{il} \int_{a_{il}} \phi_{lg}\phi_{ik} \Big) \\
&- \sum_{j=1}^{d_i} p_{ij}\ \hat{\mathbf{v}} \cdot \int_{K_i} \phi_{ij} \nabla \phi_{ik}.
\end{aligned}
\tag{3.59}
$$

The total electron and hole currents appearing in volume integrals are

$$\begin{aligned}\mathbf{J}_{n_i} &= \mathbf{J}_{n_{i,drift}} + \mathbf{J}_{n_{i,diff}} = q\,\left[n\;\mu_n\;\mathbf{E}\right]_i \;+\; q\,\left[D_n\;\boldsymbol{\alpha}\right]_i,\\ \mathbf{J}_{p_i} &= \mathbf{J}_{p_{i,drift}} + \mathbf{J}_{p_{i,diff}} = q\,\left[p\;\mu_p\;\mathbf{E}\right]_i \;-\; q\,\left[D_p\;\boldsymbol{\beta}\right]_i.\end{aligned} \tag{3.60}$$

Then, recovering (3.43), (3.49), (3.50) and (3.51), the drift-diffusion numerical flux is assembled as

$$\begin{aligned}\mathbf{J}_{n_*} &= \frac{\mathbf{J}_{n_{l,drift}} + \mathbf{J}_{n_{i,drift}}}{2} \;+\; q\lambda_n \frac{n_l - n_i}{2}\hat{\mathbf{n}}_{il} \;+\; \theta\;\mathbf{J}_{n_{i,diff}} + (1-\theta)\;\mathbf{J}_{n_{l,diff}},\\ \mathbf{J}_{p_*} &= \frac{\mathbf{J}_{p_{l,drift}} + \mathbf{J}_{p_{i,drift}}}{2} \;-\; q\lambda_p \frac{p_l - p_i}{2}\hat{\mathbf{n}}_{il} \;+\; \theta\;\mathbf{J}_{p_{i,diff}} + (1-\theta)\;\mathbf{J}_{p_{l,diff}}.\end{aligned} \tag{3.61}$$

To put (3.57-3.59) into a matrix form, a suitable notation is introduced. Let $v \in \{x, y, z\}$, $j \in \{1, 2,\ \ldots d_i\}$, $g \in \{1, 2,\ \ldots d_l\}$, where as usual $i, l \in \{1,\ \ldots N_{\mathcal{T}}\}$. We define the following symbols for vectors and block vectors of numerical coefficients, as well as stiffness, mass and flux matrices and block matrices:

$$\mathbf{A}_i^v = \begin{bmatrix} A_{i1}^v \\ \vdots \\ A_{id_i}^v \end{bmatrix}, \quad \overline{\mathbf{A}}_i = \begin{bmatrix} \mathbf{A}_i^x \\ \mathbf{A}_i^y \\ \mathbf{A}_i^z \end{bmatrix}, \quad \mathbf{a}_i = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{id_i} \end{bmatrix}, \quad \overline{\mathbf{a}}_i = \begin{bmatrix} \mathbf{a}_i \\ \mathbf{a}_i \\ \mathbf{a}_i \end{bmatrix},$$

$$\begin{aligned}\left(\mathbb{K}_i^v\right)_{kj} &= \int_{K_i} \phi_{ij}\frac{\partial}{\partial v}\phi_{ik}\;, & \overline{\mathbb{K}}_i &= \begin{bmatrix} \mathbb{K}_i^x \\ \mathbb{K}_i^y \\ \mathbb{K}_i^z \end{bmatrix},\\ \left(\mathbb{M}_i^\mu\right)_{kj} &= \int_{K_i} \mu\;\phi_{ij}\phi_{ik}\;, & \overline{\mathbb{M}}_i^\mu &= diag(\mathbb{M}_i^\mu, \mathbb{M}_i^\mu, \mathbb{M}_i^\mu),\\ \left(\mathbb{M}_i^\epsilon\right)_{kj} &= \int_{K_i} \epsilon\;\phi_{ij}\phi_{ik}\;, & \overline{\mathbb{M}}_i^\epsilon &= diag(\mathbb{M}_i^\epsilon, \mathbb{M}_i^\epsilon, \mathbb{M}_i^\epsilon),\\ \left(\mathbb{M}_i\right)_{kj} &= \int_{K_i} \phi_{ij}\phi_{ik}\;, & \overline{\mathbb{M}}_i &= diag(\mathbb{M}_i, \mathbb{M}_i, \mathbb{M}_i),\\ \left(\mathbb{S}_{il}\right)_{kj} &= \int_{a_{il}} \phi_{ij}\phi_{ik}\;, & \overline{\mathbb{S}}_{il} &= diag(\mathbb{S}_{il}, \mathbb{S}_{il}, \mathbb{S}_{il}),\\ \left(\mathbb{S}_{il}^*\right)_{kg} &= \int_{a_{il}} \phi_{lg}\phi_{ik}\;, & \overline{\mathbb{S}}_{il}^* &= diag(\mathbb{S}_{il}^*, \mathbb{S}_{il}^*, \mathbb{S}_{il}^*),\end{aligned} \tag{3.62}$$

$$\begin{aligned}&\hat{\mathbf{n}}_{il} = diag\Big(\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{x}}\;\mathbb{I}_{d_i}, \hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{y}}\;\mathbb{I}_{d_i}, \hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{z}}\;\mathbb{I}_{d_i}\Big),\\ &\mathbb{I}_{d_i} = d_i \times d_i \text{ identity matrix.}\end{aligned}$$

An important point to address is integration of nonlinear terms. We shall opt for a quadrature method relying on the basis $\{\phi_{ik} : k = 1, 2,\ \ldots\ , d_i\}$, which is assumed to satisfy the fundamental

property

$$\phi_{ij}(\mathbf{x}_k) = \delta_{jk}, \tag{3.63}$$

and plug the basis expansion of each unknown in (3.60), (3.61), $R(n_i, p_i)$. For example, considering the electron drift current, (3.63) implies that the r-th numerical component in the basis is calculated as

$$\Big[q\ \sum_{j=1}^{d_i} n_{ij}\phi_{ij}\ \ \mu_n\Big(\sum_{j=1}^{d_i}\overline{\mathbf{E}}_{ij}\phi_{ij}\Big)\ \ \sum_{j=1}^{d_i}\mathbf{E}_{ij}\ \phi_{ij}\Big]_{\mathbf{x}=\mathbf{x}_r} = q\ n_{ir}\ \ \mu_n\big(\overline{\mathbf{E}}_{ir}\big)\ \ \mathbf{E}_{ir} \qquad \forall r \in \{1, 2, \ldots d_i\}, \tag{3.64}$$

hence

$$\begin{aligned}
\int_{K_i} \mathbf{J}_{n_{i,drift}}\ \phi_{ik} &= \int_{K_i} q\ n_i\ \mu_n\big(\mathbf{E}_i\big)\ \mathbf{E}_i\ \phi_{ik} \approx \int_{K_i} q\ \sum_{j=1}^{d_i} n_{ij}\phi_{ij}\ \ \mu_n\Big(\sum_{j=1}^{d_i}\overline{\mathbf{E}}_{ij}\phi_{ij}\Big)\ \ \sum_{j=1}^{d_i}\mathbf{E}_{ij}\ \phi_{ij}\ \phi_{ik} \\
&\approx \int_{K_i} \sum_{r=1}^{d_i}\Big[q\ \sum_{j=1}^{d_i} n_{ij}\phi_{ij}\ \ \mu_n\Big(\sum_{j=1}^{d_i}\overline{\mathbf{E}}_{ij}\phi_{ij}\Big)\ \ \sum_{j=1}^{d_i}\mathbf{E}_{ij}\ \phi_{ij}\Big]_{\mathbf{x}=\mathbf{x}_r}\ \phi_{ir}\ \phi_{ik} \\
&= q\sum_{r=1}^{d_i}\ n_{ir}\ \mu_n\big(\overline{\mathbf{E}}_{ir}\big)\ \mathbf{E}_{ir}\int_{K_i}\phi_{ir}\phi_{ik},
\end{aligned} \tag{3.65}$$

and similarly for the other integrals. In essence, $\mathbf{J}_{n_{i,drift}}$ is replaced by its nodal interpolant. The choice is computationally attractive as no special matrix is introduced for the purpose of interpolation – the mass matrix is used.

The aforementioned developments lead to the following space discretization of Maxwell's equations:

$$\begin{aligned}
\overline{\mathbb{M}}_i^{\mu}\frac{\partial\overline{\mathbf{H}}_i}{\partial t} &= \overline{\mathbb{K}}_i \times \overline{\mathbf{E}}_i + \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}_{il}(\overline{\mathbf{E}}_{*,i}\times\hat{\mathbf{n}}_{il}) + \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}^*_{il}(\overline{\mathbf{E}}_{*,l}\times\hat{\mathbf{n}}_{il}), \\
\overline{\mathbb{M}}_i^{\epsilon}\frac{\partial\overline{\mathbf{E}}_i}{\partial t} &= -\overline{\mathbb{K}}_i \times \overline{\mathbf{H}}_i - \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}_{il}(\overline{\mathbf{H}}_{*,i}\times\hat{\mathbf{n}}_{il}) - \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}^*_{il}(\overline{\mathbf{H}}_{*,l}\times\hat{\mathbf{n}}_{il}) - \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{n_i} - \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{p_i}.
\end{aligned} \tag{3.66}$$

with the matrix-vector cross products

$$\begin{aligned}
\overline{\mathbb{K}}_i \times \overline{\mathbf{E}}_i &= \begin{bmatrix} \mathbb{K}_i^y\mathbf{E}_i^z - \mathbb{K}_i^z\mathbf{E}_i^y \\ -\mathbb{K}_i^x\mathbf{E}_i^z + \mathbb{K}_i^z\mathbf{E}_i^x \\ \mathbb{K}_i^x\mathbf{E}_i^y - \mathbb{K}_i^y\mathbf{E}_i^x \end{bmatrix}, \\
\overline{\mathbf{E}}_{*,i}\times\hat{\mathbf{n}}_{il} &= -\hat{\mathbf{n}}_{il}\times\overline{\mathbf{E}}_{*,i}\ .
\end{aligned} \tag{3.67}$$

Note that according to (3.62) the block vectors $\overline{\mathbf{E}}_{*,i}$ and $\overline{\mathbf{H}}_{*,i}$ contain the numerical coefficients of the i part of (3.33). Similarly, $\overline{\mathbf{E}}_{*,l}$ and $\overline{\mathbf{H}}_{*,l}$.

Moving to the semiconductor equations, we state

$$
\begin{aligned}
q\mathbb{M}_i\frac{\partial \mathbf{n}_i}{\partial t} &= -\overline{\mathbb{K}}_i\cdot\overline{\mathbf{J}}_{n_i} + \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}(\overline{\mathbf{J}}_{n*,i}\cdot\hat{\mathbf{n}}_{il}) + \sum_{l\in\mathcal{N}_i}\mathbb{S}^*_{il}(\overline{\mathbf{J}}_{n*,l}\cdot\hat{\mathbf{n}}_{il}) - q\mathbb{M}_i\mathbf{R}_i + q\mathbb{M}_i\mathbf{G}_i,\\
q\mathbb{M}_i\frac{\partial \mathbf{p}_i}{\partial t} &= \overline{\mathbb{K}}_i\cdot\overline{\mathbf{J}}_{p_i} - \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}(\overline{\mathbf{J}}_{p*,i}\cdot\hat{\mathbf{n}}_{il}) - \sum_{l\in\mathcal{N}_i}\mathbb{S}^*_{il}(\overline{\mathbf{J}}_{p*,l}\cdot\hat{\mathbf{n}}_{il}) - q\mathbb{M}_i\mathbf{R}_i + q\mathbb{M}_i\mathbf{G}_i,\\
\overline{\mathbb{M}}_i\overline{\boldsymbol{\alpha}}_i &= \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}_{il}\ (\mathbf{n}_{*,i}\ \hat{\mathbf{n}}_{il}) + \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}^*_{il}\ (\mathbf{n}_{*,l}\ \hat{\mathbf{n}}_{il}) - \overline{\mathbb{K}}_i\overline{\mathbf{n}}_i,\\
\overline{\mathbb{M}}_i\overline{\boldsymbol{\beta}}_i &= \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}_{il}\ (\mathbf{p}_{*,i}\ \hat{\mathbf{n}}_{il}) + \sum_{l\in\mathcal{N}_i}\overline{\mathbb{S}}^*_{il}\ (\mathbf{p}_{*,l}\ \hat{\mathbf{n}}_{il}) - \overline{\mathbb{K}}_i\overline{\mathbf{p}}_i,
\end{aligned}
\tag{3.68}
$$

with the dot products

$$
\begin{aligned}
\overline{\mathbb{K}}_i\cdot\overline{\mathbf{J}}_{n_i} &= \mathbb{K}_i^x\mathbf{J}^x_{n_i} + \mathbb{K}_i^y\mathbf{J}^y_{n_i} + \mathbb{K}_i^z\mathbf{J}^z_{n_i},\\
\overline{\mathbf{J}}_{n*,i}\cdot\hat{\mathbf{n}}_{il} &= (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{x}})\mathbf{J}^x_{n*,i} + (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{y}})\mathbf{J}^y_{n*,i} + (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{z}})\mathbf{J}^z_{n*,i},
\end{aligned}
\tag{3.69}
$$

and similarly for their p counterparts. The i and l terms in the numerical flux are

$$
\begin{aligned}
\mathbf{J}^v_{n*,i} &= \frac{1}{2}\mathbf{J}^v_{n_{i,drift}} - \frac{q}{2}\lambda_n\mathbf{n}_i\ (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{v}}) + \theta\ \mathbf{J}^v_{n_{i,diff}},\\
\mathbf{J}^v_{n*,l} &= \frac{1}{2}\mathbf{J}^v_{n_{l,drift}} + \frac{q}{2}\lambda_n\mathbf{n}_l\ (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{v}}) + (1-\theta)\ \mathbf{J}^v_{n_{l,diff}},\\
\mathbf{J}^v_{p*,i} &= \frac{1}{2}\mathbf{J}^v_{p_{i,drift}} + \frac{q}{2}\lambda_p\mathbf{p}_i\ (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{v}}) + \theta\ \mathbf{J}^v_{p_{i,diff}},\\
\mathbf{J}^v_{p*,l} &= \frac{1}{2}\mathbf{J}^v_{p_{l,drift}} - \frac{q}{2}\lambda_p\mathbf{p}_l\ (\hat{\mathbf{n}}_{il}\cdot\hat{\mathbf{v}}) + (1-\theta)\ \mathbf{J}^v_{p_{l,diff}},
\end{aligned}
\tag{3.70}
$$

with the electron diffusion part being "balanced" in the auxiliary equations by

$$
\begin{aligned}
\mathbf{n}_{*,i} &= (1-\theta)\ \mathbf{n}_i,\\
\mathbf{n}_{*,l} &= \theta\ \mathbf{n}_l,
\end{aligned}
\tag{3.71}
$$

and similarly for holes.

Inclusion of optical absorption (and dispersion) is straightforward – it suffices to add the polarization current to the Ampère-Maxwell law (the second equation of 3.66) the same way as the drift and diffusion ones:

$$
\overline{\mathbb{M}}^\epsilon_i\frac{\partial\overline{\mathbf{E}}_i}{\partial t} = \ \ldots\ - \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{n_i} - \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{p_i} - \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{l_i}.
\tag{3.72}
$$

Note however that here $\epsilon = \epsilon_0\epsilon_\infty$ inside $\overline{\mathbb{M}}^\epsilon_i$.
The auxiliary differential equations are discretized in space as

$$
\begin{aligned}
\overline{\mathbb{M}}_i\frac{\partial\overline{\mathbf{P}}_i}{\partial t} &= \overline{\mathbb{M}}_i\overline{\mathbf{J}}_{l,i},\\
\overline{\mathbb{M}}_i\frac{\partial\overline{\mathbf{J}}_{l,i}}{\partial t} &= \epsilon_0\omega_p^2\overline{\mathbb{M}}_i\overline{\mathbf{E}}_i - \gamma\overline{\mathbb{M}}_i\overline{\mathbf{J}}_{l,i} - \omega_r^2\overline{\mathbb{M}}_i\overline{\mathbf{P}}_i,
\end{aligned}
\tag{3.73}
$$

but they can actually be simplified by applying $\overline{\mathbb{M}}_i^{-1}$ to both sides.

In the end, the semi-discrete DG formulation on an element $K_i \subset \Omega$ reads as follows.

$$
\begin{aligned}
\overline{\mathbb{M}}_i^{\mu} \frac{\partial \overline{\mathbf{H}}_i}{\partial t} &= \overline{\mathbb{K}}_i \times \overline{\mathbf{E}}_i + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} (\overline{\mathbf{E}}_{*,i} \times \hat{\mathbf{n}}_{il}) + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} (\overline{\mathbf{E}}_{*,l} \times \hat{\mathbf{n}}_{il}), \\
\overline{\mathbb{M}}_i^{\epsilon} \frac{\partial \overline{\mathbf{E}}_i}{\partial t} &= -\overline{\mathbb{K}}_i \times \overline{\mathbf{H}}_i - \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} (\overline{\mathbf{H}}_{*,i} \times \hat{\mathbf{n}}_{il}) - \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} (\overline{\mathbf{H}}_{*,l} \times \hat{\mathbf{n}}_{il}) - \overline{\mathbb{M}}_i \overline{\mathbf{J}}_{n_i} - \overline{\mathbb{M}}_i \overline{\mathbf{J}}_{p_i} - \overline{\mathbb{M}}_i \overline{\mathbf{J}}_{l_i}, \\
q \mathbb{M}_i \frac{\partial \mathbf{n}_i}{\partial t} &= -\overline{\mathbb{K}}_i \cdot \overline{\mathbf{J}}_{n_i} + \sum_{l \in \mathcal{N}_i} \mathbb{S}_{il} (\overline{\mathbf{J}}_{n_*,i} \cdot \hat{\mathbf{n}}_{il}) + \sum_{l \in \mathcal{N}_l} \mathbb{S}^*_{il} (\overline{\mathbf{J}}_{n_*,l} \cdot \hat{\mathbf{n}}_{il}) - q \mathbb{M}_i \mathbf{R}_i + q \mathbb{M}_i \mathbf{G}_i, \\
q \mathbb{M}_i \frac{\partial \mathbf{p}_i}{\partial t} &= \overline{\mathbb{K}}_i \cdot \overline{\mathbf{J}}_{p_i} - \sum_{l \in \mathcal{N}_i} \mathbb{S}_{il} (\overline{\mathbf{J}}_{p_*,i} \cdot \hat{\mathbf{n}}_{il}) - \sum_{l \in \mathcal{N}_l} \mathbb{S}^*_{il} (\overline{\mathbf{J}}_{p_*,l} \cdot \hat{\mathbf{n}}_{il}) - q \mathbb{M}_i \mathbf{R}_i + q \mathbb{M}_i \mathbf{G}_i, \\
\overline{\mathbb{M}}_i \overline{\boldsymbol{\alpha}}_i &= \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} \ (\mathbf{n}_{*,i} \ \hat{\mathbf{n}}_{il}) + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} \ (\mathbf{n}_{*,l} \ \hat{\mathbf{n}}_{il}) - \overline{\mathbb{K}}_i \overline{\mathbf{n}}_i, \\
\overline{\mathbb{M}}_i \overline{\boldsymbol{\beta}}_i &= \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} \ (\mathbf{p}_{*,i} \ \hat{\mathbf{n}}_{il}) + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} \ (\mathbf{p}_{*,l} \ \hat{\mathbf{n}}_{il}) - \overline{\mathbb{K}}_i \overline{\mathbf{p}}_i, \\
\frac{\partial \overline{\mathbf{P}}_i}{\partial t} &= \overline{\mathbf{J}}_{l,i}, \\
\frac{\partial \overline{\mathbf{J}}_{l,i}}{\partial t} &= \epsilon_0 \omega_p^2 \overline{\mathbf{E}}_i - \gamma \overline{\mathbf{J}}_{l,i} - \omega_r^2 \overline{\mathbf{P}}_i.
\end{aligned}
\tag{3.74}
$$

For the sake of completeness, the case $K_i \subset \Omega_M$ is reported as well:

$$
\begin{aligned}
\overline{\mathbb{M}}_i^{\mu} \frac{\partial \overline{\mathbf{H}}_i}{\partial t} &= \overline{\mathbb{K}}_i \times \overline{\mathbf{E}}_i + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} (\overline{\mathbf{E}}_{*,i} \times \hat{\mathbf{n}}_{il}) + \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} (\overline{\mathbf{E}}_{*,l} \times \hat{\mathbf{n}}_{il}), \\
\overline{\mathbb{M}}_i^{\epsilon} \frac{\partial \overline{\mathbf{E}}_i}{\partial t} &= -\overline{\mathbb{K}}_i \times \overline{\mathbf{H}}_i - \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}_{il} (\overline{\mathbf{H}}_{*,i} \times \hat{\mathbf{n}}_{il}) - \sum_{l \in \mathcal{N}_i} \overline{\mathbb{S}}^*_{il} (\overline{\mathbf{H}}_{*,l} \times \hat{\mathbf{n}}_{il}) - \overline{\mathbb{M}}_i \overline{\mathbf{J}}_{d_i}, \\
\frac{\partial \overline{\mathbf{J}}_{d,i}}{\partial t} &= \epsilon_0 \omega_d^2 \overline{\mathbf{E}}_i - \gamma_d \overline{\mathbf{J}}_{d,i},
\end{aligned}
\tag{3.75}
$$

with $\epsilon = \epsilon_0 \epsilon_\infty$ inside $\overline{\mathbb{M}}_i^{\epsilon}$.[9]
The case $K_i \subset \Omega_e \setminus (\Omega_M \cup \Omega)$ is omitted as it is simply (3.75) with $\overline{\mathbf{J}}_{d,i} = \mathbf{0}$ and $\epsilon = \epsilon_0 \epsilon_r$.

3.1.3 Semi-discrete DG formulation: global form

As mentioned in Section 3.1, a notable approach to DG approximation of elliptic problems – which has been followed in the present work – consists in introducing an auxiliary variable for the gradient of the unknown and solving the resulting additional equation. The main reference on the topic is [ABCM02], in which a so-called *primal formulation* is derived – a global DG formulation condensed in a single equation.
The purpose of this section is to apply such a procedure to drift-diffusion equations, and then

[9] It may be worthwhile to recall that the value of ϵ_∞ depends on the material, thus it is expected not to be the same in Ω and Ω_M.

assemble a global DG formulation of the MDD system, in order to pave the way for future stability analysis.
With reference to the mathematical framework introduced in Section 3.1 we consider for example the equations governing n along the following steps.

Step 1 – Summation of the element-wise weak formulation over the whole mesh.
This yields, for all $\boldsymbol{\tau} \in \Sigma_h$, $v \in V_h$,

$$\begin{aligned}
\int_\Omega \boldsymbol{\alpha}_h \cdot \boldsymbol{\tau} &= -\int_\Omega n_h \nabla_h \cdot \boldsymbol{\tau} + \sum_{K\in\mathcal{T}_h} \int_{\partial K} n_* \, \hat{\mathbf{n}} \cdot \boldsymbol{\tau}, \\
\int_\Omega \frac{\partial n_h}{\partial t} v &= \sum_{K\in\mathcal{T}_h} \int_{\partial K} \left(n\mu_n \mathbf{E}\right)_* \cdot \hat{\mathbf{n}} \; v + \sum_{K\in\mathcal{T}_h} \int_{\partial K} \left(D_n \boldsymbol{\alpha}\right)_* \cdot \hat{\mathbf{n}} \; v \\
&\quad - \int_\Omega n_h \mu_n \mathbf{E} \cdot \nabla_h v - \int_\Omega D_n \boldsymbol{\alpha}_h \cdot \nabla_h v + \int_\Omega (G - R) \; v,
\end{aligned} \tag{3.76}$$

where $\int_\Omega n_h \nabla_h \cdot \boldsymbol{\tau} = \sum_{K\in\mathcal{T}_h} \int_K n_h \nabla \cdot \boldsymbol{\tau}$, and similarly for the other integrals on Ω involving ∇_h.

Step 2 – Rewriting the boundary integrals in terms of jump and average operators.
The step exploits the identity

$$\sum_{K\in\mathcal{T}_h} \int_{\partial K} q_K \boldsymbol{\varphi}_K \cdot \hat{\mathbf{n}} = \int_\Gamma [\![q]\!] \cdot \{\boldsymbol{\varphi}\} + \int_{\Gamma^0} \{q\} [\![\boldsymbol{\varphi}]\!], \qquad \forall q \in T(\Gamma), \boldsymbol{\varphi} \in [T(\Gamma)]^2, \tag{3.77}$$

where $q_K \boldsymbol{\varphi}_K$ is the value of $q\boldsymbol{\varphi}$ on the interior page of ∂K and $\hat{\mathbf{n}}$ the outward unit normal vector to ∂K.

By applying (3.77), we obtain for all $\boldsymbol{\tau} \in \Sigma_h$, $v \in V_h$

$$\begin{aligned}
\int_\Omega \boldsymbol{\alpha}_h \cdot \boldsymbol{\tau} &= -\int_\Omega n_h \nabla_h \cdot \boldsymbol{\tau} + \int_\Gamma [\![n_*]\!] \cdot \{\boldsymbol{\tau}\} + \int_{\Gamma^0} \{n_*\} [\![\boldsymbol{\tau}]\!], \\
\int_\Omega \frac{\partial n_h}{\partial t} v &= \int_\Gamma [\![v]\!] \cdot \{\left(n\mu_n \mathbf{E}\right)_*\} + \int_{\Gamma^0} \{v\} [\![\left(n\mu_n \mathbf{E}\right)_*]\!] \\
&\quad + \int_\Gamma [\![v]\!] \cdot \{\left(D_n \boldsymbol{\alpha}\right)_*\} + \int_{\Gamma^0} \{v\} [\![\left(D_n \boldsymbol{\alpha}\right)_*]\!] \\
&\quad - \int_\Omega n_h \mu_n \mathbf{E} \cdot \nabla_h v - \int_\Omega D_n \boldsymbol{\alpha}_h \cdot \nabla_h v + \int_\Omega (G - R) \; v.
\end{aligned} \tag{3.78}$$

Step 3 – Expressing $\boldsymbol{\alpha}_h$ in terms of the sole n_h.
Another important identity, that holds for all $\boldsymbol{\tau} \in [H^1(\mathcal{T}_h)]^3$, $v \in H^1(\mathcal{T}_h)$, is

$$-\int_\Omega v \nabla_h \cdot \boldsymbol{\tau} = \int_\Omega \boldsymbol{\tau} \cdot \nabla_h v - \int_\Gamma \{\boldsymbol{\tau}\} \cdot [\![v]\!] - \int_{\Gamma^0} [\![\boldsymbol{\tau}]\!] \{v\}. \tag{3.79}$$

Taking $v = n_h$ in it allows to write

$$\begin{aligned}
\int_\Omega \boldsymbol{\alpha}_h \cdot \boldsymbol{\tau} &= \int_\Omega \boldsymbol{\tau} \cdot \nabla_h n_h - \int_\Gamma \{\boldsymbol{\tau}\} \cdot [\![n_h]\!] - \int_{\Gamma^0} [\![\boldsymbol{\tau}]\!] \{n_h\} + \int_\Gamma [\![n_*]\!] \cdot \{\boldsymbol{\tau}\} + \int_{\Gamma^0} \{n_*\} [\![\boldsymbol{\tau}]\!] \\
&= \int_\Omega \boldsymbol{\tau} \cdot \nabla_h n_h + \int_\Gamma [\![n_* - n_h]\!] \cdot \{\boldsymbol{\tau}\} + \int_{\Gamma_0} \{n_* - n_h\} [\![\boldsymbol{\tau}]\!].
\end{aligned} \tag{3.80}$$

If we introduce the lifting operators $\mathbf{r} : [L^2(\Gamma)]^3 \to \Sigma_h$ and $\mathbf{l} : L^2(\Gamma^0) \to \Sigma_h$ such that for all $\boldsymbol{\tau} \in \Sigma_h$

$$\int_\Omega \mathbf{r}(\boldsymbol{\varphi}) \cdot \boldsymbol{\tau} = -\int_\Gamma \boldsymbol{\varphi} \cdot \{\boldsymbol{\tau}\}, \qquad \int_\Omega \mathbf{l}(q) \cdot \boldsymbol{\tau} = -\int_{\Gamma_0} q[\![\boldsymbol{\tau}]\!], \tag{3.81}$$

whose existence is proved in [ABCM02], and recall that $\nabla_h n_h \in \nabla_h V_h \subset \Sigma_h$, we can state the weak identity in Σ_h

$$\boldsymbol{\alpha}_h = \boldsymbol{\alpha}_h(n_h) := \nabla_h n_h - \mathbf{r}([\![n_* - n_h]\!]) - \mathbf{l}(\{n_* - n_h\}). \tag{3.82}$$

Step 4 – Substituting (3.82) in the continuity equation.
The volume integral featuring the diffusion current is preliminarily rewritten as:

$$\begin{aligned}\int_\Omega D_n \boldsymbol{\alpha}_h \cdot \nabla_h v &= \int_\Omega D_n \nabla_h n_h \cdot \nabla_h v - \int_\Omega D_n \mathbf{r}([\![n_* - n_h]\!]) \cdot \nabla_h v - \int_\Omega D_n \mathbf{l}(\{n_* - n_h\}) \cdot \nabla_h v \\ &= \int_\Omega D_n \nabla_h n_h \cdot \nabla_h v + \int_\Gamma [\![n_* - n_h]\!] \cdot \{D_n \ \nabla_h v\} + \int_{\Gamma_0} \{n_* - n_h\} [\![D_n \ \nabla_h v]\!].\end{aligned} \tag{3.83}$$

Then, we plug the new expression in the continuity equation and obtain

$$\begin{aligned}\int_\Omega \frac{\partial n_h}{\partial t} v = &\int_\Gamma [\![v]\!] \cdot \{\left(n\mu_n \mathbf{E}\right)_*\} + \int_{\Gamma^0} \{v\} [\![\left(n\mu_n \mathbf{E}\right)_*]\!] + \int_\Gamma [\![v]\!] \cdot \{\left(D_n \boldsymbol{\alpha}\right)_*\} + \int_{\Gamma^0} \{v\} [\![\left(D_n \boldsymbol{\alpha}\right)_*]\!] \\ &- \int_\Omega n_h \mu_n \mathbf{E} \cdot \nabla_h v - \int_\Omega D_n \nabla_h n_h \cdot \nabla_h v - \int_\Gamma [\![n_* - n_h]\!] \cdot \{D_n \ \nabla_h v\} \\ &- \int_{\Gamma_0} \{n_* - n_h\} [\![D_n \ \nabla_h v]\!] + \int_\Omega (G - R) \ v.\end{aligned} \tag{3.84}$$

Grouping integrals and bringing all terms depending on n to the left-hand side leads to the primal formulation

$$\begin{aligned}\int_\Omega \frac{\partial n_h}{\partial t} v &- \int_\Omega \Big(- n_h \mu_n \mathbf{E} \cdot \nabla_h v - D_n \nabla_h n_h \cdot \nabla_h v \Big) \\ &- \int_\Gamma \Big([\![v]\!] \cdot \{\left(n\mu_n \mathbf{E}\right)_*\} + [\![v]\!] \cdot \{\left(D_n \boldsymbol{\alpha}\right)_*\} - [\![n_* - n_h]\!] \cdot \{D_n \ \nabla_h v\} \Big) \\ &- \int_{\Gamma^0} \Big(\{v\} [\![\left(n\mu_n \mathbf{E}\right)_*]\!] + \{v\} [\![\left(D_n \boldsymbol{\alpha}\right)_*]\!] - \{n_* - n_h\} [\![D_n \ \nabla_h v]\!] \Big) \\ &+ \int_\Omega R \ v = \int_\Omega G \ v.\end{aligned} \tag{3.85}$$

To complete the picture we conceive an analogous equation to (3.85) for holes and assemble the global formulation of the MDD system.

$$
\begin{aligned}
\int_\Omega \mu \frac{\partial \mathbf{H}_h}{\partial t} \cdot \boldsymbol{\tau} &= -\int_\Omega \mathbf{E}_h \cdot \nabla \times \boldsymbol{\tau} + \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathbf{E}_* \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau}, \\
\int_\Omega \epsilon \frac{\partial \mathbf{E}_h}{\partial t} \cdot \boldsymbol{\tau} &= \int_\Omega \mathbf{H}_h \cdot \nabla \times \boldsymbol{\tau} - \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathbf{H}_* \times \hat{\mathbf{n}} \cdot \boldsymbol{\tau} - \int_\Omega (q\mu_n n_h + q\mu_p p_h)\mathbf{E}_h \cdot \boldsymbol{\tau} \\
&\quad - \int_\Omega q D_n \nabla_h n_h \cdot \boldsymbol{\tau} - \int_\Gamma q[\![n_* - n_h]\!] \cdot \{D_n\ \boldsymbol{\tau}\} - \int_{\Gamma_0} q\{n_* - n_h\}[\![D_n\ \boldsymbol{\tau}]\!] \\
&\quad + \int_\Omega q D_p \nabla_h p_h \cdot \boldsymbol{\tau} + \int_\Gamma q[\![p_* - p_h]\!] \cdot \{D_p\ \boldsymbol{\tau}\} + \int_{\Gamma_0} q\{p_* - p_h\}[\![D_p\ \boldsymbol{\tau}]\!], \\
\int_\Omega \frac{\partial n_h}{\partial t} v &- \int_\Omega \Big(- n_h \mu_n \mathbf{E} \cdot \nabla_h v - D_n \nabla_h n_h \cdot \nabla_h v \Big) \\
&- \int_\Gamma \Big([\![v]\!] \cdot \{ \big(n\mu_n \mathbf{E}\big)_* \} + [\![v]\!] \cdot \{ \big(D_n \boldsymbol{\alpha}\big)_* \} - [\![n_* - n_h]\!] \cdot \{D_n\ \nabla_h v\} \Big) \\
&- \int_{\Gamma^0} \Big(\{v\} [\![\big(n\mu_n \mathbf{E}\big)_*]\!] + \{v\} [\![\big(D_n \boldsymbol{\alpha}\big)_*]\!] - \{n_* - n_h\} [\![D_n\ \nabla_h v]\!] \Big) \\
&+ \int_\Omega R\ v = \int_\Omega G\ v, \\
\int_\Omega \frac{\partial p_h}{\partial t} v &- \int_\Omega \Big(+ p_h \mu_p \mathbf{E} \cdot \nabla_h v - D_p \nabla_h p_h \cdot \nabla_h v \Big) \\
&- \int_\Gamma \Big(- [\![v]\!] \cdot \{ \big(p\mu_p \mathbf{E}\big)_* \} + [\![v]\!] \cdot \{ \big(D_p \boldsymbol{\beta}\big)_* \} - [\![p_* - p_h]\!] \cdot \{D_p\ \nabla_h v\} \Big) \\
&- \int_{\Gamma^0} \Big(- \{v\} [\![\big(p\mu_p \mathbf{E}\big)_*]\!] + \{v\} [\![\big(D_p \boldsymbol{\beta}\big)_*]\!] - \{p_* - p_h\} [\![D_p\ \nabla_h v]\!] \Big) \\
&+ \int_\Omega R\ v = \int_\Omega G\ v.
\end{aligned}
\tag{3.86}
$$

Note that Maxwell's equations feature diffusion currents in terms of n_* and p_*, thanks to the elimination of the auxiliary unknowns α and β.
Further exploring this path would require to transform the integrals involving $\mathbf{E}_*$ and $\mathbf{H}_*$ into suitable counterparts defined on the skeleton of the mesh, possibly by means of the average and jump operators (3.77).

This first global formulation could serve as a basic framework for a possible theoretical stability study that would complement empirical stability analyses such as the one proposed in Section 4.2. This deserves a consequent theoretical numerical analysis work that will not be addressed in this work.

3.1.4 Discrete numerical formulation

With the semi-discrete equations in place it is finally possible to address time discretization. The presence of nonlinear terms and the aim of exploiting parallel computing suggest to look at the explicit DGTD methods, as these allow local problems to be solved independently (i.e. in parallel) at each iteration. The notable price to pay is an upper bound on the time step.

The choice falls on explicit Runge-Kutta schemes, which offer higher-order convergence than forward Euler integration, while still having a one-step structure.
The core problem (3.74) is discussed here – the simpler (3.75) can be addressed similarly.

3.1.4.1 Runge-Kutta

For an ordinary differential equation of the kind

$$\begin{cases} U'(t) &= G(t, U(t)), \\ U(t_0) &= U^0, \end{cases} \tag{3.87}$$

the (explicit) Runge-Kutta scheme of order $M \in \mathbb{N}_0$, that takes U^m as an approximation of $U(t_m)$ and calculates U^{m+1}, is defined by the M steps

$$\begin{aligned} \psi_1 &= G(t_m, U^m), \\ \psi_j &= G(\ t_m + \Delta t\ c_j\ ,\ U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\ \psi_s\ ), \qquad j = 2, 3,\ \dots, M, \\ U^{m+1} &= U^m + \Delta t \sum_{j=1}^{M} b_j\ \psi_j, \end{aligned} \tag{3.88}$$

the sets of coefficients $\{a_{js}\}, \{b_j\}, \{c_j\}$ being defined on a per-order basis according to a so-called *Butcher table*.[10]

In fact, (3.74) cannot be represented by (3.87) as it features auxiliary equations with no time derivatives. The right model is rather

$$\begin{cases} U'(t) &= G(t, U(t), W(t)), \\ W(t) &= F(U(t)), \\ U(t_0) &= U^0, \end{cases} \tag{3.89}$$

or equivalently

$$\begin{cases} U'(t) &= G(t, U(t), F(U(t))), \\ U(t_0) &= U^0, \end{cases} \tag{3.90}$$

for which (3.88) becomes

$$\begin{aligned} \psi_1 &= G(t_m, U^m, F(U^m)), \\ \psi_j &= G(\ t_m + \Delta t\ c_j\ ,\ U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\ \psi_s\ ,\ F(U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\ \psi_s)\ ), \qquad j = 2, 3,\ \dots, M, \\ U^{m+1} &= U^m + \Delta t \sum_{j=1}^{M} b_j\ \psi_j. \end{aligned} \tag{3.91}$$

The cases $M = 2, 4$ have been implemented in the present work and are detailed below; their convergence rates for time discretization are 2 and 4 respectively [Qua13].

[10] Discussing the nature of (3.88) is beyond the scope of the present work; the interested reader can refer to [But87].

- Runge-Kutta 2:

$$\begin{aligned} \psi_1 &= G(t_m, U^m, F(U^m)), \\ \psi_2 &= G(\ t_m + \Delta t\ ,\ U^m + \Delta t\ \psi_1\ ,\ F(U^m + \Delta t\ \psi_1)\ ), \\ U^{m+1} &= U^m + \frac{\Delta t}{2}(\psi_1 + \psi_2). \end{aligned} \tag{3.92}$$

- Runge-Kutta 4:

$$\begin{aligned} \psi_1 &= G(t_m, U^m, F(U^m)), \\ \psi_2 &= G(\ t_m + \frac{\Delta t}{2},\ U^m + \frac{\Delta t}{2}\psi_1,\ F(U^m + \frac{\Delta t}{2}\psi_1)\ ), \\ \psi_3 &= G(\ t_m + \frac{\Delta t}{2},\ U^m + \frac{\Delta t}{2}\psi_2,\ F(U^m + \frac{\Delta t}{2}\psi_2)\ ), \\ \psi_4 &= G(\ t_m + \Delta t,\ U^m + \Delta t\ \psi_3,\ F(U^m + \Delta t\ \psi_3)\ ), \\ U^{m+1} &= U^m + \frac{\Delta t}{6}(\psi_1 + 2\psi_2 + 2\psi_3 + \psi_4). \end{aligned} \tag{3.93}$$

More generally, as further discussed in [Qua13] and proven in [But87]:

- the convergence rate of an explicit RKM method cannot exceed M;
- there exists no explicit RKM method converging at rate M if $M \geq 5$.

Hence, convergence at rate M requires:

- at least M stages for $M = 1, 2, 3, 4$;
- more than M stages for $M \geq 5$.

To apply explicit RKM algorithms in our context we recall that the local DG formulation on K_i involves the unknowns on both K_i and its direct neighbors $\{K_l : l \in \mathcal{N}_i\}$; as a consequence, each of the M steps must be applied in parallel to all mesh elements, $K_1, K_2,\ ... K_{N_\mathcal{T}}$. To better highlight this point, (3.74) is synthesized as follows.
Let us first introduce some auxiliary quantities for the local unknowns and for the right-hand sides:

$$U_i = \begin{bmatrix} \overline{\mathbf{H}}_i & \overline{\mathbf{E}}_i & \mathbf{n}_i & \mathbf{p}_i & \overline{\mathbf{P}}_i & \overline{\mathbf{J}}_{l,i} \end{bmatrix}^T,$$

$$G_i = \begin{bmatrix} G_{i,H} \\ G_{i,E} \\ G_{i,n} \\ G_{i,p} \\ G_{i,l} \end{bmatrix} = \begin{bmatrix} \overline{\mathbb{M}}_i^{\mu,-1} \left[\text{R.H.S. of the 1st equation of (3.74)} \right] \\ \overline{\mathbb{M}}_i^{\epsilon,-1} \left[\text{R.H.S. of the 2nd equation of (3.74)} \right] \\ q^{-1}\mathbb{M}_i^{-1} \left[\text{R.H.S. of the 3rd equation of (3.74)} \right] \\ q^{-1}\mathbb{M}_i^{-1} \left[\text{R.H.S. of the 4th equation of (3.74)} \right] \\ \left[\text{R.H.S. of the Lorentz model} \right] \end{bmatrix}, \tag{3.94}$$

$$F_i = \begin{bmatrix} \overline{\mathbb{M}}_i^{-1} \left[\text{R.H.S. of the 5th equation of (3.74)} \right] \\ \overline{\mathbb{M}}_i^{-1} \left[\text{R.H.S. of the 6th equation of (3.74)} \right] \end{bmatrix}.$$

These are assembled into the global vectors

$$U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_{N_\mathcal{T}} \end{bmatrix}, \qquad G = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_{N_\mathcal{T}} \end{bmatrix}, \qquad F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_{N_\mathcal{T}} \end{bmatrix}. \tag{3.95}$$

With such notation, the expression $U'(t) = G(t, U(t), F(U(t)))$ describes a global problem built by collecting all of the $\mathcal{N}_\mathcal{T}$ semi-discrete local problems like (3.74); to this global problem we shall apply the Runge-Kutta algorithms (3.91).
In practice, assuming that the initial state includes the auxiliary unknowns $\boldsymbol{\alpha}, \boldsymbol{\beta}$, the implementation consists in the steps below.

1. Calculate ψ_1 as an array made of the following 4 blocks (each calculated by a dedicated subroutine).
   (a) $\{G_{i,H}(t_m, U^m) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
   (b) $\{G_{i,E}(t_m, U^m, F(U^m)) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
   (c) $\{G_{i,n}(t_m, U^m, F(U^m)) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
   (d) $\{G_{i,p}(t_m, U^m, F(U^m)) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
   (e) $\{G_{i,l}(t_m, U^m) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
2. Recall such subroutines to calculate ψ_j, $j = 2, 3, \ \dots\ , M$, with intermediate inputs.
   (a) $\{G_{i,H}(t_m + c_j\Delta t, U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$
   (b) $\{G_{i,E}(t_m + c_j\Delta t, U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s, F(U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s)) : i = 1, 2, \ \dots\ , N_\mathcal{T}\}$

(c) $\{G_{i,n}(t_m + c_j\Delta t, U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s, F(U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s)) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

(d) $\{G_{i,p}(t_m + c_j\Delta t, U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s, F(U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s)) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

(e) $\{G_{i,l}(t_m + c_j\Delta t, U^m + \Delta t \sum_{s=1}^{j-1} a_{js}\psi_s) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

3. Calculate U^{m+1} from U^m and $\psi_1, \psi_2, \ ... \ , \psi_M$.

4. Calculate the auxiliary unknowns, that is $F(U^{m+1})$, based on the knowledge of U^{m+1}.

3.1.4.2 Low-Storage Runge-Kutta

When mesh elements and degrees of freedom increase, memory consumption can become a severe issue, as (3.91) requires the storage of M arrays of the dimension of U at each time step. This has led to the conception of methods known as Low-Storage Runge-Kutta (LSRK) [Wil80] and based on the algorithm

$$\left.\begin{aligned} \psi_1 &= U^m, \\ \psi_2 &= A_j\psi_2 + \Delta t\ G(\ t_m + \Delta t\ C_j\ ,\ \psi_1\ ,\ F(\psi_1)\ ), \\ \psi_1 &= \psi_1 + B_j\psi_2, \\ U^{m+1} &= \psi_1. \end{aligned}\right\} \quad j = 1, 2, \ ... \ , M, \tag{3.96}$$

The explicit expressions of the new coefficients are omitted – they are derived from those appearing in (3.91) and can be found in the original paper. The focus here is rather on the fact that the number of stored arrays per time step is reduced to 2, regardless of M, since ψ_1 and ψ_2 are overwritten at each iteration j. The price to pay is a generally lower convergence rate; for example, it has been proved that RKM, $M = 1, 2, 3, 4$, is accurate to order M whereas LSRK requires $M = 5$ to achieve order 4 [But87], [Wil80].

The application of (3.96) to our system goes as follow.

1. Calculate ψ_2 as an array made of the following 4 blocks (each calculated by a dedicated subroutine).

   (a) $\{G_{i,H}(t_m, \psi_1) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

   (b) $\{G_{i,E}(t_m, \psi_1, F(\psi_1)) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

   (c) $\{G_{i,n}(t_m, \psi_1, F(\psi_1)) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

   (d) $\{G_{i,p}(t_m, \psi_1, F(\psi_1)) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

   (e) $\{G_{i,l}(t_m, \psi_1) : i = 1, 2, \ ... \ , N_{\mathcal{T}}\}$

2. Calculate ψ_1 based on its current value and ψ_2.

3. Repeat the steps until $j = M$.

4. Store the new ψ_1 as U^{m+1}.

5. Calculate the auxiliary unknowns, that is $F(U^{m+1})$, based on the knowledge of U^{m+1}.

3.1.4.3 Discussion on stability

In the case of explicit integration schemes, drift and diffusion concur in limiting the time step. To better understand this, it is convenient to consider the two limiting cases of pure advection and pure diffusion.

Neglecting diffusion leads to a classical conservation law (as pointed out in Section 3.1.1.1) which demands to satisfy the notable condition by Courant-Friedrichs-Lewy

$$\Delta t \leq C_p^a \frac{h}{v}. \tag{3.97}$$

The coefficient C_p^a depends on the interpolation degree p and is generally determined empirically; v is the maximum advection velocity in the mesh, to be drawn from the characteristic speeds (3.17). More specifically

$$v = \max\left\{ \max_{\mathcal{T}_h} \frac{1}{\sqrt{\epsilon\mu}},\ \max_{\mathcal{T}_h}(\mu_n||\mathbf{E}||),\ \max_{\mathcal{T}_h}(\mu_p||\mathbf{E}||) \right\}. \tag{3.98}$$

As pointed out in Section 2.1.1, increasing the electric field can increase drift velocity up to a saturation value, which is smaller than the speed of light. Thus, $v = \max_{\mathcal{T}_h}\{\frac{1}{\sqrt{\epsilon\mu}}\}$ and stability is determined by the electromagnetic part of the Maxwell-Drift system.[11]

Neglecting drift leads to a coupled pair of diffusion-reaction equations (one for n, one for p). As for the case of the heat equation, the stability condition is [HW02]

$$\Delta t \leq C_p^d \frac{h^2}{D}, \tag{3.99}$$

where

$$D = \max\{\max_{\mathcal{T}_h} D_n, \max_{\mathcal{T}_h} D_p\} \tag{3.100}$$

and C_p^d is a coefficient depending on the interpolation degree p and, like C_p^a, is generally determined empirically. From a practical standpoint, $D_n = V_T\mu_n$, $D_p = V_T\mu_p$ are calculated at the beginning of a simulation, then drift is disabled by setting $\mu_n = \mu_p = 0$.[12]

Going back to the original system, one could be tempted to assemble the two constraints as

$$\Delta t \leq \min\left\{ C_p^a \frac{h}{v},\ C_p^d \frac{h^2}{D} \right\}, \tag{3.101}$$

but it was pointed out in [AD06] that the approach may lead to erroneous results in problems with mixed boundary conditions and should be therefore waived in favor of a continuous interpolation between the two constraints. The literature offers references on time step selection in linear advection-diffusion with uniform velocity and diffusivity: in a multidimensional DG framework (e.g. [AD06]); in a one-dimensional finite-volume framework (e.g. [FLLO89]). Based on these, a sensible constraint on the time step of the MDD system should be

$$\Delta t \leq \frac{h}{\dfrac{v}{C_p^a} + \dfrac{D}{C_p^d h}}, \tag{3.102}$$

[11] For example, $v_{n,sat} = 0.72 \times 10^5$ m s^{-1} in GaAs at 300K [QMPS00].

[12] Note that physics impose $\mu_n > \mu_p$ and hence $D = \max_{\mathcal{T}_h} D_n$.

with which becomes (3.97) for $D = 0$ and (3.99) for $v = 0$.

It is worthwhile to remark the benefit of assuming static μ_n and μ_p (hence D_n and D_p). If it were not the case, the right-hand sides of the stability conditions just presented would be time-varying, thereby potentially introducing the need of adjusting the time step at each iteration.

3.1.4.4 Optimized time discretization

When physical values are assigned to simulation parameters, the speed of light outnumbers drift velocity and actually dominates (3.102), even for a mesh size as low as a few nanometers (as it may be required in applications where optical radiation interacts with a nanoscaled device). In other words, electromagnetic field propagation sets a stricter constraint on the time step than drift and diffusion, suggesting to implement two different time grids: one for Maxwell's equations ($\Delta t = \Delta t_{EM}$), one for the continuity equations ($\Delta t = \Delta t_{np}$). The general idea, proposed in [CB20a] (to which we refer for the detailed algorithm), is to let $\Delta t_{np} = N_{skip}\ \Delta t_{EM}$ for a suitable integer N_{skip} defining the number of electromagnetic iterations to be skipped before the next update of the n and p. This idea has been tested in the present work, namely in the test case shown in Section 4.4.3, whereas its generalization to real device simulation appears less trivial, as will be discussed in Section 6.2.

.

Chapter 4

Implementation

The foundations for the numerical resolution of the MDD model in a three-dimensional DGTD framework were set in the foregoing sections. Over the course of the project, different avenues have been pondered as to how to translate such a work into software. Integrating a MDD module into DIOGENeS [Atl], the in-house DGTD solver of Maxwell's equations in three dimensions, was the most fascinating option in the first place. However, it was estimated that this would entail devoting significant effort to integration itself, at the expense of a key requirement: agile exploration of the possible strategies for accomplishing the core task of numerically solving the MDD model. Based on such a requirement, reducing the number of space dimensions has appeared to be a sensible way forward, considering also that a second, basic in-house DGTD code solving two-dimensional (TM) Maxwell's equations, was available.
Working on the latter solver has actually demanded to implement from scratch numerous key features that were already available in DIOGENeS, such as: the upwind flux (3.33), RK/LSRK schemes, a Total Field / Scattered Field (TFSF) apparatus (introduced in Section 5.1), Fourier analysis, and reading an arbitrary unstructured mesh.

In the end, handling the latter task alongside with the core one has been challenging enough to confine the project to the two-dimensional environment. This might seem a drawback at first glance, but overall the experience has been quite valuable from an educational standpoint and has indeed served to the original purpose of setting the guidelines for a future three-dimensional implementation.

This chapter is organized as follows. The first section introduces the numerical formulation of the system in two dimensions. The second section presents an itinerary of test cases, from the elementary building blocks (Maxwell's equations, advection, diffusion) to the final MDD equations. Over the course of the project, each test has entailed code verification, e.g. by means of convergence analysis; the three most relevant and representative cases have been selected, and the respective simulation results are proposed in the third section along with the criteria behind their designation.[1]

[1] A choice has been made to illustrate the whole repertoire of test cases in Section 4.2, rather than restrict the attention to those selected in Section 4.4, the rationale being to set a reference for an implementation in three dimensions. Indeed, each case can be easily generalized to the latter framework.

4.1 Two-dimensional formulation

An important point to discuss when considering dimension reduction is the choice of a TE model versus a TM one. From a mathematical standpoint, the TE model is preferable because it has been the object of studies on existence and uniqueness (e.g. [Joc98]) while no references on the TM counterpart have been found in the literature. From a physical standpoint, having only one component of the electric field would limit charge transport to one direction and spatial variable, thereby making it impossible to simulate any geometry of practical interest. For example, it is apparent that a photo-conductive antenna requires (at least) two-dimensional charge and electric field distributions.

Having established the importance of the TE model, we proceed to its derivation and numerical resolution. This section presents the DGTD formulation in two-dimensions, while Section 4.2 is devoted to verifying the corresponding solver.

The two-dimensional TE version of (2.1) is derived by setting $E^z = H^x = H^y = \alpha^z = \beta^z = 0$, $\hat{\mathbf{n}}_{il} = [n_{il}^x \; n_{il}^y \; 0]$.

$$
\begin{aligned}
-\mu \, \partial_t H_z &= \partial_x E_y - \partial_y E_x, \\
\epsilon \, \partial_t \begin{bmatrix} E_x \\ E_y \end{bmatrix} &= \begin{bmatrix} \partial_y H_z \\ -\partial_x H_z \end{bmatrix} - q\mu_n n \begin{bmatrix} E_x \\ E_y \end{bmatrix} - qD_n \boldsymbol{\alpha} \; - q\mu_p p \begin{bmatrix} E_x \\ E_y \end{bmatrix} + qD_p \boldsymbol{\beta}, \\
\partial_t n &= \nabla \cdot \left(\mu_n n \begin{bmatrix} E_x \\ E_y \end{bmatrix} + D_n \boldsymbol{\alpha} \right) + G - R, \\
\partial_t p &= -\nabla \cdot \left(\mu_p p \begin{bmatrix} E_x \\ E_y \end{bmatrix} - D_p \boldsymbol{\beta} \right) + G - R, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p, \\
&\text{in } \Omega \times [0, +\infty[,
\end{aligned}
\tag{4.1}
$$

where Ω is a two-dimensional set in the (x, y) plane like the one already depicted in Figure 2.1. As usual, the problem is mathematically closed with initial and boundary conditions:

$$
\begin{aligned}
\mathbf{E}(x,y,0) &= \mathbf{E}_0(x,y) && (x,y) \in \bar{\Omega}, \\
H_z(x,y,0) &= H_{z0}(x,y) && (x,y) \in \bar{\Omega}, \\
n(x,y,0) &= n_0(x,y) && (x,y) \in \bar{\Omega}, \\
p(x,y,0) &= p_0(x,y) && (x,y) \in \bar{\Omega}, \\
\\
n(x,y,t) &= n_D(x,y) && (x,y,t) \in \Gamma_D \times [0, +\infty[, \\
p(x,y,t) &= p_D(x,y) && (x,y,t) \in \Gamma_D \times [0, +\infty[, \\
n_x E_y(x,y,t) - n_y E_x(x,y,t) &= 0 && (x,y,t) \in \Gamma_D \times [0, +\infty[, \\
\\
\hat{\mathbf{n}} \cdot \left(n\mu_n \mathbf{E}(x,y,t) + D_n \boldsymbol{\alpha}(x,y,t) \right) &= 0 && (x,y,t) \in \Gamma_N \times [0, +\infty[, \\
\hat{\mathbf{n}} \cdot \left(p\mu_p \mathbf{E}(x,y,t) - D_p \boldsymbol{\beta}(x,y,t) \right) &= 0 && (x,y,t) \in \Gamma_N \times [0, +\infty[, \\
H_z(x,y,t) &= H_{zN}(x,y,t) && (x,y,t) \in \Gamma_N \times [0, +\infty[.
\end{aligned}
\tag{4.2}
$$

To obtain the semi-discrete DG formulation we can refer to (3.74) and again set $E^z = H^x = H^y = \alpha^z = \beta^z = 0$, $\hat{\mathbf{n}}_{il} = [n_{il}^x \; n_{il}^y \; 0]$. Then, we take $\hat{\mathbf{v}} = \hat{\mathbf{z}}$ in the first equation of (3.57) and $\hat{\mathbf{v}} = \hat{\mathbf{x}}, \hat{\mathbf{y}}$ in the second one, as well as in (3.59).
To ease the task we make use of some preliminary identities.

$$\begin{aligned} \mathbf{E} \cdot \hat{\mathbf{z}} \times \nabla\phi = \mathbf{E} \cdot \begin{bmatrix} -\partial_y\phi \\ \partial_x\phi \\ 0 \end{bmatrix} = E^y \partial_x\phi - E^x \partial_y\phi, \\ \mathbf{E} \times \hat{\mathbf{n}} \cdot \hat{\mathbf{z}} = E^x n^y - E^y n^x, \end{aligned} \tag{4.3}$$

$$\begin{aligned} -\begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \cdot \hat{\mathbf{x}} \times \nabla\phi = -\begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -\partial_z\phi \\ \partial_y\phi \end{bmatrix} = -H^z \partial_y\phi, \\ \begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \times \hat{\mathbf{n}} \cdot \hat{\mathbf{x}} = H^z \begin{bmatrix} -n^y \\ n^x \\ 0 \end{bmatrix} \cdot \hat{\mathbf{x}} = -H^z n^y, \end{aligned} \tag{4.4}$$

$$\begin{aligned} -\begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \cdot \hat{\mathbf{y}} \times \nabla\phi = -\begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \cdot \begin{bmatrix} \partial_z\phi \\ 0 \\ -\partial_x\phi \end{bmatrix} = H^z \partial_x\phi, \\ \begin{bmatrix} 0 \\ 0 \\ H^z \end{bmatrix} \times \hat{\mathbf{n}} \cdot \hat{\mathbf{y}} = H^z \begin{bmatrix} -n^y \\ n^x \\ 0 \end{bmatrix} \cdot \hat{\mathbf{y}} = H^z n^x. \end{aligned} \tag{4.5}$$

The semi-discrete DG formulation of TE Maxwell's equations reads

$$\begin{aligned} \sum_{j=1}^{d_i} \frac{\partial H_{ij}^z}{\partial t} \int_{K_i} \mu \; \phi_{ij}\phi_{ik} = &+ \sum_{j=1}^{d_i} E_{ij}^y \int_{K_i} \phi_{ij} \frac{\partial}{\partial x}\phi_{ik} - E_{ij}^x \int_{K_i} \phi_{ij} \frac{\partial}{\partial y}\phi_{ik} \\ &+ \sum_{l \in \mathcal{N}_i} \Bigg(\sum_{j=1}^{d_i} \frac{Y_i E_{ij}^x n_{il}^y - Y_i E_{ij}^y n_{il}^x - \alpha H_{ij}^z}{Y_i + Y_l} \int_{a_{il}} \phi_{ij}\phi_{ik} \\ &+ \sum_{g=1}^{d_l} \frac{Y_l E_{lg}^x n_{il}^y - Y_l E_{lg}^y n_{il}^x + \alpha H_{lg}^z}{Y_i + Y_l} \int_{a_{il}} \phi_{lg}\phi_{ik} \Bigg), \end{aligned} \tag{4.6}$$

$$\begin{aligned} \sum_{j=1}^{d_i} \frac{\partial E_{ij}^x}{\partial t} \int_{K_i} \epsilon \; \phi_{ij}\phi_{ik} = &- \sum_{j=1}^{d_i} H_{ij}^z \int_{K_i} \phi_{ij} \frac{\partial}{\partial y}\phi_{ik} \\ &+ \sum_{l \in \mathcal{N}_i} \Bigg(\sum_{j=1}^{d_i} \frac{-Z_i H_{ij}^z n_{il}^y + \alpha[E_{ij}^x - n_{il}^x(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{ij})]}{Z_i + Z_l} \int_{a_{il}} \phi_{ij}\phi_{ik} \\ &+ \sum_{g=1}^{d_l} \frac{-Z_l H_{lg}^z n_{il}^y - \alpha[E_{lg}^x - n_{il}^x(\hat{\mathbf{n}}_{il} \cdot \mathbf{E}_{lg})]}{Z_i + Z_l} \int_{a_{il}} \phi_{lg}\phi_{ik} \Bigg) \\ &- \int_{K_i} (J_{n_i}^x + J_{p_i}^x) \; \phi_{ik}, \end{aligned} \tag{4.7}$$

$$
\begin{aligned}
\sum_{j=1}^{d_i} \frac{\partial E_{ij}^y}{\partial t} \int_{K_i} \epsilon\ \phi_{ij}\phi_{ik} = &+ \sum_{j=1}^{d_i} H_{ij}^z \int_{K_i} \phi_{ij} \frac{\partial}{\partial x}\phi_{ik} \\
&- \sum_{l\in\mathcal{N}_i} \Big(\sum_{j=1}^{d_i} \frac{Z_i H_{ij}^z n_{il}^x + \alpha[E_{ij}^y - n_{il}^y(\hat{\mathbf{n}}_{il}\cdot\mathbf{E}_{ij})]}{Z_i + Z_l} \int_{a_{il}} \phi_{ij}\phi_{ik} \\
&+ \sum_{g=1}^{d_l} \frac{Z_l H_{lg}^z n_{il}^x - \alpha[E_{lg}^y - n_{il}^y(\hat{\mathbf{n}}_{il}\cdot\mathbf{E}_{lg})]}{Z_i + Z_l} \int_{a_{il}} \phi_{lg}\phi_{ik} \Big) \\
&- \int_{K_i} (J_{n_i}^y + J_{p_i}^y)\ \phi_{ik}.
\end{aligned} \tag{4.8}
$$

Finally, we write the semi-discrete DG formulation of the whole TE system, including dispersion (with the abbreviation $v = x, y$ when possible). On an element $K_i \subset \Omega$ we have

$$
\begin{aligned}
\mathbb{M}_i^{\mu} \frac{\mathbf{H}_i^z}{\partial t} &= \mathbb{K}_i^x \mathbf{E}_i^y - \mathbb{K}_i^y \mathbf{E}_i^x + \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il} \frac{Y_i \mathbf{E}_i^x n_{il}^y - Y_i \mathbf{E}_i^y n_{il}^x - \alpha \mathbf{H}_i^z}{Y_i + Y_l} \\
&\quad + \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}^* \frac{Y_l \mathbf{E}_l^x n_{il}^y - Y_l \mathbf{E}_l^y n_{il}^x + \alpha \mathbf{H}_l^z}{Y_i + Y_l}, \\
\mathbb{M}_i^{\epsilon} \frac{\mathbf{E}_i^x}{\partial t} &= -\mathbb{K}_i^y \mathbf{H}_i^z - \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il} \frac{-Z_i \mathbf{H}_i^z n_{il}^y + \alpha[\mathbf{E}_i^x - n_{il}^x(\hat{\mathbf{n}}_{il}\cdot\overline{\mathbf{E}}_i)]}{Z_i + Z_l} \\
&\quad - \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}^* \frac{-Z_l \mathbf{H}_l^z n_{il}^y - \alpha[\mathbf{E}_l^x - n_{il}^x(\hat{\mathbf{n}}_{il}\cdot\overline{\mathbf{E}}_l)]}{Z_i + Z_l} - \mathbb{M}_i \mathbf{J}_{p_i}^x - \mathbb{M}_i \mathbf{J}_{n_i}^x - \mathbb{M}_i \mathbf{J}_{l_i}^x, \\
\mathbb{M}_i^{\epsilon} \frac{\mathbf{E}_i^y}{\partial t} &= \mathbb{K}_i^x \mathbf{H}_i^z - \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il} \frac{Z_i \mathbf{H}_i^z n_{il}^x + \alpha[\mathbf{E}_i^y - n_{il}^y(\hat{\mathbf{n}}_{il}\cdot\overline{\mathbf{E}}_i)]}{Z_i + Z_l} \\
&\quad - \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}^* \frac{Z_l \mathbf{H}_l^z n_{il}^x - \alpha[\mathbf{E}_l^y - n_{il}^y(\hat{\mathbf{n}}_{il}\cdot\overline{\mathbf{E}}_l)]}{Z_i + Z_l} - \mathbb{M}_i \mathbf{J}_{p_i}^y - \mathbb{M}_i \mathbf{J}_{n_i}^y - \mathbb{M}_i \mathbf{J}_{l_i}^y, \\
q\mathbb{M}_i \frac{\partial \mathbf{n}_i}{\partial t} &= -\mathbb{K}_i^x \mathbf{J}_{n_i}^x - \mathbb{K}_i^y \mathbf{J}_{n_i}^y + \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}\ (\overline{\mathbf{J}}_{n_{*,i}}\cdot\hat{\mathbf{n}}_{il}) + \mathbb{S}_{il}^*\ (\overline{\mathbf{J}}_{n_{*,l}}\cdot\hat{\mathbf{n}}_{il}) - q\mathbb{M}_i \mathbf{R}_i + q\mathbb{M}_i \mathbf{G}_i, \\
q\mathbb{M}_i \frac{\partial \mathbf{p}_i}{\partial t} &= \mathbb{K}_i^x \mathbf{J}_{p_i}^x + \mathbb{K}_i^y \mathbf{J}_{p_i}^y - \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}\ (\overline{\mathbf{J}}_{p_{*,i}}\cdot\hat{\mathbf{n}}_{il}) - \mathbb{S}_{il}^*\ (\overline{\mathbf{J}}_{p_{*,l}}\cdot\hat{\mathbf{n}}_{il}) - q\mathbb{M}_i \mathbf{R}_i + q\mathbb{M}_i \mathbf{G}_i, \\
\mathbb{M}_i \boldsymbol{\alpha}_i^v &= \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}\ (\mathbf{n}_{*,i}\ n_{il}^v) + \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}^*\ (\mathbf{n}_{*,l}\ n_{il}^v) - \mathbb{K}_i^v \mathbf{n}_i, \\
\mathbb{M}_i \boldsymbol{\beta}_i^v &= \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}\ (\mathbf{p}_{*,i}\ n_{il}^v) + \sum_{l\in\mathcal{N}_i} \mathbb{S}_{il}^*\ (\mathbf{p}_{*,l}\ n_{il}^v) - \mathbb{K}_i^v \mathbf{p}_i, \\
\frac{\partial \mathbf{P}_i^v}{\partial t} &= \mathbf{J}_{l,i}^v, \\
\frac{\partial \mathbf{J}_{l,i}^v}{\partial t} &= \epsilon_0 \omega_p^2 \mathbf{E}_i^v - \gamma \mathbf{J}_{l,i}^v - \omega_r^2 \mathbf{P}_i^v.
\end{aligned} \tag{4.9}
$$

On an a metallic element $K_i \subset \Omega_M$,

$$
\begin{aligned}
\mathbb{M}_i^{\mu}\frac{\mathbf{H}_i^z}{\partial t} &= \mathbb{K}_i^x\mathbf{E}_i^y - \mathbb{K}_i^y\mathbf{E}_i^x + \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}\frac{Y_i\mathbf{E}_i^x n_{il}^y - Y_i\mathbf{E}_i^y n_{il}^x - \alpha\mathbf{H}_i^z}{Y_i+Y_l} \\
&\quad + \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}^*\frac{Y_l\mathbf{E}_l^x n_{il}^y - Y_l\mathbf{E}_l^y n_{il}^x + \alpha\mathbf{H}_l^z}{Y_i+Y_l}, \\
\mathbb{M}_i^{\epsilon}\frac{\mathbf{E}_i^x}{\partial t} &= -\mathbb{K}_i^y\mathbf{H}_i^z - \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}\frac{-Z_i\mathbf{H}_i^z n_{il}^y + \alpha[\mathbf{E}_i^x - n_{il}^x(\hat{\mathfrak{n}}_{il}\cdot\overline{\mathbf{E}}_i)]}{Z_i+Z_l} \\
&\quad - \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}^*\frac{-Z_l\mathbf{H}_l^z n_{il}^y - \alpha[\mathbf{E}_l^x - n_{il}^x(\hat{\mathfrak{n}}_{il}\cdot\overline{\mathbf{E}}_l)]}{Z_i+Z_l} - \mathbb{M}_i\mathbf{J}_{p_i}^x - \mathbb{M}_i\mathbf{J}_{n_i}^x - \mathbb{M}_i\mathbf{J}_{d_i}^x, \qquad (4.10) \\
\mathbb{M}_i^{\epsilon}\frac{\mathbf{E}_i^y}{\partial t} &= \mathbb{K}_i^x\mathbf{H}_i^z - \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}\frac{Z_i\mathbf{H}_i^z n_{il}^x + \alpha[\mathbf{E}_i^y - n_{il}^y(\hat{\mathfrak{n}}_{il}\cdot\overline{\mathbf{E}}_i)]}{Z_i+Z_l} \\
&\quad - \sum_{l\in\mathcal{N}_i}\mathbb{S}_{il}^*\frac{Z_l\mathbf{H}_l^z n_{il}^x - \alpha[\mathbf{E}_l^y - n_{il}^y(\hat{\mathfrak{n}}_{il}\cdot\overline{\mathbf{E}}_l)]}{Z_i+Z_l} - \mathbb{M}_i\mathbf{J}_{p_i}^y - \mathbb{M}_i\mathbf{J}_{n_i}^y - \mathbb{M}_i\mathbf{J}_{d_i}^y, \\
\frac{\partial\mathbf{J}_{d,i}^v}{\partial t} &= \epsilon_0\omega_d^2\mathbf{E}_i^v - \gamma_d\mathbf{J}_{d,i}^v.
\end{aligned}
$$

We recall the expression $\epsilon = \epsilon_0\epsilon_\infty$ inside $\mathbb{M}_i^{\epsilon}$ in both (4.9) and (4.10), with generally different values. Finally, $K_i \subset \Omega_e \setminus (\Omega_M \cup \Omega)$ is omitted as it is a sub-case of (4.10) with $\mathbf{J}_{d,i}^v = \mathbf{0}$ and $\epsilon = \epsilon_0\epsilon_r$. Analogously to the three-dimensional case, the above formulation is meant to be coupled with an explicit Runge-Kutta time-stepping scheme.

4.2 Code verification

There are three crucial tasks in the development of a DG solver for the MDD model: the implementation of the semiconductor equations, the coupling with Maxwell's equations in a bounded domain, the extension to a scattering problem. The process is broken down into elementary steps, which are outlined below.[2]

The sequence begins in the first subsection with test cases admitting analytical solutions: source-free TE Maxwell's equations, uniform advection, uniform diffusion, uniform advection-diffusion. The attention is restricted to electron transport in virtue of the similarity between the two semiconductor equations. The second subsection features more complex cases, for which manufactured solutions are considered: electron non-uniform advection-diffusion, unipolar and bipolar MDD in a bounded, then (ideally) unbounded domain, the inclusion of DC components and a simple field-dependent Poynting-like model for optical generation. The Maxwell-Lorentz system concludes the verification flow (the Drude model being a particular case of the Lorentz one).

To keep the notation light, the AC subscript is lifted in this section so that $\mathbf{E}, \mathbf{H}, n, p, \boldsymbol{\alpha}, \boldsymbol{\beta}$ denote the unknowns and are implicitly time-varying. Also, $R := R''$. The DC subscript is retained for the static counterparts, which are assumed to be known a priori.

[2]The two-dimensional formulation is referenced here but, as anticipated at the beginning of the chapter, all the steps can be easily adapted into three dimensions.

4.2.1 Verification with analytical solutions

We present here a sequence of test cases admitting analytical solutions and discuss relevant implementation details.

4.2.1.1 TE Maxwell's equations

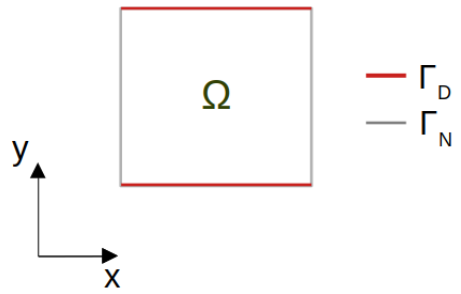


Figure 4.1: The square bounded domain Ω for the verification of source-free Maxwell's equations.

The TE version of Maxwell's equations in a two-dimensional source-free region Ω (Figure 4.1) reads

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\ \nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t}, \\ \mathbf{H}(x,y,t) &= \begin{bmatrix} 0 \\ 0 \\ H_z(x,y,t) \end{bmatrix}, \\ \mathbf{E}(x,y,t) &= \begin{bmatrix} E_x(x,y,t) \\ E_y(x,y,t) \\ 0 \end{bmatrix}. \end{aligned} \tag{4.11}$$

To the purpose of code verification it is assumed that $\Omega = [0,1]^2$ is the cross section of a square waveguide whose walls are made of a perfect electric conductor. The usual initial and boundary conditions are

$$\begin{aligned} \mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) && (x,y) \in \bar{\Omega}, \\ \mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0) && (x,y) \in \bar{\Omega}, \\ \hat{\mathbf{n}} \times \mathbf{E} &= \mathbf{0} && \text{on } \partial\Omega \times [0,+\infty[, \end{aligned} \tag{4.12}$$

where $\hat{\mathbf{n}}$ denotes the outward unit normal vector to $\partial\Omega$ as usual.

The problem admits analytical solutions, for example the TE11 mode

$$\begin{aligned}
\tilde{\mathbf{H}}(x,y,t) &= \begin{bmatrix} 0 \\ 0 \\ H_{nm}\cos(\frac{n\pi}{a}x)\ \cos(\frac{m\pi}{b}y)\ \cos(\omega t) \end{bmatrix}, \\
\tilde{\mathbf{E}}(x,y,t) &= \begin{bmatrix} -\frac{m\pi}{b}\frac{1}{\omega\epsilon}\ H_{nm}\ \cos(\frac{n\pi}{a}x)\ \sin(\frac{m\pi}{b}y)\ \sin(\omega t) \\ \frac{n\pi}{a}\frac{1}{\omega\epsilon}\ H_{nm}\ \sin(\frac{n\pi}{a}x)\ \cos(\frac{m\pi}{b}y)\ \sin(\omega t) \\ 0 \end{bmatrix}, \\
\omega\sqrt{\mu\epsilon} &= \sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2}, \\
n &= m = 1, \\
a &= b = 1\ \mathrm{m}, \\
H_{nm} &= 1\ \frac{\mathrm{A}}{\mathrm{m}}.
\end{aligned} \tag{4.13}$$

In fact, (4.13) is a particular case of the TE11 mode of a square waveguide, in that the general expression depends on z as well, through factors describing field propagation in the latter direction. Given our interest in solutions that only depend on x, y, an elegant way to get rid of the z dependence is by setting $\omega\sqrt{\mu\epsilon} = \sqrt{2\pi^2}$ (equivalently, $f = \frac{2\pi}{\omega} \approx 0.212$ GHz), which is the cut-off condition for the mode and hence prevents it from propagating in the waveguide [Bal12].

Noticing that (4.2) requires to enforce $\hat{\mathbf{n}} \times \mathbf{E}|_{\Gamma_D}$ and $\hat{\mathbf{n}} \times \mathbf{H}|_{\Gamma_N}$, it appears more appropriate to implement and test the non-standard conditions

$$\begin{aligned}
\mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) && (x,y) \in \bar{\Omega}, \\
\mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0) && (x,y) \in \bar{\Omega}, \\
\hat{\mathbf{n}} \times \mathbf{E} &= \hat{\mathbf{n}} \times \tilde{\mathbf{E}} && \text{on } \Gamma_D \times [0, +\infty[, \\
\hat{\mathbf{n}} \times \mathbf{H} &= \hat{\mathbf{n}} \times \tilde{\mathbf{H}} && \text{on } \Gamma_N \times [0, +\infty[,
\end{aligned} \tag{4.14}$$

which bring the additional advantage of making $(\tilde{\mathbf{E}}, \tilde{\mathbf{H}})$ the exact solution even when Ω is an arbitrary rectangle, rather than the unit square.

4.2.1.2 Diffusion (uniform diffusivity)

Let us consider the heat equation, again in the domain Ω depicted in Figure 4.1. It is known that the equation admits analytical solutions, e.g. the 11 mode

$$\tilde{n}(x,y,t) = \sin(\pi x)\ \sin(\pi y)\ e^{-2D_n\pi^2 t}. \tag{4.15}$$

Such a solution is imposed via the initial condition as well as Dirichlet and Neumann boundary conditions on Γ_D and Γ_N respectively, where $\partial\Omega = \Gamma_D \cup \Gamma_N$ as usual. The problem on which diffusion is tested is therefore

$$\begin{cases}
\frac{\partial}{\partial t} n = D_n \nabla^2 n & \text{in } \Omega \times [0, +\infty[, \\
n(x,y,0) = \tilde{n}(x,y,0) & (x,y) \in \Omega \cup \partial\Omega, \\
n(x,y,t) = \tilde{n}(x,y,t) & (x,y,t) \in \Gamma_D \times [0, +\infty[, \\
\hat{\mathbf{n}} \cdot \nabla n(x,y,t) = \hat{\mathbf{n}} \cdot \nabla \tilde{n}(x,y,t) & (x,y,t) \in \Gamma_N \times [0, +\infty[.
\end{cases} \tag{4.16}$$

4.2.1.3 Advection (uniform velocity)

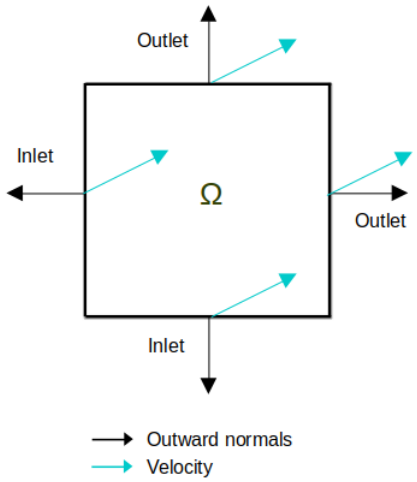


Figure 4.2: Test environment for uniform advection. Ω, a portion of a conceptually infinite domain, is crossed by a Gaussian pulse of particle density. Advected particles enter (leave) from edges on which the velocity field points inward (outward). Mathematically, an inlet/outlet is identified by looking at the scalar product between the velocity and the outward normal.

The first test case for advection is the propagation of a Gaussian pulse across the domain Ω (Figure 4.2) due to a uniform velocity field $\boldsymbol{v}_n$, from an initial position (x_0, y_0), in formulae

$$\begin{cases} \dfrac{\partial n}{\partial t} = -\boldsymbol{v}_n \cdot \nabla n & \text{in } \Omega \times [0, +\infty[, \\ n(x, y, 0) = \tilde{n}(x, y, 0) & (x, y) \in \Omega \cup \partial\Omega, \\ n(x, y, t) = \tilde{n}(x, y, t) & (x, y, t) \in \partial\Omega \times [0, +\infty[\text{ such that } \hat{\mathbf{n}} \cdot \boldsymbol{v}_n \leq 0. \end{cases} \tag{4.17}$$

Dirichlet boundary conditions are imposed at inlets, which are identified by the condition $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n < 0$. At outlets ($\hat{\mathbf{n}} \cdot \boldsymbol{v}_n > 0$) the numerical solution is let free, i.e. $n = n_i$ is coded on an outlet edge belonging to element K_i. This ensures that outlets perfectly absorb particles that are set to leave the computational domain.[3]
The analytical solution is

$$\tilde{n}(x, y, t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(x - x_0 - \boldsymbol{v}_n \cdot \hat{\mathbf{x}}\ t)^2 + (y - y_0 - \boldsymbol{v}_n \cdot \hat{\mathbf{y}}\ t)^2}{2\sigma^2} \right). \tag{4.18}$$

4.2.1.4 Advection-diffusion (uniform velocity and diffusivity)

To test both advection and diffusion, we let the uniform velocity field $\boldsymbol{v}_n$ advect (4.15) and obtain

$$\tilde{n}(x, y, t) = \sin(\pi(x - x_0 - \boldsymbol{v}_n \cdot \hat{\mathbf{x}}\ t))\ \sin(\pi(y - y_0 - \boldsymbol{v}_n \cdot \hat{\mathbf{y}}\ t))\ e^{-2D_n \pi^2 t}. \tag{4.19}$$

[3] $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n = 0$ is intentionally avoided in the simulation setup, with the idea of allowing it later, in the (more interesting) advection-diffusion case.

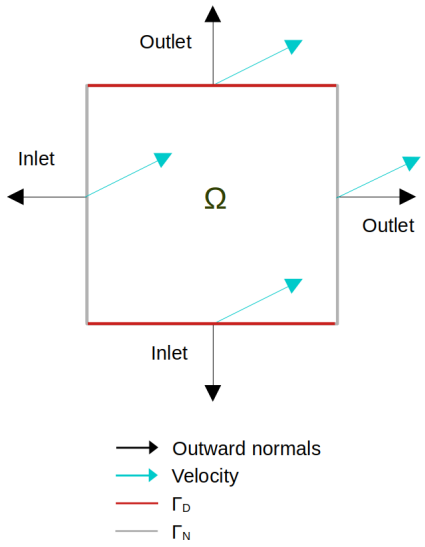


Figure 4.3: The bounded domain Ω in which uniform advection of the 11 mode of the heat equation is tested. Dirichlet and Robin boundary conditions define two types of inlets and outlets.

On Γ_D, Dirichlet boundary conditions are imposed as in (4.17); on Γ_N, Robin and Neumann boundary conditions are imposed at inlets and outlets respectively [LNX+20]. The resulting problem reads

$$\begin{cases} \dfrac{\partial n}{\partial t} = -\boldsymbol{v}_n \cdot \nabla n + D_n \nabla^2 n & \text{in } \Omega \times (0, +\infty), \\ n(x, y, 0) = \tilde{n}(x, y, 0) & (x, y) \in \Omega \cup \partial\Omega, \\ n = \tilde{n} & \text{on } \Gamma_D \times [0, +\infty[\text{ such that } \hat{\mathbf{n}} \cdot \boldsymbol{v}_n \leq 0, \\ -\hat{\mathbf{n}} \cdot \boldsymbol{v}_n \; n + \hat{\mathbf{n}} \cdot D_n \nabla n = g_{in} & \text{on } \Gamma_N \times [0, +\infty[\text{ such that } \hat{\mathbf{n}} \cdot \boldsymbol{v}_n \leq 0, \\ \hat{\mathbf{n}} \cdot D_n \nabla n = g_{out} & \text{on } \Gamma_N \times [0, +\infty[\text{ such that } \hat{\mathbf{n}} \cdot \boldsymbol{v}_n > 0, \end{cases} \tag{4.20}$$

where

$$\begin{aligned} g_{in} &= -\hat{\mathbf{n}} \cdot \boldsymbol{v}_n \; \tilde{n} + \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n}, \\ g_{out} &= \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n}. \end{aligned} \tag{4.21}$$

It is interesting to exploit this environment to separately simulate the corner cases $\boldsymbol{v}_n = \mathbf{0}$ and $D_n = 0$. To this purpose, the eventuality $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n = 0$ needs to be included in the Dirichlet condition, so that when $\boldsymbol{v}_n = \mathbf{0}$, $n = \tilde{n}$ is imposed everywhere on Γ_D and the heat equation is correctly closed. Note that when $\boldsymbol{v}_n \neq \mathbf{0}$, imposing $n = \tilde{n}$ at the points of Γ_D where $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n = 0$ is harmless - it entails no undesired injection or reflection of particles.
The issue has little relevance when it comes to Γ_N, as $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n = 0$ can be indifferently incorporated at inlets or outlets because inflow and outflow boundary conditions become identical.

To numerically solve (4.20) with a DG method, an auxiliary variable $\boldsymbol{\alpha} = \nabla \tilde{n}$ is introduced along with a dedicated equation as in (4.9). Since both the original and the auxiliary equations feature boundary integrals of n, one could be tempted to employ inflow and outflow concepts for

all of them; in fact, that would be incorrect in the case of the integrals appearing in the auxiliary equations, as these stem from the way diffusion is handled in the numerical scheme and thus are mathematical artifacts with no physical meaning. Hence we shall reason as follows.

- On Γ_D, n is assigned at inlets, thus in the transport equation one should code $\hat{\mathbf{n}} \cdot \mathbf{J}_{drift}$ with $n = \tilde{n}$ on such edges and take $n = n|_{K_i}$ at outlets, for each boundary element K_i touching Γ_D. In the auxiliary equation, the sensible choice is $n|_{\Gamma_D} = \tilde{n}$ regardless of $\hat{\mathbf{n}} \cdot \boldsymbol{v}_n$ because a) this allows to have a closed problem when drift is neglected and b) no flow is defined for diffusion.
- On Γ_N, $\hat{\mathbf{n}} \cdot \mathbf{J}_{drift}$ is constrained at inlets and let free at outlets, thus in the transport equation one should code $\hat{\mathbf{n}} \cdot \mathbf{J}_{drift}$ with $n = \tilde{n}$ at inlets and with the numerical n at outlets. There is no boundary condition on n itself, which is therefore let free in the auxiliary equation.

The reader might have noticed at this point that the boundary conditions implemented for code verification are flow-dependent, contrarily to the flow-independent conditions required by the physical model and stated in (2.59). The reason lies in the wish of having the code ready for possible future works entailing theoretically unlimited (and artificially truncated) semiconductor devices, as well as non-Ohmic contacts.

4.2.2 Verification with manufactured solutions

We now consider more complex test cases, in which analytical resolution is tricky or unfeasible, so that it is convenient to *manufacture* a solution. The procedure consists in *fixing* the system of equations under test, in that a designated function is made the exact solution by introducing *fictitious* source terms and enforcing suitable initial and boundary conditions.

4.2.2.1 Advection-diffusion (non-uniform velocity, uniform diffusivity)

We start by considering a non-uniform advection velocity, i.e. $\boldsymbol{v}_n = \boldsymbol{v}_n(x, y, t)$, as depicted in Figure 4.4, while keeping D_n constant. In the test, the system has the following form:

$$\begin{cases} \dfrac{\partial n}{\partial t} = \nabla \cdot (-\boldsymbol{v}_n \; n + D_n \nabla n) + \tilde{\mathbf{B}} & \text{in } \Omega \times [0, +\infty[, \\ n(x, y, 0) = \tilde{n}(x, y, 0) & (x, y) \in \Omega \cup \partial\Omega, \\ n = \tilde{n} & \text{on } \Gamma_D \times [0, +\infty[\;\text{ such that }\; \hat{\mathbf{n}} \cdot \boldsymbol{v}_n \leq 0, \\ -\hat{\mathbf{n}} \cdot \boldsymbol{v}_n \; n + \hat{\mathbf{n}} \cdot D_n \nabla n = g_{in} & \text{on } \Gamma_N \times [0, +\infty[\;\text{ such that }\; \hat{\mathbf{n}} \cdot \boldsymbol{v}_n \leq 0, \\ \hat{\mathbf{n}} \cdot D_n \nabla n = g_{out} & \text{on } \Gamma_N \times [0, +\infty[\;\text{ such that }\; \hat{\mathbf{n}} \cdot \boldsymbol{v}_n > 0, \end{cases} \tag{4.22}$$

where $\tilde{n}$ is the manufactured solution, for example

$$\tilde{n}(x, y, t) = \sin(\pi(x - x_0 - \boldsymbol{v}_n^u \cdot \hat{\mathbf{x}} \; t)) \; \sin(\pi(y - y_0 - \boldsymbol{v}_n^u \cdot \hat{\mathbf{y}} \; t)) \; e^{-2D_n \pi^2 t}, \tag{4.23}$$

which is noticeably the analytical solution to the advection-diffusion problem seen in the previous section with $\boldsymbol{v}_n = \boldsymbol{v}_n^u$ (an arbitrary uniform velocity). Then, to make the designated function (4.23) the exact solution to (4.22), we introduce the fictitious source

$$\tilde{\mathbf{B}} = -\Big(\nabla \cdot (-\boldsymbol{v}_n \; \tilde{n}) + \boldsymbol{v}_n^u \cdot \nabla \tilde{n} \Big). \tag{4.24}$$

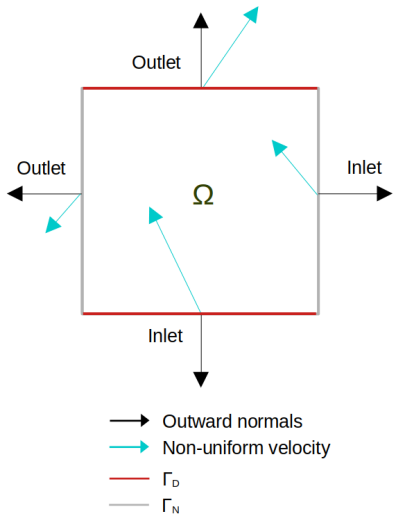


Figure 4.4: Test environment for non-uniform advection. The velocity field is derived from mid-points for purely illustrative purposes – it is space-varying across the whole domain and its boundary.

4.2.2.2 Unipolar Maxwell-Drift-Diffusion

On the ground of the previous case, Maxwell's equations can be coupled to electron transport by letting $\boldsymbol{v}_n(x,y,t) = -\mu_n \mathbf{E}(x,y,t)$. Also, the Einstein relation $D_n = V_T \mu_n$ is used. This novelty is compatible with the existing manufactured solution (4.23), which is therefore retained. For the electromagnetic field we take (4.13) and (4.14). The resulting system is

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t} + q n \mu_n \mathbf{E} + q D_n \boldsymbol{\alpha} + \tilde{\mathbf{A}}, \\
\frac{\partial n}{\partial t} &= \nabla \cdot (n \mu_n \mathbf{E}) + \nabla \cdot (D_n \boldsymbol{\alpha}) + \tilde{\mathbf{B}}, \\
\boldsymbol{\alpha} &= \nabla n, \\
&\text{in } \Omega \times [0, +\infty[,
\end{aligned} \tag{4.25}$$

$$
\begin{aligned}
\mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) && (x,y) \in \bar{\Omega}, \\
\mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0 && (x,y) \in \bar{\Omega}, \\
n(x,y,0) &= \tilde{n}(x,y,0) && (x,y) \in \bar{\Omega}, \\
\hat{\mathbf{n}} \times \mathbf{E} &= \hat{\mathbf{n}} \times \tilde{\mathbf{E}} && \text{on } \Gamma_D \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{H} &= \hat{\mathbf{n}} \times \tilde{\mathbf{H}} && \text{on } \Gamma_N \times [0,+\infty[, \\
n &= \tilde{n} && \text{on } \Gamma_D \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0, \\
n\ \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} \ + \ \hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \tilde{n}\ \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} \ + \ \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0, \\
\hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} < 0,
\end{aligned}
\tag{4.26}
$$

$$
\begin{aligned}
\tilde{\mathbf{A}} &= -\Big(q\tilde{n}\mu_n \tilde{\mathbf{E}} + qD_n \nabla \tilde{n} \Big), \\
\tilde{\mathbf{B}} &= -\Big(\nabla \cdot (\mu_n \tilde{\mathbf{E}}\ \tilde{n}) + \boldsymbol{v}_n^u \cdot \nabla \tilde{n} \Big).
\end{aligned}
\tag{4.27}
$$

A few remarks on (4.25-4.27) are opportune.

- $\hat{\mathbf{n}} \cdot \mathbf{E}$ appears at the right-hand side of the inflow boundary condition, rather than the exact counterpart $\hat{\mathbf{n}} \cdot \tilde{\mathbf{E}}$, consistently with the fact that such a condition is meant to constrain n and not the velocity.
- Inlets and outlets are now stated in terms of $\mathbf{E}$, hence inlets (outlets) are those where $\mathbf{E}$ points outward (inward) from $\partial\Omega$.
- Different notations are used for the gradients of n and $\tilde{n}$ respectively: ∇n is denoted by $\boldsymbol{\alpha}$ while no shorthand is introduced for $\nabla \tilde{n}$. The aim is to keep clear that the former is unknown while the latter is assigned.
- μ_n and D_n are uniform; they will vary in space when DC components are included and (2.45) is implemented.
- Manufacturing the exact solution allows to elegantly solve the issue of assigning $\hat{\mathbf{n}} \times \mathbf{H}$ on Γ_N, as anticipated in Section 2.1.1.

4.2.2.3 Maxwell-Drift-Diffusion

The code that implements electron transport can be reused for holes by replacing μ_n with $-\mu_p$. With bipolar transport in place, recombination can be introduced to couple the semiconductor equations and the Maxwell-Drift-Diffusion system can be assembled:[4]

[4]Generation is not included here - in the development flow it was decided to test it in a later, physically realistic picture.

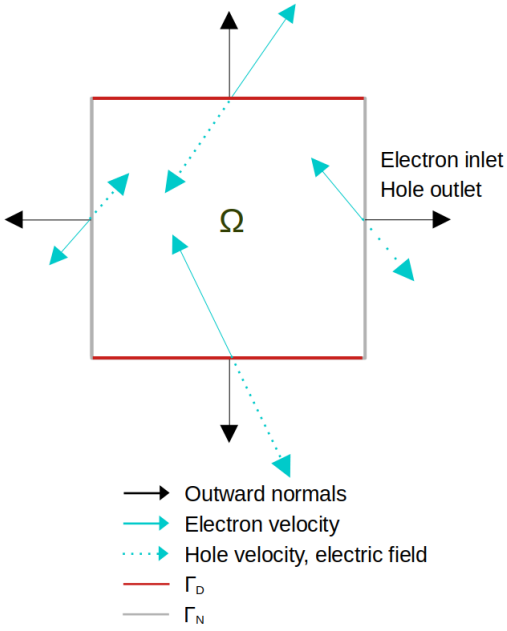


Figure 4.5: The domain Ω in which the Maxwell-Drift-Diffusion system is tested. To simplify the figure it is assumed here that $\boldsymbol{v}_p = \mathbf{E} = -\boldsymbol{v}_n$, i.e. $\mu_n = \mu_p = 1$. Note that inlets for holes are outlets for electrons and vice versa.

$$
\begin{aligned}
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t} + qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha} + qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta} + \tilde{\mathbf{A}}, \\
\frac{\partial n}{\partial t} &= \nabla \cdot (n\mu_n \mathbf{E}) + \nabla \cdot (D_n \boldsymbol{\alpha}) - R + \tilde{\mathbf{B}}, \\
\frac{\partial p}{\partial t} &= -\nabla \cdot (p\mu_p \mathbf{E}) + \nabla \cdot (D_p \boldsymbol{\beta}) - R + \tilde{\mathbf{C}}, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p, \\
&\text{in } \Omega \times [0, +\infty[,
\end{aligned}
\tag{4.28}
$$

$$
\begin{aligned}
\mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) && (x,y) \in \bar{\Omega},\\
\mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0) && (x,y) \in \bar{\Omega},\\
n(x,y,0) &= \tilde{n}(x,y,0) && (x,y) \in \bar{\Omega},\\
p(x,y,0) &= \tilde{p}(x,y,0) && (x,y) \in \bar{\Omega},\\
\hat{\mathbf{n}} \times \mathbf{E} &= \hat{\mathbf{n}} \times \tilde{\mathbf{E}} && \text{on } \Gamma_D \times [0,+\infty[,\\
\hat{\mathbf{n}} \times \mathbf{H} &= \hat{\mathbf{n}} \times \tilde{\mathbf{H}} && \text{on } \Gamma_N \times [0,+\infty[,\\
n &= \tilde{n} && \text{on } \Gamma_D \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0,\\
n\ \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} \ +\ \hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \tilde{n}\ \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} \ +\ \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0,\\
\hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} < 0,\\
p &= \tilde{p} && \text{on } \Gamma_D \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \leq 0,\\
p\ \hat{\mathbf{n}} \cdot \mu_p \mathbf{E} \ -\ \hat{\mathbf{n}} \cdot D_p \boldsymbol{\beta} &= \tilde{p}\ \hat{\mathbf{n}} \cdot \mu_p \mathbf{E} \ -\ \hat{\mathbf{n}} \cdot D_p \nabla \tilde{p} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \leq 0,\\
\hat{\mathbf{n}} \cdot D_p \boldsymbol{\beta} &= \hat{\mathbf{n}} \cdot D_p \nabla \tilde{p} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} > 0,
\end{aligned}
\tag{4.29}
$$

$$
\begin{aligned}
\tilde{\mathbf{A}} &= -\Big(q\tilde{n}\mu_n \tilde{\mathbf{E}} + qD_n \nabla \tilde{n} + q\tilde{p}\mu_p \tilde{\mathbf{E}} - qD_p \nabla \tilde{p} \Big),\\
\tilde{\mathbf{B}} &= -\Big(\nabla \cdot (\mu_n \tilde{\mathbf{E}}\ \tilde{n}) + \boldsymbol{v}_n^u \cdot \nabla \tilde{n} - \tilde{R} \Big),\\
\tilde{\mathbf{C}} &= -\Big(-\nabla \cdot (\mu_p \tilde{\mathbf{E}}\ \tilde{p}) + \boldsymbol{v}_p^u \cdot \nabla \tilde{p} - \tilde{R} \Big),
\end{aligned}
\tag{4.30}
$$

where $\tilde{p} = \tilde{n}|_{\boldsymbol{v}_n^u = \boldsymbol{v}_p^u}$ and $\tilde{R} := R(\tilde{n}, \tilde{p})$.

4.2.2.4 Immersion in a bounded dielectric

With reference to the realistic environment of Figure 2.2, we introduce a new manufactured solution for the electromagnetic field, for example a linearly polarized plane wave traveling in vacuum:

$$
\begin{aligned}
\tilde{\mathbf{E}}(x,y,t) &= \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix} E_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}),\\
\tilde{\mathbf{H}}(x,y,t) &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \frac{E_0}{\zeta_0} \cos(\omega t - \mathbf{k} \cdot \mathbf{r}),\\
\mathbf{k} = \mathbf{k}_0 &:= \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} \omega \sqrt{\epsilon_0 \mu_0},\\
\zeta_0 &= \sqrt{\frac{\mu_0}{\epsilon_0}},
\end{aligned}
\tag{4.31}
$$

where $\mathbf{r}$ denotes the position vector.

We retain the manufactured solution to transport equations to let the fictitious sources $\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}$ remain those in (4.30) (up to the new expression of $\tilde{\mathbf{E}}$).
In such a setting, we solve

$$
\begin{aligned}
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t} + \tilde{\mathbf{F}}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t} + \tilde{\mathbf{L}}, \\
&\qquad \text{in } \Omega_e \setminus \Omega \ \times [0, +\infty[, \\
\nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t} + \tilde{\mathbf{F}}, \\
\nabla \times \mathbf{H} &= \epsilon \frac{\partial \mathbf{E}}{\partial t} + qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha} + qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta} + \tilde{\mathbf{A}} + \tilde{\mathbf{L}}, \\
\frac{\partial n}{\partial t} &= \nabla \cdot (n\mu_n \mathbf{E}) + \nabla \cdot (D_n \boldsymbol{\alpha}) - R + \tilde{\mathbf{B}}, \\
\frac{\partial p}{\partial t} &= -\nabla \cdot (p\mu_p \mathbf{E}) + \nabla \cdot (D_p \boldsymbol{\beta}) - R + \tilde{\mathbf{C}}, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p, \\
&\qquad \text{in } \Omega \times [0, +\infty[,
\end{aligned}
\tag{4.32}
$$

with

$$
\begin{aligned}
\tilde{\mathbf{F}} &= \nabla \times \tilde{\mathbf{E}} + \mu \frac{\partial \tilde{\mathbf{H}}}{\partial t} = \mu_0(\mu_r - 1) \frac{\partial \tilde{\mathbf{H}}}{\partial t}, \\
\tilde{\mathbf{L}} &= \nabla \times \tilde{\mathbf{H}} - \epsilon \frac{\partial \tilde{\mathbf{E}}}{\partial t} = \epsilon_0(1 - \epsilon_r) \frac{\partial \tilde{\mathbf{E}}}{\partial t}.
\end{aligned}
\tag{4.33}
$$

The boundary conditions are

$$
\begin{aligned}
n(x,y,0) &= \tilde{n}(x,y,0) && (x,y) \in \bar{\Omega}, \\
p(x,y,0) &= \tilde{p}(x,y,0) && (x,y) \in \bar{\Omega}, \\
n &= \tilde{n} && \text{on } \Gamma_D \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0, \\
n\, \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} + \hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \tilde{n}\, \hat{\mathbf{n}} \cdot \mu_n \mathbf{E} + \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \geq 0, \\
\hat{\mathbf{n}} \cdot D_n \boldsymbol{\alpha} &= \hat{\mathbf{n}} \cdot D_n \nabla \tilde{n} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} < 0, \\
p &= \tilde{p} && \text{on } \Gamma_D \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \leq 0, \\
p\, \hat{\mathbf{n}} \cdot \mu_p \mathbf{E} - \hat{\mathbf{n}} \cdot D_p \boldsymbol{\beta} &= \tilde{p}\, \hat{\mathbf{n}} \cdot \mu_p \mathbf{E} - \hat{\mathbf{n}} \cdot D_p \nabla \tilde{p} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} \leq 0, \\
\hat{\mathbf{n}} \cdot D_p \boldsymbol{\beta} &= \hat{\mathbf{n}} \cdot D_p \nabla \tilde{p} && \text{on } \Gamma_N \times [0,+\infty[\text{ such that } \hat{\mathbf{n}} \cdot \mathbf{E} > 0, \\
\mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) && (x,y) \in \bar{\Omega}_e, \\
\mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0) && (x,y) \in \bar{\Omega}_e, \\
\hat{\mathbf{n}} \times \mathbf{E} &= \hat{\mathbf{n}} \times \tilde{\mathbf{E}} && \text{on } \Gamma_D \times [0,+\infty[, \\
\hat{\mathbf{n}} \times \mathbf{E} = \hat{\mathbf{n}} \times \tilde{\mathbf{E}} \ \text{ or } \ \hat{\mathbf{n}} \times \mathbf{H} &= \hat{\mathbf{n}} \times \tilde{\mathbf{H}} && \text{on } \Gamma_e \times [0,+\infty[.
\end{aligned}
\tag{4.34}
$$

There are two notable differences with respect to (2.11). Firstly, Γ_D has lost its physical meaning (a perfect electric conductor) due to its incompatibility with the manufactured solution – the exact

solution is $\mathbf{E} = \tilde{\mathbf{E}}$ on $\bar{\Omega}_e$ and thus the code is expected to produce $\hat{\mathbf{n}} \times \mathbf{E} = \hat{\mathbf{n}} \times \tilde{\mathbf{E}} \neq \mathbf{0}$ on Γ_D. As a matter of fact this is not really an issue, as the new condition is anyhow compliant with the fundamental requirement of verifying the code when the exact tangential component of the electric field is enforced on some internal edges, and whether the value is zero or not is not important. Secondly, the computational domain Ω_e is precisely the physical domain, so the boundary condition on Γ_e is not an ABC.
Implementation-wise, (3.33) is enforced on Γ_N as this is not a boundary in the electromagnetic problem. As to Γ_D, it can be regarded as the boundary of a collapsed subset of Ω_e excluded from the computational domain – in the original problem, a perfect electric conductor – on which the tangential component of the electric field is assigned as a boundary condition, while that of the magnetic field is let free, that is: in the local formulation on element K_i, one takes $\hat{\mathbf{n}}_{il} \times \mathbf{H}_* = \hat{\mathbf{n}}_{il} \times \mathbf{H}_i$ on edges that belong to Γ_D.

4.2.2.5 Immersion in an infinite dielectric

To simulate electromagnetic scattering, a few adjustments to the last test case are made.
Let Γ_e from Figure 2.2 become the artificial truncation of a theoretically unbounded propagation domain. We retain (4.31) as the manufactured solution, thus the fictitious sources just defined, but under slightly different roles: $\tilde{\mathbf{E}}, \tilde{\mathbf{H}}$ now represent the incident field, $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{L}}$ are meant to cancel the scattered one.
On Γ_e the following Silver-Muller ABC is imposed:

$$\hat{\mathbf{n}} \times \mathbf{E} + \hat{\mathbf{n}} \times \sqrt{\frac{\mu_0}{\epsilon_0}}(\hat{\mathbf{n}} \times \mathbf{H}) = \hat{\mathbf{n}} \times \tilde{\mathbf{E}} + \hat{\mathbf{n}} \times \sqrt{\frac{\mu_0}{\epsilon_0}}(\hat{\mathbf{n}} \times \tilde{\mathbf{H}}) \quad \text{on } \Gamma_e \times [0, +\infty[. \tag{4.35}$$

The interpretation of (4.35) is: when the outward normal to Γ_e is $\hat{\mathbf{n}} = -\hat{\mathbf{k}}_0$ the incident plane wave is injected into the computational domain, while when $\hat{\mathbf{n}} = \hat{\mathbf{k}}_0$ it is absorbed, i.e. it leaves with no reflection.
There are two key differences to remark with respect to a real scenario.

- Since the fictitious sources cancel the scattered field, only the incident plane wave hits Γ_e. In a real simulation - no exact solution available - the total field (see (2.9)) would have to be absorbed by means of the Silver-Muller condition stated in (2.11).
- The impedance of free space $\sqrt{\frac{\mu_0}{\epsilon_0}}$ appears in (4.35) regardless of the medium actually filling $\Omega_e \setminus \Omega$. This is motivated by the fact that the field to be absorbed or injected has the expression of a plane wave propagating in vacuum. Real ABCs would feature the impedance of $\Omega_e \setminus \Omega$.[5]

4.2.2.6 Adding DC components

The verification method can be finally applied to the simulation of a biased semiconductor device. The AC subscript is lifted for simplicity, thus $\mathbf{E}, \mathbf{H}, n, p, \boldsymbol{\alpha}, \boldsymbol{\beta}$ denote the unknowns, while $\mathbf{E}_{DC}, \mathbf{H}_{DC}, n_{DC}, p_{DC}$ have the usual meaning. The relations (2.45) imply that $\mu_i = \mu_i(\mathbf{E}_{DC})$, $D_i = \mu_i \, V_T$, $i = n, p$, are now space-varying.

[5] It would be licit of course for $(\tilde{\mathbf{E}}, \tilde{\mathbf{H}})$ to be a plane wave traveling in a dielectric medium with $\epsilon > \epsilon_0$ and/or $\mu > \mu_0$; then, such constants would appear in (4.35) instead of ϵ_0 and μ_0.

In the expressions of $\tilde{n}, \tilde{p}$ used so far, D_n, D_p were uniform by definition. To make the old manufactured solution work in the new picture, arbitrary uniform diffusivities D_n^u, D_p^u are introduced, yielding

$$\begin{aligned}\tilde{n}(x,y,t) &= \sin(\pi(x - x_0 - \boldsymbol{v}_n^u \cdot \hat{\mathbf{x}}\ t))\ \sin(\pi(y - y_0 - \boldsymbol{v}_n^u \cdot \hat{\mathbf{y}}\ t))\ e^{-2D_n^u \pi^2 t},\\ \tilde{p}(x,y,t) &= \sin(\pi(x - x_0 - \boldsymbol{v}_p^u \cdot \hat{\mathbf{x}}\ t))\ \sin(\pi(y - y_0 - \boldsymbol{v}_p^u \cdot \hat{\mathbf{y}}\ t))\ e^{-2D_p^u \pi^2 t}.\end{aligned} \tag{4.36}$$

Moreover, the simulation data are such that

$$\begin{aligned}\boldsymbol{v}_n^u &= -\mu_{n0}\mathbf{E}^u, & D_n^u &= \mu_{n0}\ V_T,\\ \boldsymbol{v}_p^u &= \mu_{p0}\mathbf{E}^u, & D_p^u &= \mu_{p0}\ V_T,\end{aligned} \tag{4.37}$$

with an arbitrary uniform electrostatic field $\mathbf{E}^u$.
The fictitious sources $\tilde{\mathbf{B}}, \tilde{\mathbf{C}}$ have now the role of canceling non-uniform drift and diffusion and replacing them by their uniform counterparts. Initial and boundary conditions are imposed on the sole AC quantities and are unaffected by the change in $\tilde{n}, \tilde{p}$, thus they retain the form (4.29) and are not repeated.
In the end the new system is

$$\begin{aligned}\nabla \times \mathbf{E} &= -\mu\frac{\partial \mathbf{H}}{\partial t} + \tilde{\mathbf{F}},\\ \nabla \times \mathbf{H} &= \epsilon\frac{\partial \mathbf{E}}{\partial t} + \tilde{\mathbf{L}},\\ &\quad \text{in } \Omega_e \setminus \Omega\ \times [0, +\infty[,\\ \nabla \times \mathbf{E} &= -\mu\frac{\partial \mathbf{H}}{\partial t} + \tilde{\mathbf{F}},\\ \nabla \times \mathbf{H} &= \epsilon\frac{\partial \mathbf{E}}{\partial t} + qn\mu_n(\mathbf{E} + \mathbf{E}_{DC}) + qD_n\boldsymbol{\alpha} + qp\mu_p(\mathbf{E} + \mathbf{E}_{DC}) - qD_p\boldsymbol{\beta}\\ &\quad + q\ n_{DC}\ \mu_n\mathbf{E} + q\ p_{DC}\ \mu_p\mathbf{E} + \tilde{\mathbf{A}} + \tilde{\mathbf{L}}\\ \frac{\partial n}{\partial t} &= \nabla \cdot (n\mu_n(\mathbf{E} + \mathbf{E}_{DC})) + \nabla \cdot (D_n\boldsymbol{\alpha}) + G - R\\ &\quad + \nabla \cdot (n_{DC}\ \mu_n\mathbf{E}) + \tilde{\mathbf{B}}\\ \frac{\partial p}{\partial t} &= -\nabla \cdot (p\mu_p(\mathbf{E} + \mathbf{E}_{DC})) + \nabla \cdot (D_p\boldsymbol{\beta}) + G - R\\ &\quad - \nabla \cdot (p_{DC}\ \mu_p\mathbf{E}) + \tilde{\mathbf{C}}\\ \boldsymbol{\alpha} &= \nabla n,\\ \boldsymbol{\beta} &= \nabla p,\\ &\quad \text{in } \Omega \times [0, +\infty[.\end{aligned} \tag{4.38}$$

$$
\begin{aligned}
\tilde{\mathbf{A}} &= -\Big(q\ \tilde{n}\ \mu_n(\tilde{\mathbf{E}} + \mathbf{E}_{DC}) + qD_n\nabla\tilde{n} + q\ \tilde{p}\ \mu_p(\tilde{\mathbf{E}} + \mathbf{E}_{DC}) - qD_p\nabla\tilde{p} \\
&\quad + q\ n_{DC}\ \mu_n\tilde{\mathbf{E}} + q\ p_{DC}\ \mu_p\tilde{\mathbf{E}}\Big), \\
\tilde{\mathbf{B}} &= -\Big(\nabla\cdot\big[\tilde{n}\ \mu_n(\tilde{\mathbf{E}} + \mathbf{E}_{DC})\big] + \nabla\cdot(n_{DC}\mu_n\tilde{\mathbf{E}}) \\
&\quad + \nabla\cdot(D_n\nabla\tilde{n}) + \boldsymbol{v}_n^u\cdot\nabla\tilde{n} - D_n^u\nabla^2\tilde{n} + \tilde{G} - \tilde{R}\Big), \\
\tilde{\mathbf{C}} &= -\Big(-\nabla\cdot\big[\tilde{p}\ \mu_p(\tilde{\mathbf{E}} + \mathbf{E}_{DC})\big] - \nabla\cdot(p_{DC}\mu_p\tilde{\mathbf{E}}) \\
&\quad + \nabla\cdot(D_p\nabla\tilde{p}) + \boldsymbol{v}_p^u\cdot\nabla\tilde{p} - D_p^u\nabla^2\tilde{p} + \tilde{G} - \tilde{R}\Big).
\end{aligned}
\tag{4.39}
$$

The meaning of G and $\tilde{G}$ demands clarification – the point is that (4.39) is common to two consecutive but quite different stages of the verification flow.

- The one in this section, which concludes a sequence of tests on the MDD system with no generation, in temporary geometrical and physical frameworks, different from the final ones; the opaque G's are set to zero.
- The one in the next section, in which the same equations are tested in a realistic geometrical and physical environment; the opaque G's are implemented.

In the end, opacity here simply avoids rewriting long formulae with minor changes.

The DC quantities being known a priori, they should be assigned in order to test the code. For example, as a simple debug test, Gaussian electrostatic field and linear charge densities have been considered:

$$
\begin{aligned}
g(x,y) &:= -\frac{(x-x_{0,DC})^2 + (y-y_{0,DC})^2}{2\sigma_{DC}^2}, \\
\frac{\partial g}{\partial x} &= -\frac{x-x_{0,DC}}{\sigma_{DC}^2}, \qquad \frac{\partial g}{\partial y} = -\frac{y-y_{0,DC}}{\sigma_{DC}^2}, \\
\mathbf{E}_{DC} &= \frac{\exp(g)}{\sqrt{2\pi\sigma_{DC}^2}}\begin{bmatrix}1\\1\end{bmatrix}, \\
\frac{\partial}{\partial x}\mathbf{E}_{DC} &= \mathbf{E}_{DC}\ \frac{\partial g}{\partial x}, \qquad \frac{\partial}{\partial y}\mathbf{E}_{DC} = \mathbf{E}_{DC}\ \frac{\partial g}{\partial y}, \\
n_{DC}(x,y) &= p_{DC}(x,y) = x - y + 1, \\
\boldsymbol{\alpha}_{DC}(x,y) &= \boldsymbol{\beta}_{DC}(x,y) = \begin{bmatrix}1\\-1\end{bmatrix},
\end{aligned}
\tag{4.40}
$$

the rationale being to ensure that the totality of the terms appearing in the unfolded expressions

of $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ be activated:

$$\begin{aligned}\tilde{\mathbf{B}} &= \boldsymbol{v}_n^u \cdot \nabla\tilde{n} + D_n^u \nabla^2\tilde{n} - \nabla\tilde{n} \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \mu_n - \nabla\mu_n \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \tilde{n} - \nabla \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \mu_n\ \tilde{n} \\ &\quad - \nabla D_n \cdot \nabla\tilde{n} - D_n \nabla^2\tilde{n} - \nabla n_{DC} \cdot \mathbf{E}\ \mu_n - \nabla\mu_n \cdot \mathbf{E}\ n_{DC} - \nabla \cdot \mathbf{E}\ \mu_n\ n_{DC} - \tilde{G} + \tilde{R}, \\ \tilde{\mathbf{C}} &= -\boldsymbol{v}_p^u \cdot \nabla\tilde{p} + D_p^u \nabla^2\tilde{p} + \nabla\tilde{p} \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \mu_p + \nabla\mu_p \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \tilde{p} + \nabla \cdot (\mathbf{E} + \mathbf{E}_{DC})\ \mu_p\ \tilde{p} \\ &\quad - \nabla D_p \cdot \nabla\tilde{p} - D_p \nabla^2\tilde{p} + \nabla p_{DC} \cdot \mathbf{E}\ \mu_p + \nabla\mu_p \cdot \mathbf{E}\ p_{DC} + \nabla \cdot \mathbf{E}\ \mu_p\ p_{DC} - \tilde{G} + \tilde{R}.\end{aligned} \tag{4.41}$$

The Caughey-Thomas model (2.45) was considered for mobility with $\beta_n = \beta_p = 1/2$, along with the Einstein relations to define diffusivity, so that

$$\begin{aligned}\nabla\mu_n &= -\frac{\mu_{n0}^3}{v_{n,sat}^2}\left[1 + \left(\frac{||\mathbf{E}||\ \mu_{n0}}{v_{n,sat}}\right)^2\right]^{-\frac{3}{2}} \nabla\mathbf{E}\ \mathbf{E}, \\ \nabla D_n &= V_T \nabla\mu_n,\end{aligned} \tag{4.42}$$

with $\nabla\mathbf{E}$ denoting the Jacobian matrix of $\mathbf{E}$.

4.2.3 Towards a physically meaningful setting

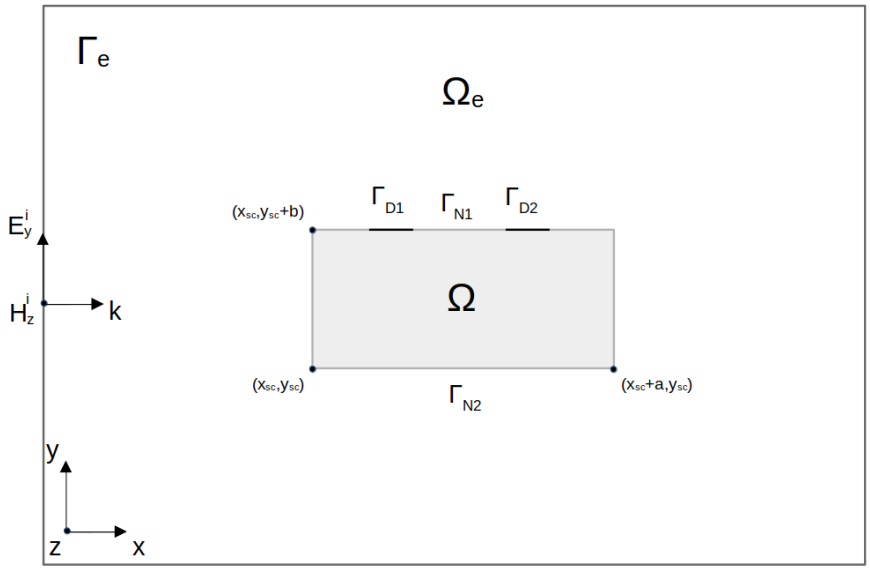


Figure 4.6: A bounded semiconductor domain $\Omega = [x_{sc}, x_{sc}+a] \times [y_{sc}, y_{sc}+b]$ surrounded by Ω_e, the truncation of a theoretically infinite dielectric medium. For illustration purposes, the propagation direction of $(\tilde{\mathbf{E}}, \tilde{\mathbf{H}})$ is assumed parallel to the x axis; therefore, the wave is injected from the western wall of Γ_e and absorbed by the eastern one.

We wish now to consider a more realistic setting for our tests, with the aim of preparing the ground for THz PCA simulation. To this purpose, space, time and frequency are scaled, realistic

values are set for physical parameters and the photo-generation function is finally introduced. However, the plan may not be fully compatible with the wish to track an exact solution – the issue is highlighted below.

The geometry under test is depicted in Figure 4.6 with $\Omega = [x_{sc}, x_{sc}+a] \times [y_{sc}, y_{sc}+b]$, and the assumption that a, b measure a few μm.
Physical parameters are set in agreement with a typical material of interest, e.g. LT-GaAs [MPG+14]:

$$\begin{aligned} \mu_{n0} &= 8000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}, & \mu_{p0} &= 400 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}, \\ D_{n0} &\approx 208 \text{ cm}^2 \text{ s}^{-1}, & D_{p0} &\approx 10 \text{ cm}^2 \text{ s}^{-1}, \\ n_i &= 9 \times 10^6 \text{ cm}^{-3}, & C_{An} &= C_{Ap} = 7 \times 10^{-30} \text{ cm}^6 \text{ s}^{-1}, \\ \tau_n &= 0.3 \text{ ps}, & \tau_p &= 0.4 \text{ ps}, \\ n_1 &= 4.5 \times 10^6 \text{ cm}^{-3}, & p_1 &= n_i^2/n_1, \end{aligned} \tag{4.43}$$

For simplicity, static quantities are assumed uniform and such that $n_{DC} = n_i = p_{DC}$. The electrostatic field is such that $||\mathbf{E}_{DC}|| \sim 10^6$ V m^{-1} so that carrier mobility can legitimately be assumed independent of $\tilde{\mathbf{E}}$. For example, $\mathbf{E}_{DC} = (-0.5, -1.0,\ 0) \times 10^6$ V m^{-1}.

Given the interest in pulsed excitation, the new manufactured electromagnetic field has the form

$$\begin{aligned} \tau &= t - t_0 - \frac{\mathbf{k}_0 \cdot \mathbf{r}}{\omega}, \\ \mathcal{G}(\tau) &= E_0 \exp\left(- \frac{\tau^2}{2\sigma_t^2} \right), \\ \tilde{\mathbf{E}}(x,y,t) &= \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix} \mathcal{G}(\tau) \cos(\omega\tau), \\ \tilde{\mathbf{H}}(x,y,t) &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \frac{\mathcal{G}(\tau)}{\zeta_0} \cos(\omega\tau), \end{aligned} \tag{4.44}$$

with

$$\begin{aligned} E_0 &= 10^4 \text{ V m}^{-1}, & \lambda &= \frac{2\pi}{\omega} = 800 \text{ nm}, \\ t_0 &= 3 \times 10^{-13} \text{ s}, & \sigma_t &= 8 \times 10^{-14} \text{ s}. \end{aligned} \tag{4.45}$$

Space modulation might be considered as well, but this would add complexity to $\tilde{\mathbf{L}}, \tilde{\mathbf{F}}$ without making tests any more probative.[6] Also, the angle of incidence (the propagation direction $\hat{\mathbf{k}}$) is not relevant as scattering is still suppressed by the fictitious sources.

Manufactured charge densities (4.36) serve well independently of domain coordinates as long as they are enforced by initial and boundary conditions; however some scaling is necessary to make them appreciably evolve in the new spatial and temporal framework. For example we take the 11 mode of the heat equation in $[0, a] \times [0, b]$.[7]

[6] An advantage of (4.44) is that it still is an analytical solution to Maxwell's equations in vacuum – it wouldn't be if space modulation were introduced.

[7] For simplicity we set $x_{sc} = 0 = y_{sc}$ in the manufactured solution with no loss of generality – the functions comply with the key requirement of witnessing evolution in space and time, and they solve the problem thanks to fictitious sources and to the enforcement exerted by initial and boundary conditions.

To get closer to the underlying physics it is sensible to let the uniform velocity be proportional to a uniform static electric field via the low-field mobility; also, the uniform diffusivity is taken equal to the low-field diffusivity $D_0 = V_T \mu_0$. As a result,

$$\begin{aligned}
\tilde{n}(x,y,t) &= \sin\left(\frac{\pi}{a}(x - x_0 + \mu_{n0}\mathbf{E}^u \cdot \hat{\mathbf{x}}\ t)\right)\ \sin\left(\frac{\pi}{b}(y - y_0 + \mu_{n0}\mathbf{E}^u \cdot \hat{\mathbf{y}}\ t)\right) \\
&\quad \times \exp\left(-D_{n0}\left[\left(\frac{\pi}{a}\right)^2 + \left(\frac{\pi}{b}\right)^2\right] t\right), \\
\tilde{p}(x,y,t) &= \sin\left(\frac{\pi}{a}(x - x_0 - \mu_{p0}\mathbf{E}^u \cdot \hat{\mathbf{x}}\ t)\right)\ \sin\left(\frac{\pi}{b}(y - y_0 - \mu_{p0}\mathbf{E}^u \cdot \hat{\mathbf{y}}\ t)\right) \\
&\quad \times \exp\left(-D_{p0}\left[\left(\frac{\pi}{a}\right)^2 + \left(\frac{\pi}{b}\right)^2\right] t\right).
\end{aligned} \tag{4.46}$$

The uniform electric field should be strong enough to make the waveforms $\tilde{n}, \tilde{p}$ travel distances comparable with the device size within the illumination window, which is in the order of $2t_0$. The simplest choice is $\mathbf{E}^u = \mathbf{E}_{DC}$.

An instantaneous photo-generation model inspired to (2.21) is considered at this stage, in which dispersion is not yet in the game:

$$G(x,y,z,t) = \frac{\alpha}{\hbar\omega}\frac{||\mathbf{E}(x,y,z,t)||^2}{2\zeta_0}. \tag{4.47}$$

Even if not physically rigorous, such a model is probative – it is instantaneous and field-dependent, and allows to highlight the core issue of this subsection.
Looking for example at electrons, the transport equation has the form

$$\frac{\partial n}{\partial t} = \ldots + G - \tilde{G} = \ldots + \frac{\alpha}{2\ \zeta_0\ \hbar\ \omega}\left(||\mathbf{E}||^2 - ||\tilde{\mathbf{E}}||^2\right). \tag{4.48}$$

Given that $\zeta_0 \approx 377\ \Omega$ and $\hbar\omega \sim 10^{19}$ J for $\lambda = 800$ nm, if $\alpha \sim \mu\text{m}^{-1}$ [MPG+14], it turns out that a factor as high as 10^{22} (m J Ω)$^{-1}$ amplifies the difference between the squared norms of the numerical and the exact electric field. Hence $G - \tilde{G}$ becomes a gigantic, undesired source term that a) strongly pulls n away from $\tilde{n}$ (and towards overflow) throughout the whole illumination window and b) is impossible to annihilate with any realistic interpolation order and mesh size.

A spontaneous fix could be to also consider a physically meaningful value for $\tilde{n}$, possibly an amplitude much larger than $\max(|G - \tilde{G}|)$, but that would only lift the problem from an equation to land it into another, namely

$$\nabla \times \mathbf{H} = \ldots + n\mu_n\mathbf{E} - \tilde{n}\mu_n\tilde{\mathbf{E}} \approx \ldots + \tilde{n}\mu_n\left(\mathbf{E} - \tilde{\mathbf{E}}\right), \tag{4.49}$$

with a new unbearable amplification, this time of $\mathbf{E} - \tilde{\mathbf{E}}$. Alternatively, one could opt for a normalization like $\frac{G-\tilde{G}}{\tilde{G}}$ stemming from a legitimate interest in the relative - rather than absolute - error on the field, but division by a Gaussian function is not convenient. The same philosophy actually suggests a simpler, viable option: multiplying $G - \tilde{G}$ by an attenuation factor, i.e. replacing $G - \tilde{G}$ by $\xi\ (G - \tilde{G})$ for some ξ.

To highlight the effects of ξ, let us consider a fixed geometrical and parametric setting, ideally in the drift-only limit, so that it is possible to monitor the relative error

$$\frac{||n - \tilde{n}||_{L^2(\Omega)}}{||\tilde{n}||_{L^2(\Omega)}}, \tag{4.50}$$

which becomes uninteresting in the presence of diffusion (the denominator asymptotically tends to zero).[8]

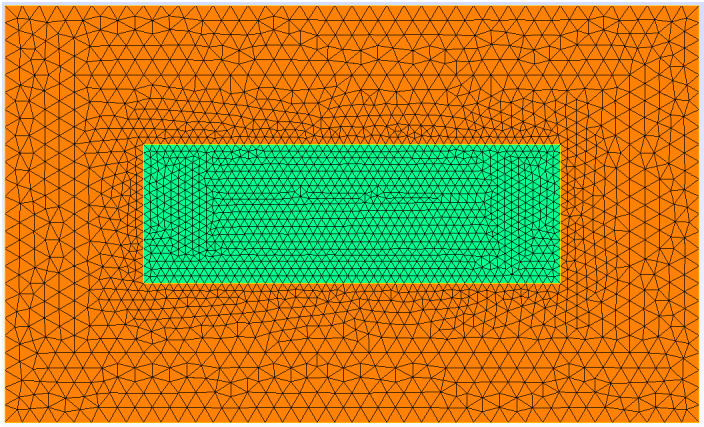


Figure 4.7: The unstructured mesh used to observe the effects of different orders of magnitude of the generation function. In green, the semiconductor domain $\Omega = [-4.8, 4.8] \times [-1.6, 1.6]$ μm, bounded by $\partial\Omega = \Gamma_N$. In orange, the surrounding dielectric, terminated with ABCs at 0.8 μm from $\partial\Omega$.

For example, based on the physical setting (4.43-4.45), we let $\Omega = [-a, a] \times [-b, b]$, $a = 4.8$ μm and $b = 1.6$ μm, $\partial\Omega = \Gamma_N$ (electrically insulating boundary).
A non-uniform mesh (Figure 4.7) is generated with $h = 0.05$ μm inside Ω and $h = 0.1$ μm at the boundary of the computational domain Γ_e, which lies at a distance of one wavelength (0.8 μm) from the semiconductor.
Time discretization relies on a RK2 scheme with $\Delta t = 4.5 \times 10^{-14}$ s; linear Lagrange polynomials are used for discretization in space.

Figure 4.8 illustrates the evolution of the relative error (4.50) throughout the simulation time $[0, 2t_0]$ for different values of an attenuation factor ξ multiplying $G - \tilde{G}$. A strong increase is observed in the interval $[0, t_0]$ in which the electromagnetic field arises and reaches its peak ($t = t_0$); from this time on, the field fades out and the error settles to a constant value. This suggests that if the field was still monochromatic, rather than pulsed, the error would keep growing at such a rapid rate. Tests have led to conclude that with the current set of parameters the generation function needs to be attenuated by at least 10^{-19} to keep (4.50) below 5%.

It should be noted that the tests on G were run with DC components switched off, as these can generate stability issues as well. Indeed, repeating the above analysis for n_{DC} and p_{DC} would lead to the need for attenuation factors. The details would be redundant and are therefore omitted

[8] As usual, observing n entails no loss of generality – similar conclusions can be drawn on p.

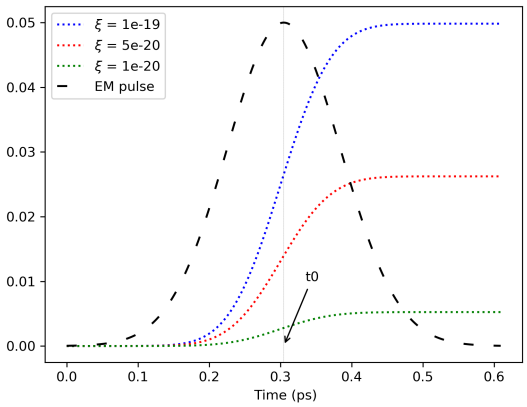


Figure 4.8: The electromagnetic pulse (re-scaled for illustration purposes) and the resulting evolution in $[0, t_0]$ of the relative error on electron concentration (4.50) for different orders of magnitude of $\xi(G - \tilde{G})$, set by means of the attenuation factor ξ. The error settles to a constant value after the generation peak $t = t_0$.

– the discussion is concluded by elaborating on the nature of such instability. For example, the electron transport equation has the form

$$\begin{aligned} \frac{\partial n}{\partial t} &= \ldots + \nabla \cdot \left(n_{DC} \mu_n \mathbf{E}\right) - \nabla \cdot \left(n_{DC} \mu_n \tilde{\mathbf{E}}\right) \\ &= \ldots + n_i \ \nabla \cdot \left(\mu_n \mathbf{E} - \mu_n \tilde{\mathbf{E}}\right). \end{aligned} \tag{4.51}$$

Then, retaining the setting $n_{DC} = n_i$, with a typical value of the order of 10^6 cm$^{-3} = 10^{12}$ m^{-3} in GaAs, the divergence of the deviation $\boldsymbol{v}_{n,drift} - \tilde{\boldsymbol{v}}_{n,drift}$ is likely to be unbearably amplified unless artificially attenuated.

Simulation results for this test case, with both drift and diffusion and suitable attenuation factors for $G - \tilde{G}$, n_{DC}, p_{DC} are shown in Section 4.4.

4.2.4 Verifying dispersion models

In the context of semiconductor simulation, dispersion logically comes in the game after optical generation. The polarization current appears as a source in charge continuity equations, but is affected by them via the electric field, hence indirectly. The key point is that the auxiliary differential equations describe the dielectric response of the medium regardless of its nature (a metal, a semiconductor or an insulator). Thus, to embed dispersion in the verification flow, n and p don't need to be calculated - it suffices to just simulate Maxwell's equations and the Lorentz ones. Moreover, there is no need to consider the Drude model as it is a simpler sub-case of the tested one.

To investigate the interaction of a $\mathbf{E}, \mathbf{H}$ with $\mathbf{J}_l, \mathbf{P}$ at optical frequencies, we consider an artificially truncated dispersive semiconductor regarded as a dielectric (i.e. with no DD equations) and let the manufactured electromagnetic solution be the monochromatic plane wave (4.31) with $\omega = \tilde{\omega} = 2\pi/\lambda$, $\lambda = 800$ nm, $\mathbf{k}_0 = \tilde{\omega}\sqrt{\epsilon_0\mu_0}$. The computational domain can be as simple as a square $\Omega_e = [-\lambda, \lambda]^2$, terminated with the ABCs (4.35), and discretized with a uniform structured mesh.
Fixing an exact solution for $\mathbf{E}$ allows to calculate $\mathbf{P}$ and $\mathbf{J}_l$ analytically; to this purpose it is convenient to deal in terms of the complex electric susceptibility

$$\chi(\omega) = \frac{\omega_p^2}{\omega_r^2 - \omega^2 + i\omega\gamma} = \chi'(\omega) - j\chi''(\omega), \tag{4.52}$$

with

$$\chi'(\omega) = \rho(\omega)(\omega_r^2 - \omega^2), \qquad \chi''(\omega) = \rho(\omega)\ \omega\gamma, \tag{4.53}$$

$$\rho(\omega) := \frac{\omega_p^2}{(\omega_r^2 - \omega^2)^2\ +\ \omega^2\gamma^2}, \tag{4.54}$$

and to define

$$\begin{aligned} \mathbf{E}_0 &= (-\sin\theta, \cos\theta,\ 0)\ E_0, \\ \mathcal{E}(\mathbf{r}, t) &= \cos(\tilde{\omega}t - \mathbf{k}_0 \cdot \mathbf{r}), \end{aligned} \tag{4.55}$$

so that the exact electric field is rewritten as

$$\tilde{\mathbf{E}}(\mathbf{r}, t) = \mathbf{E}_0\mathcal{E}(\mathbf{r}, t). \tag{4.56}$$

The exact polarization density is then:[9]

$$\tilde{\mathbf{P}}(\mathbf{r}, t) = \mathfrak{F}^{-1}\left[\tilde{\mathbf{P}}(\mathbf{r}, \omega)\right] = \mathfrak{F}^{-1}\left[\epsilon_0\chi(\omega)\ \tilde{\mathbf{E}}(\mathbf{r}, \omega)\right] = \epsilon_0\mathbf{E}_0\ \mathfrak{F}^{-1}\left[\chi(\omega)\ \mathcal{E}(\mathbf{r}, \omega)\right]. \tag{4.57}$$

Dirac delta function's properties and hermiticity of $\chi(\omega)$ allow to easily calculate the inverse Fourier

[9]For the purpose of readability, in these calculations the initial notation $\overline{\mathbf{P}}(\mathbf{r}, \omega) := \mathfrak{F}[\mathbf{P}](\mathbf{r}, \omega)$ is simplified and the Fourier transform of $\mathbf{P}(\mathbf{r}, t)$ is implicitly denoted by $\mathbf{P}(\mathbf{r}, \omega)$. Similarly for scalar quantities.

transform

$$
\begin{aligned}
&\mathfrak{F}^{-1}\left[\chi(\omega)\ \mathcal{E}(\mathbf{r},\omega)\right] = \mathfrak{F}^{-1}\left[\chi(\omega)\ 2\pi\ \frac{\delta(\omega-\tilde{\omega})\ e^{-i\mathbf{k_0}\cdot\mathbf{r}} + \delta(\omega+\tilde{\omega})\ e^{i\mathbf{k_0}\cdot\mathbf{r}}}{2}\right] \\
&= \mathfrak{F}^{-1}\left[\chi(\tilde{\omega})\ 2\pi\ \frac{\delta(\omega-\tilde{\omega})\ e^{-i\mathbf{k_0}\cdot\mathbf{r}}}{2}\right] + \mathfrak{F}^{-1}\left[\chi(-\tilde{\omega})\ 2\pi\ \frac{\delta(\omega+\tilde{\omega})\ e^{+i\mathbf{k_0}\cdot\mathbf{r}}}{2}\right] \\
&= \left[\chi(\tilde{\omega})\ \frac{e^{i(\tilde{\omega}t-\mathbf{k_0}\cdot\mathbf{r})}}{2}\right] + \left[\chi^*(\tilde{\omega})\ \frac{e^{-i(\tilde{\omega}t-\mathbf{k_0}\cdot\mathbf{r})}}{2}\right] \\
&= 2Re\left\{\chi(\tilde{\omega})\ \frac{e^{i(\tilde{\omega}t-\mathbf{k_0}\cdot\mathbf{r})}}{2}\right\} \\
&= \chi'(\tilde{\omega})\cos(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r}) + \chi''(\tilde{\omega})\sin(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r}).
\end{aligned}
\tag{4.58}
$$

In the end, the analytical solution to the dispersion equations is

$$
\begin{aligned}
\tilde{\mathbf{P}}(\mathbf{r},t) &= \epsilon_0\mathbf{E}_0\left[\chi'(\tilde{\omega})\cos(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r}) + \chi''(\tilde{\omega})\sin(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r})\right], \\
\tilde{\mathbf{J}}_l(\mathbf{r},t) &= \frac{\partial\tilde{\mathbf{P}}}{\partial t} = \tilde{\omega}\ \epsilon_0\mathbf{E}_0\left[\chi''(\tilde{\omega})\cos(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r}) - \chi'(\tilde{\omega})\sin(\tilde{\omega}t-\mathbf{k}_0\cdot\mathbf{r})\right],
\end{aligned}
\tag{4.59}
$$

and the assembled model reads

$$
\begin{aligned}
\nabla\times\mathbf{E} &= -\mu\frac{\partial\mathbf{H}}{\partial t} + \tilde{\mathbf{F}}, \\
\nabla\times\mathbf{H} &= \epsilon\frac{\partial\mathbf{E}}{\partial t} + \mathbf{J}_l + \tilde{\mathbf{L}} - \tilde{\mathbf{J}}_l, \\
\frac{\partial\mathbf{P}}{\partial t} &= \mathbf{J}_l, \\
\frac{\partial\mathbf{J}_l}{\partial t} &= \epsilon_0\omega_p^2\mathbf{E} - \gamma\mathbf{J}_l - \omega_r^2\mathbf{P}, \\
&\text{in } \Omega_e\times[0,+\infty[,
\end{aligned}
\tag{4.60}
$$

$$
\begin{aligned}
\mathbf{E}(x,y,0) &= \tilde{\mathbf{E}}(x,y,0) \qquad (x,y)\in\bar{\Omega}_e, \\
\mathbf{H}(x,y,0) &= \tilde{\mathbf{H}}(x,y,0) \qquad (x,y)\in\bar{\Omega}_e, \\
\mathbf{P}(x,y,0) &= \tilde{\mathbf{P}}(x,y,0) \qquad (x,y)\in\bar{\Omega}_e, \\
\mathbf{J}_l(x,y,0) &= \tilde{\mathbf{J}}_l(x,y,0) \qquad (x,y)\in\bar{\Omega}_e,
\end{aligned}
\tag{4.61}
$$

and is closed with the absorbing condition (4.35) on the boundary of the computational domain $\Gamma_e = \partial\Omega_e$. The fictitious sources $\tilde{\mathbf{F}}, \tilde{\mathbf{L}}$ are precisely those in (4.33) and their role is to cancel the scattered field; however an extra term is needed here to complete the task, namely $-\tilde{\mathbf{J}}_l$ in the Ampère-Maxwell law. No fictitious sources are needed in the dispersion equations of course, as these are solved analytically.

Following the discussion from the previous section, the order of magnitude of γ, ω_r, ω_p is to be carefully monitored. Looking at ω_p, for example the key fact is that (4.59) analytically solves

the equations as long as $\mathbf{E} = \tilde{\mathbf{E}}$, which is only approximately true. The consequences can be immediately grasped if $\Delta\mathbf{E} := \mathbf{E} - \tilde{\mathbf{E}}$ is introduced in the last equation of (4.60):

$$\frac{\partial \mathbf{J}_l}{\partial t} = \epsilon_0 \omega_p^2 \tilde{\mathbf{E}} - \gamma \mathbf{J}_l - \omega_r^2 \mathbf{P} + \epsilon_0 \omega_p^2 \Delta\mathbf{E}. \tag{4.62}$$

The numerical error on the electric field behaves as an additional, undesired source, and drifts (4.59) off target – the actual exact solution to the dispersion equations changes (and becomes unknown). Assigning realistic values to ω_p exacerbates the effect, as these can easily be in the order of 10^{17} rad/s. Similar effects can be observed when tuning the state terms $\gamma\mathbf{J}_l$, and $\omega_r^2\mathbf{P}$.

It should be noted that the aforementioned effects stem from an inherent pitfall to tests in which an exact solution needs to be tracked, and will disappear when simulating real scenarios. The point is further discussed in the Section 4.4.

4.3 Measurement units in the implementation

The implemented DGTD solver follows usual conventions, that can be found for example in [Koe11] and [Viq18], and are applied in the DIOGENeS software [Atl], which make μ_0 and ϵ_0 implicit in Maxwell's equations by a suitable change in units. The procedure is outlined below and extended to the DD model.
Bearing in mind the identities

$$\zeta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \qquad c_0 = \sqrt{\frac{1}{\mu_0\epsilon_0}} \qquad \zeta_0 c_0 = \frac{1}{\epsilon_0} \qquad \frac{\zeta_0}{c_0} = \mu_0, \tag{4.63}$$

and defining

$$\mathbf{E} = \mathbf{E}, \qquad \mathbf{H} = \zeta_0 \mathbf{H}, \qquad t = c_0 t, \tag{4.64}$$

Maxwell's equations are modified as follows:

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mu_r \frac{\zeta_0}{c_0} \frac{\partial \mathbf{H}}{\partial t} && \iff \nabla \times \mathbf{E} = -\mu_r \frac{\partial \mathbf{H}}{\partial t}, \\ \nabla \times \mathbf{H} &= \epsilon_r \frac{1}{\zeta_0 c_0} \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_n + \mathbf{J}_p + \mathbf{J}_l && \iff \nabla \times \mathbf{H} = \epsilon_r \frac{\partial \mathbf{E}}{\partial t} + \zeta_0 \big(\mathbf{J}_n + \mathbf{J}_p + \mathbf{J}_l\big), \end{aligned} \tag{4.65}$$

where as usual

$$\begin{aligned} \mathbf{J}_n &= \mathbf{J}_{n,drift} + \mathbf{J}_{n,diff} = qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha}, \\ \mathbf{J}_p &= \mathbf{J}_{p,drift} + \mathbf{J}_{p,diff} = qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta}. \end{aligned} \tag{4.66}$$

The transformation is convenient computation-wise as it avoids multiplication by $\mu_0 = 4\pi \times 10^{-7}$ H/m and $\epsilon_0 \approx 8.854 \times 10^{-12}$ F/m at each iteration, whereas multiplying by $\zeta_0 \approx 377\ \Omega$ is less of a concern.

To extend the change in units to the DD equations we start by introducing t, which yields

$$\begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{c_0}\left(\frac{1}{q}\nabla \cdot \mathbf{J}_n + G - R\right) = \nabla \cdot \left(n\frac{\mu_n}{c_0}\mathbf{E} + \frac{D_n}{c_0}\boldsymbol{\alpha}\right) + \frac{G-R}{c_0}, \\ \frac{\partial p}{\partial t} &= \frac{1}{c_0}\left(-\frac{1}{q}\nabla \cdot \mathbf{J}_p + G - R\right) = -\nabla \cdot \left(p\frac{\mu_p}{c_0}\mathbf{E} + \frac{D_p}{c_0}\boldsymbol{\beta}\right) + \frac{G-R}{c_0}. \end{aligned} \tag{4.67}$$

At this point the currents appear divided by c_0 (and also by $q \approx 1.602 \times 10^{-19}$ C) in the DD equations and multiplied by ζ_0 in Maxwell's equations. Taking advantage of the fact that q does not play any role in the former, we set

$$q = \zeta_0 c_0 \; q = \frac{q}{\epsilon_0} \approx 1.81 \times 10^{-8} \text{ V m}. \tag{4.68}$$

The resulting system, reads

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\mu_r \frac{\partial \mathbf{H}}{\partial t}, \\
\nabla \times \mathbf{H} &= \epsilon_r \frac{\partial \mathbf{E}}{\partial t} + qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha} + qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta} + \mathbf{J}_l, \\
\frac{\partial n}{\partial t} &= \nabla \cdot \left(n\mu_n \mathbf{E} + D_n \boldsymbol{\alpha} \right) + G - R, \\
\frac{\partial p}{\partial t} &= -\nabla \cdot \left(p\mu_p \mathbf{E} - D_p \boldsymbol{\beta} \right) + G - R, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p,
\end{aligned} \tag{4.69}$$

with

$$\begin{aligned}
\mathbf{E} &= \mathbf{E}, & \mathbf{H} &= \zeta_0 \mathbf{H}, \\
t &= c_0 t, & q &= \frac{q}{\epsilon_0}, \\
\mu_n &= \frac{\mu_n}{c_0}, & D_n &= \frac{D_n}{c_0}, \\
\mu_p &= \frac{\mu_p}{c_0}, & D_p &= \frac{D_p}{c_0}, \\
G &= \frac{G}{c_0}, & R &= \frac{R}{c_0}.
\end{aligned} \tag{4.70}$$

The procedure is easily extended to initial and boundary conditions, as well as to the numerical flux.

When including the fictitious sources encountered in Section 4.2, the equations become

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\mu_r \frac{\partial \mathbf{H}}{\partial t} + \tilde{\mathbf{F}}, \\
\nabla \times \mathbf{H} &= \epsilon_r \frac{\partial \mathbf{E}}{\partial t} + qn\mu_n \mathbf{E} + qD_n \boldsymbol{\alpha} + qp\mu_p \mathbf{E} - qD_p \boldsymbol{\beta} + \tilde{\mathbf{A}} + \tilde{\mathbf{L}}, \\
\frac{\partial n}{\partial t} &= \nabla \cdot \left(n\mu_n \mathbf{E} + D_n \boldsymbol{\alpha} \right) + G - R + \tilde{\mathbf{B}}, \\
\frac{\partial p}{\partial t} &= -\nabla \cdot \left(p\mu_p \mathbf{E} - D_p \boldsymbol{\beta} \right) + G - R + \tilde{\mathbf{C}}, \\
\boldsymbol{\alpha} &= \nabla n, \\
\boldsymbol{\beta} &= \nabla p,
\end{aligned} \tag{4.71}$$

where

$$\begin{aligned}
\tilde{\mathbf{A}} &= \tilde{\mathbf{A}}(q, \mu_n, \mu_p, D_n, D_p) \\
\tilde{\mathbf{B}} &= \tilde{\mathbf{B}}(\mu_n, D_n) \\
\tilde{\mathbf{C}} &= \tilde{\mathbf{C}}(\mu_p, D_p) \\
\tilde{\mathbf{F}} &= \tilde{\mathbf{F}}, \text{ as } \mu_0 = \frac{\zeta_0}{c_0}, \\
\tilde{\mathbf{L}} &= \zeta_0 \tilde{\mathbf{L}}, \text{ as } \epsilon_0 \zeta_0 = \frac{1}{c_0},
\end{aligned} \tag{4.72}$$

and the manufactured solution becomes $\tilde{\mathbf{H}} = \zeta_0 \tilde{\mathbf{H}}$.
Moving to the Lorentz model (2.29), the introduction of t and $\mathbf{J}_l$ yields

$$\begin{aligned}
\zeta_0 c_0 \frac{\partial \mathbf{P}}{\partial t} &= \mathbf{J}_l, \\
\frac{\partial \mathbf{J}_l}{\partial t} &= \frac{\zeta_0}{c_0} \epsilon_0 \omega_p^2 \mathbf{E} - \frac{\gamma}{c_0} \mathbf{J}_l - \frac{\zeta_0}{c_0} \omega_r^2 \mathbf{P},
\end{aligned} \tag{4.73}$$

and therefore

$$\begin{aligned}
\frac{\partial \mathbf{P}}{\partial t} &= \mathbf{J}_l, \\
\frac{\partial \mathbf{J}_l}{\partial t} &= \omega_p^2 \mathbf{E} - \gamma \mathbf{J}_l - \omega_r^2 \mathbf{P},
\end{aligned} \tag{4.74}$$

with

$$\mathbf{P} = \zeta_0 c_0 \mathbf{P}, \qquad \omega = \frac{\omega}{c_0}, \qquad \gamma = \frac{\gamma}{c_0}, \tag{4.75}$$

which transform the manufactured solution into $\tilde{\mathbf{P}} = \zeta_0 c_0 \tilde{\mathbf{P}}$ and $\tilde{\mathbf{J}}_l = \zeta_0 \tilde{\mathbf{J}}_l$.
The real-time generation function can be written in the new terms as

$$G = \frac{\gamma}{\epsilon_0 \omega_p^2} \frac{||\mathbf{J}_l||^2}{h\nu_0} = \frac{\gamma}{\epsilon_0 \omega_p^2} \frac{||\mathbf{J}_l||^2}{h\nu_0} \frac{1}{c_0 \zeta_0^2} = \frac{\gamma}{\omega_p^2} \frac{||\mathbf{J}_l||^2}{h\nu_0} \frac{1}{\zeta_0}, \tag{4.76}$$

having used the identity $\zeta_0 c_0 = \epsilon_0^{-1}$ in the last equality. Then, the expression of G is

$$G = \frac{G}{c_0} = \frac{\gamma}{\omega_p^2} \frac{||\mathbf{J}_l||^2}{h\nu_0} \frac{1}{c_0 \zeta_0}. \tag{4.77}$$

The expression of the test generation function is simply:

$$G = \frac{G}{c_0} = \frac{\alpha}{2\zeta_0 c_0} \frac{||\mathbf{E}||^2}{h\nu_0}. \tag{4.78}$$

A comparison between the old and the new units is shown in Table 4.1.

Quantity	Original units	Implemented units
t	s	m
ω	rad s^{-1}	rad m^{-1}
γ	s^{-1}	m^{-1}
q	C	V m
E	V/m	V/m
H	A/m	V/m
J	A m^{-2}	V m^{-2}
P	C m^{-2}	V m
n, p	m^{-3}	m^{-3}
μ_n, μ_p	m^2 V^{-1} s^{-1}	m V^{-1}
D_n, D_p	m^2 s^{-1}	m
R, G	m^{-3} s^{-1}	m^{-4}

Table 4.1: Original and implemented measurement units

4.4 Simulation results

In this section we present simulation results for three milestones identified in the verification process illustrated in Section 4.2:

- the MDD model (4.28)-(4.29) of Section 4.2.1;
- the Maxwell-Lorentz model of Section 4.2.4;
- the MDD model of Section 4.2.3.

The choices stem from the following considerations.

At the implementation stage, convergence analyses have been run on each test case presented in Sections 4.2.1 and 4.2.2. Out of all such cases, the MDD model (4.28)-(4.29) of Section 4.2.1 can be considered the milestone, as it represents the turning point between individual tests on elementary building blocks (Maxwell's equations, advection, diffusion) and global ones, on the assembled system. For this reason, the convergence results obtained in its verification are representative of those emerged from the elementary bricks, and are proposed in this section.

A similar analysis has been carried out on the Maxwell-Lorentz model of Section 4.2.4, whose relevance is not only due to the introduction of dispersion but also to the different structure of the electromagnetic problem – the MDD model (4.28) entails low-frequency stationary waves in a bounded domain, while in (4.60) the field is a plane wave oscillating at optical frequencies and propagating in an artificially truncated domain (ABCs are applied).

To the purpose of convergence analysis, uniform structured meshes are considered and multiple simulations for decreasing values of the mesh size h (and a fixed simulation time $[0, t_f]$) are performed. With reference to a particular unknown, for each point m in the sequence of meshes, the evolution of the L^2 error $\mathfrak{E}^{(m)}$ with respect to the exact solution is recorded. Then, between two

successive simulations m and $m+1$ the numerical order of convergence is evaluated as:

$$\text{Order}_{\frac{m+1}{m}} = \frac{\log\left(\dfrac{\max\limits_{t\in[0,t_f]} \mathfrak{E}^{(m)}}{\max\limits_{t\in[0,t_f]} \mathfrak{E}^{(m+1)}}\right)}{\log\left(\dfrac{h_m}{h_{m+1}}\right)}.$$

The time step Δt is a function of h through the condition (3.102) in the MDD case and through the classical CFL condition (3.97) in the Maxwell-Lorentz case.

To conclude, we highlight the relevance of the MDD system of Section 4.2.3. With its physically meaningful time/space scales and model parameters (up to attenuation factors for generation and DC electron/hole distributions), aiming at reproducing a scattering problem at optical frequencies, it has paved the way for actual device simulation. In the proposed results, the evolution of the numerical solution is observed throughout the simulation time and compare it to its exact counterpart; also, the L^2 norm of each unknown is monitored to check stability.

4.4.1 Maxwell-Drift-Diffusion at low frequencies

The simulation results related to the model (4.28) of Section 4.2.2.3 are presented below.

This test case does not aim at reproducing a physical meaningful picture, hence it is assumed that $\mu_n = \mu_p = 10$, $D_n = D_p = 10V_T$; similarly, recombination parameters are simply $\tau_n = \tau_p = n_1 = p_1 = C_{A_n} = C_{A_p} = 10$. In this setting the two drift-diffusion equations are identical up to a sign – for example, we restrict the attention to electrons.

The computational domain $\Omega = [0,1]^2$ depicted in Figure 4.5 is discretized with a uniform structured mesh with M points per side and simulations are run for $M = 21, 31, 41, 51, 61$. For each value, linear and quadratic interpolation are tested.

An explicit RK2 scheme is considered with the stability condition given by (3.102), which fixes the maximum stable time step Δt as a function of h – stable choices of the constants are $C_p^a = 0.26$ and $C_p^d = 0.04$ for linear interpolation, $C_p^a = 0.1$ and $C_p^d = 0.01$ for quadratic interpolation.

Following Section 3.1.1.5, we take the fully upwind flux (3.33) for Maxwell's equations, the local Lax-Friedrichs flux defined in (3.49)-(3.50)-(3.51) for drift currents, and the alternate flux (3.43) for the diffusive part of the equations with $b = 1$ and $\mathbf{w} = [1, 1]$.

It is important in the simulations to keep track of the electromagnetic energy

$$\mathfrak{W}_{EM} = \int_{\Omega_e} \epsilon ||\mathbf{E}||^2 + \mu ||\mathbf{H}||^2, \tag{4.79}$$

and the norms

$$\mathfrak{W}_n = ||n||_{L^2(\Omega)}, \qquad \mathfrak{W}_p = ||p||_{L^2(\Omega)}. \tag{4.80}$$

Indeed, since the exact solution to Maxwell's equations (4.13) is a mode of a cavity resonator, a stable simulation requires the electromagnetic energy (4.79) to be constant. On the other hand,

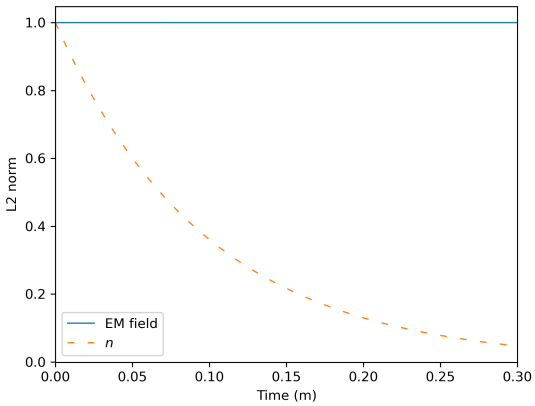


Figure 4.9: Evolution of $\mathfrak{W}_{EM}$ and $\mathfrak{W}_n$, normalized to their peak values.

the exact electron concentration is (4.23) and hence stability entails (4.80) to decrease with time as an effect of diffusion. The two behaviors are compared in Figure 4.9.[10]

The numerical solution is monitored by means of a probe positioned at the center of Ω; we track $\boldsymbol{\alpha}$ and n compared to their exact counterparts (Figure 4.10 and 4.11).

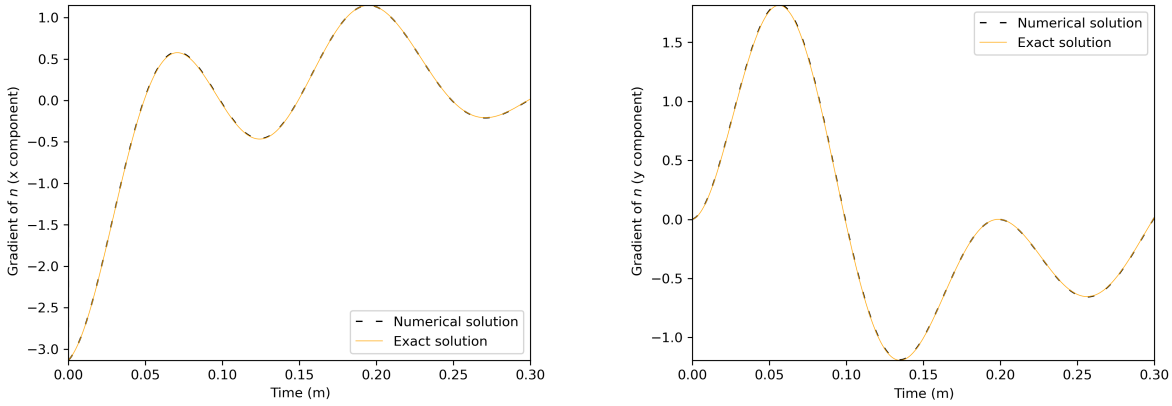


Figure 4.10: The evolution of $\boldsymbol{\alpha} \cdot \hat{\mathbf{x}}$, $\boldsymbol{\alpha} \cdot \hat{\mathbf{y}}$, at the center of the computational domain. The plots are extracted from the case of linear interpolation and $M = 41$.

[10]Note that time is multiplied by the speed of light in vacuum c_0 and hence expressed in meters the plots. This stems from the particular implementation choice outlined in Section 4.3.

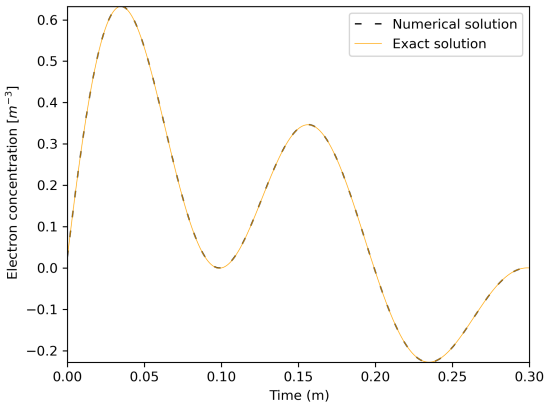


Figure 4.11: The evolution of n, at the center of the computational domain. The plot is extracted from the case of linear interpolation and $M = 41$.

Convergence results are reported below, with focus on the electromagnetic field and electron concentration. In particular, the following errors are monitored:

$$\begin{aligned} \mathfrak{E}_n &= ||n - \tilde{n}||_{L^2(\Omega)}, \\ \mathfrak{E}_{EM} &= ||\mathbf{E} - \tilde{\mathbf{E}}||_{L^2(\Omega)} + ||\mathbf{H} - \tilde{\mathbf{H}}||_{L^2(\Omega)}. \end{aligned} \tag{4.81}$$

We notice from the results that for a polynomial interpolation of order $\mathfrak{p}$, the order of convergence is $\mathfrak{p} + 1$ for the electromagnetic field and $\mathfrak{p}$ for electron concentration.

RK2, $\mathcal{P}_1$:

h	Max($\mathcal{L}_2$ error)	Numerical order
5.7×10^{-8}	0.00268289	1.1016126632202712
3.8×10^{-8}	0.0017164	1.069410316091575
2.8×10^{-8}	0.00126185	1.045646079024909
2.3×10^{-8}	0.00099925	1.036853269407647
1.9×10^{-8}	0.000827132	–

Table 4.2: Convergence analysis on electron concentration – RK2, $\mathcal{P}_1$

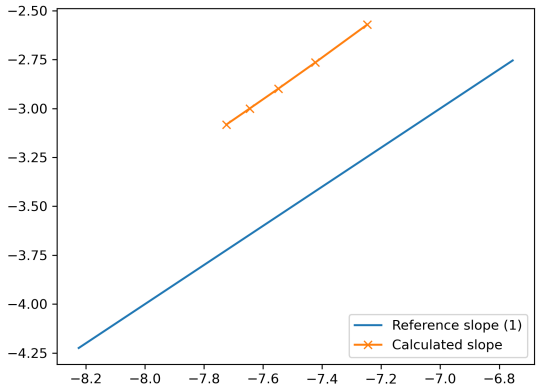


Figure 4.12: Convergence rate of $\mathfrak{E}_n$ – RK2, $\mathcal{P}_1$

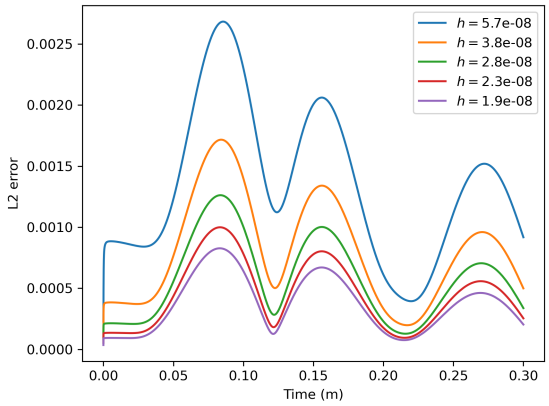


Figure 4.13: Evolution of $\mathfrak{E}_n$ in the considered meshes – RK2, $\mathcal{P}_1$

h	Max($\mathcal{L}_2$ error)	Numerical order
5.7×10^{-8}	0.00208733	2.0206584020586824
3.8×10^{-8}	0.000919964	2.0132566203923155
2.8×10^{-8}	0.00051551	2.0096462185890736
2.3×10^{-8}	0.000329217	2.0076563630666735
1.9×10^{-8}	0.000228304	–

Table 4.3: Convergence analysis on Maxwell's equations – RK2, $\mathcal{P}_1$

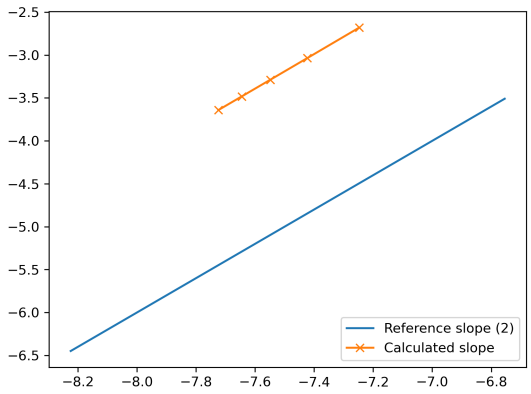


Figure 4.14: Convergence rate of $\mathfrak{E}_{EM}$ – RK2, $\mathcal{P}_1$

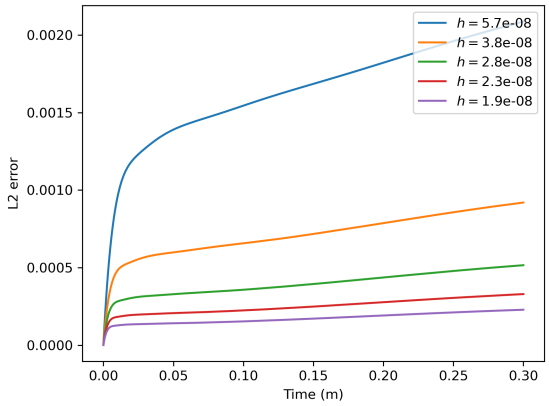


Figure 4.15: Evolution of $\mathfrak{E}_{EM}$ in the considered meshes – RK2, $\mathcal{P}_1$

h	Max($\mathcal{L}_2$ error)	Numerical order
5.7×10^{-8}	2.45709×10^{5}	3.0269863702700857
3.8×10^{-8}	7.20104×10^{6}	3.022302086719277
2.8×10^{-8}	3.01851×10^{-6}	3.01867558003563
2.3×10^{-8}	1.53905×10^{-6}	3.0160466582968346
1.9×10^{-8}	8.88052×10^{-7}	–

Table 4.4: Convergence analysis on electron concentration – RK2, $\mathcal{P}_2$

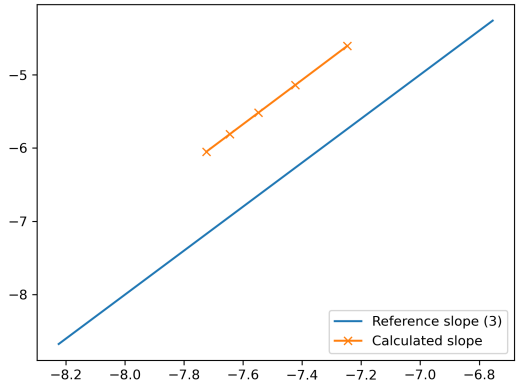


Figure 4.16: Convergence rate of $\mathfrak{E}_n$ – RK2, $\mathcal{P}_2$

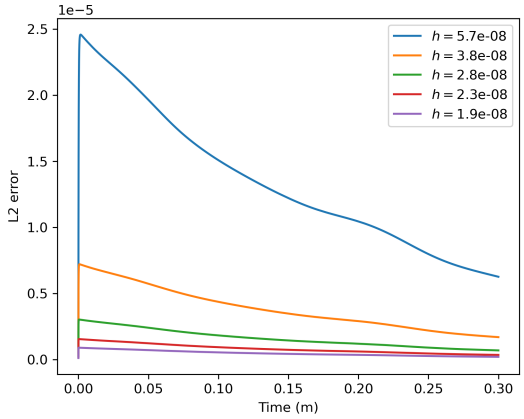


Figure 4.17: Evolution of $\mathfrak{E}_n$ in the considered meshes – RK2, $\mathcal{P}_2$

h	Max($\mathcal{L}_2$ error)	Numerical order
5.7×10^{-8}	5.12525×10^{5}	3.006671921504037
3.8×10^{-8}	1.51449×10^{5}	3.0052456859603374
2.8×10^{-8}	6.37962×10^{-6}	3.004303857751354
2.3×10^{-8}	3.26323×10^{-6}	3.003670332210577
1.9×10^{-8}	1.88718×10^{-6}	–

Table 4.5: Convergence analysis on Maxwell's equations – RK2, $\mathcal{P}_2$

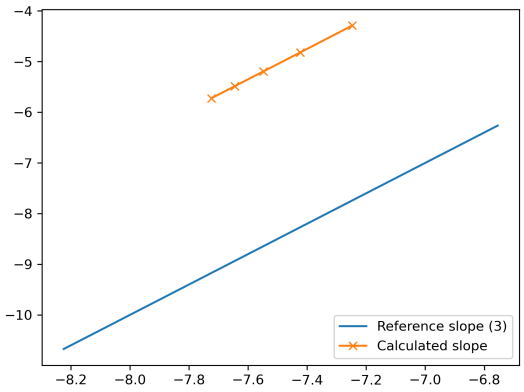


Figure 4.18: Convergence rate of $\mathfrak{E}_{EM}$ – RK2, $\mathcal{P}_2$

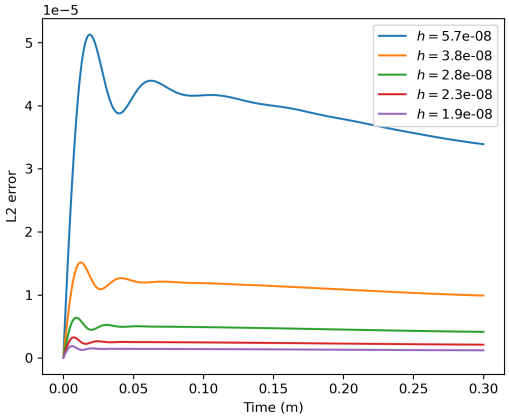


Figure 4.19: Evolution of $\mathfrak{E}_{EM}$ in the considered meshes – RK2, $\mathcal{P}_2$

4.4.2 Maxwell-Lorentz at optical frequencies

We now present the results obtained in the setting of Section 4.2.4.
The physical parameters of the dispersion model are set to $\omega_p = \gamma = \omega_r = 1$, whereas the field is based on the parameters (4.44)-(4.45).
The dispersive computational domain $\Omega = [-\lambda, \lambda]^2$, $\lambda = 800$ nm, terminated with ABCs, is discretized with a uniform structured mesh. Convergence analysis is based on four numerical settings, namely $M = 21, 41, 61, 81$, where M denotes the number of points per side of the mesh.
Time is discretized with the explicit schemes LSRK2 with linear and quadratic interpolation, and LSRK45 with quadratic interpolation. The fully upwind flux (3.33) is considered.

At each time step, the following relative errors are recorded:

$$\begin{aligned} \mathfrak{E}_{J,rel} &= \frac{||\mathbf{J}_l - \tilde{\mathbf{J}}_l||_{L^2(\Omega)}}{||\tilde{\mathbf{J}}_l||_{L^2(\Omega)}}, \\ \mathfrak{E}_{EM,rel} &= \frac{||\mathbf{E} - \tilde{\mathbf{E}}||_{L^2(\Omega)}}{||\tilde{\mathbf{E}}||_{L^2(\Omega)}} + \frac{||\mathbf{H} - \tilde{\mathbf{H}}||_{L^2(\Omega)}}{||\tilde{\mathbf{H}}||_{L^2(\Omega)}}. \end{aligned} \tag{4.82}$$

The analysis that follows show that, with the LSRK2 scheme, the asymptotic order of convergence equal is 2 for the electromagnetic field in both the linear and the quadratic case. The convergence order on the polarization current, on the other hand, seems to be 3 in the quadratic case and higher than 2 in the linear one. In the LSRK45 case, the order is 3 for both the Maxwell and the Lorentz part of the model.

It is interesting to compare these data to the analysis done in [LSV17], in which the DGTD method with the fully upwind flux (3.33) is applied to the Maxwell-Lorentz model in a perfectly electric conducting cavity, with RKn time integration and polynomial interpolation of order $\mathfrak{p}$.
The scheme is shown to converge as $\mathcal{O}(dt^n + h^{\mathfrak{p}+1})$, so that, by virtue of the CFL condition (3.97), the asymptotic order is $\min(n, \mathfrak{p} + 1)$.
The present scenario, in which wave propagation and ABCs are considered, had not been investigated so far, and yields results that are consistent with [LSV17].

LSRK2, $\mathcal{P}_1$:

h	$\max \mathfrak{E}_{EM,rel}$	Numerical order
2.8×10^{-8}	0.000127389	1.8695707687004095
1.9×10^{-8}	5.96921×10^{-5}	1.8893416888236159
1.4×10^{-8}	3.46629×10^{-5}	1.9109077783955253
1.1×10^{-8}	2.6297×10^{-5}	1.9258384015612429
9.4×10^{-9}	1.5929×10^{-5}	–

Table 4.6: Convergence analysis on Maxwell's equations – LSRK2, $\mathcal{P}_1$

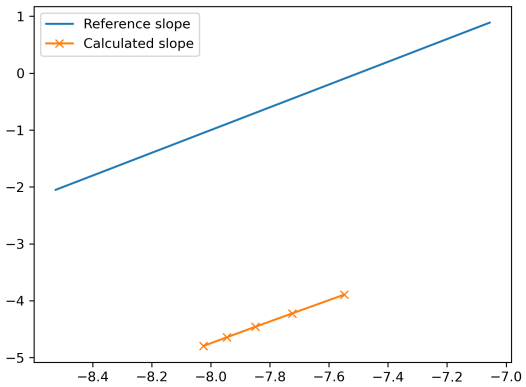


Figure 4.20: Convergence rate of $\mathfrak{E}_{EM,rel}$ – LSRK2, $\mathcal{P}_1$

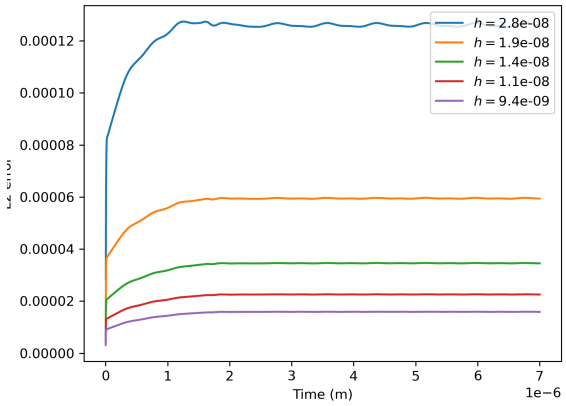


Figure 4.21: Evolution of $\mathfrak{E}_{EM,rel}$ in the considered meshes – LSRK2, $\mathcal{P}_1$

h	$\max \mathfrak{E}_{J,rel}$	Numerical order
2.8×10^{-8}	0.0466027	2.8241991617829427
1.9×10^{-8}	0.0148284	2.6839112251849624
1.4×10^{-8}	0.00685125	2.5575290089326312
1.1×10^{-8}	0.00387186	2.4522538220525743
9.4×10^{-9}	0.00247598	–

Table 4.7: Convergence analysis on the Lorentz model – LSRK2, $\mathcal{P}_1$

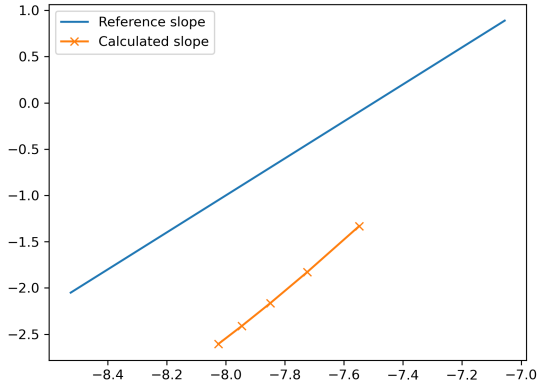


Figure 4.22: Convergence rate of $\mathfrak{E}_{J,rel}$ – LSRK2, $\mathcal{P}_1$

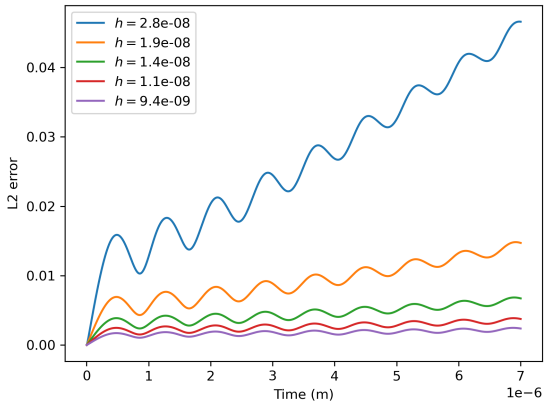


Figure 4.23: Evolution of $\mathfrak{E}_{J,rel}$ in the considered meshes – LSRK2, $\mathcal{P}_1$

LSRK2, $\mathcal{P}_2$:

h	$\max \mathfrak{E}_{EM,rel}$	Numerical order
2.8×10^{-8}	7.91896×10^{-6}	2.322072325595617
1.9×10^{-8}	3.08867×10^{-6}	2.1955647319209857
1.4×10^{-8}	1.64233×10^{-6}	2.135351083719648
1.1×10^{-8}	1.01982×10^{-6}	2.104751411350037
9.4×10^{-9}	6.94811×10^{-7}	–

Table 4.8: Convergence analysis on the Maxwell's equations – LSRK2, $\mathcal{P}_2$

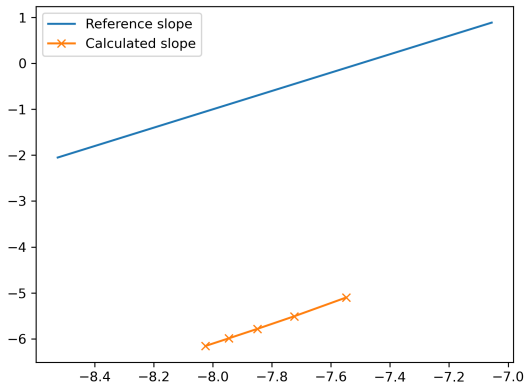


Figure 4.24: Convergence rate of $\mathfrak{E}_{EM,rel}$ – LSRK2, $\mathcal{P}_2$

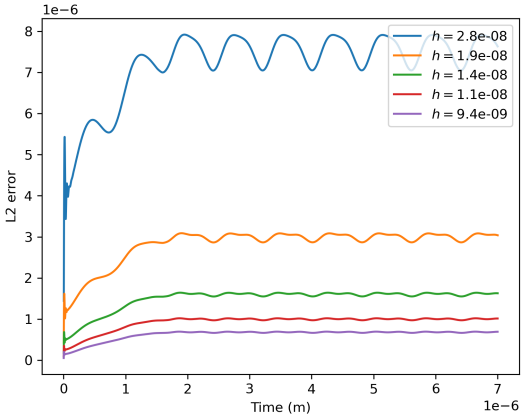


Figure 4.25: Evolution of $\mathfrak{E}_{EM,rel}$ in the considered meshes – LSRK2, $\mathcal{P}_2$

h	Max($\mathcal{L}_2$ error)	Numerical order
2.8×10^{-8}	0.00879235	2.987554564558026
1.9×10^{-8}	0.00261832	2.989517032342913
1.4×10^{-8}	0.00110794	2.988341127458159
1.1×10^{-8}	0.000568743	2.9851169149802104
9.4×10^{-9}	0.000330028	–

Table 4.9: Convergence analysis on the Lorentz model – LSRK2, $\mathcal{P}_2$

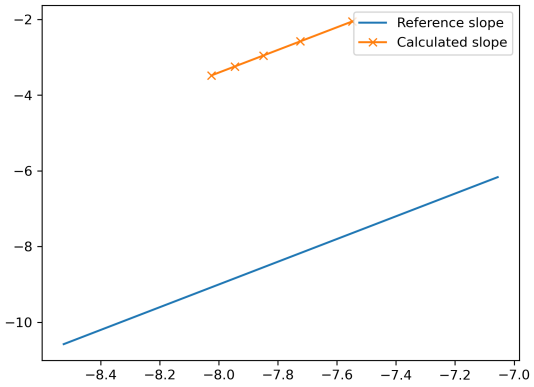


Figure 4.26: Convergence rate of $\mathfrak{E}_{J,rel}$ – LSRK2, $\mathcal{P}_2$

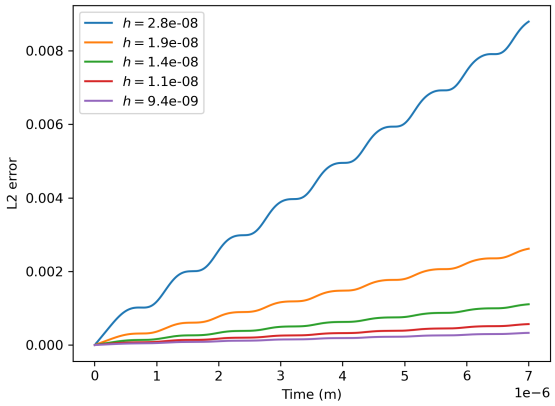


Figure 4.27: Evolution of $\mathfrak{E}_{J,rel}$ in the considered meshes – LSRK2, $\mathcal{P}_2$

LSRK45, $\mathcal{P}_2$:

h	$\max \mathfrak{E}_{EM,rel}$	Numerical order
2.8×10^{-8}	5.6012×10^{-6}	2.987187162611704
1.9×10^{-8}	1.62624×10^{-6}	2.990932444763642
1.4×10^{-8}	6.87862×10^{-7}	2.992299884460887
1.1×10^{-8}	3.52791×10^{-7}	2.9929781575903105
9.4×10^{-9}	2.04423×10^{-7}	–

Table 4.10: Convergence analysis on the Maxwell's equations – LSRK45, $\mathcal{P}_2$

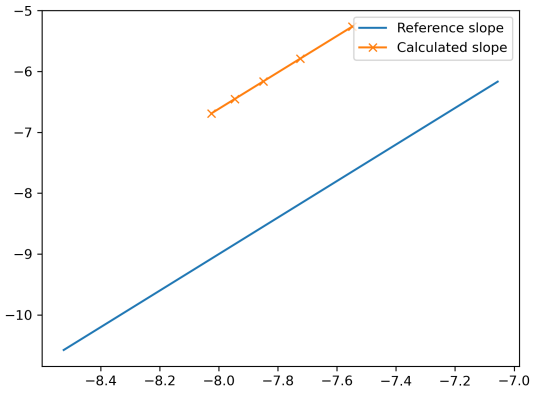


Figure 4.28: Convergence rate of $\mathfrak{E}_{EM,rel}$ – LSRK45, $\mathcal{P}_2$

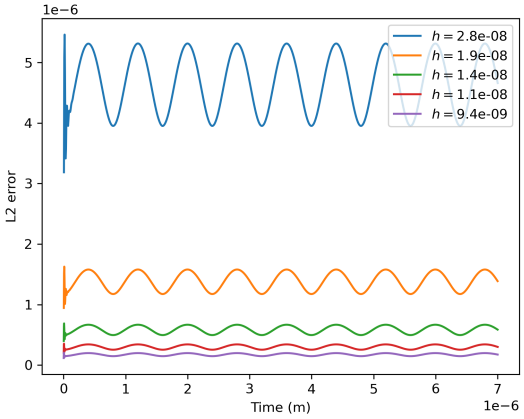


Figure 4.29: Evolution of $\mathfrak{E}_{EM,rel}$ in the considered meshes – LSRK45, $\mathcal{P}_2$

h	$\max \mathfrak{E}_{J,rel}$	Numerical order
2.8×10^{-8}	0.00878637	2.990987182986827
1.9×10^{-8}	0.0026129	2.994445567216772
1.4×10^{-8}	0.00110408	2.99576830965967
1.1×10^{-8}	0.000565823	2.997602968109476
9.4×10^{-9}	2.04423×10^{-7}	–

Table 4.11: Convergence analysis on the Lorentz model – LSRK45, $\mathcal{P}_2$

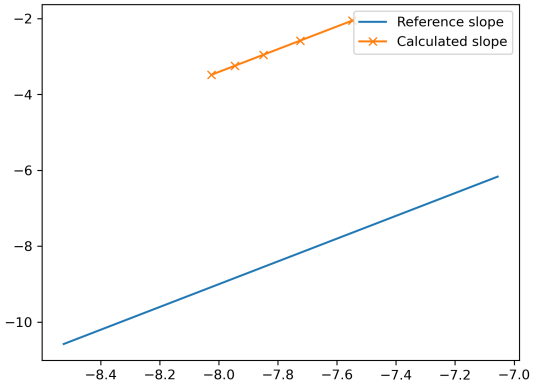


Figure 4.30: Convergence rate of $\mathfrak{E}_{J,rel}$ – LSRK45, $\mathcal{P}_2$

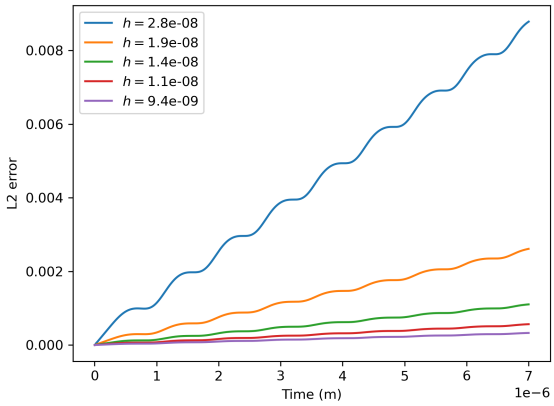


Figure 4.31: Evolution of $\mathfrak{E}_{J,rel}$ in the considered meshes – LSRK45, $\mathcal{P}_2$

We conclude the analysis with a remark on the behavior of $\mathfrak{E}_{J,rel}$, which appears to indefinitely grow with time, whereas $\mathcal{E}_{EM,rel}$ has a periodic behavior. This difference allows to take up the discussion from Section 4.2.4 on the influence of $\Delta\mathbf{E}$ on stability. To appreciate it, we propose a further convergence analysis, for example with LSRK2 and linear interpolation, in the ideal case $\Delta\mathbf{E} = \mathbf{0}$, i.e. when $\mathbf{E} = \tilde{\mathbf{E}}$ is enforced. With reference to Figure 4.32, we notice that relative error on $\mathbf{J}_l$ becomes periodic. This confirms that its increasing character in the non-ideal case is a byproduct of the numerical error on $\mathbf{E}$.

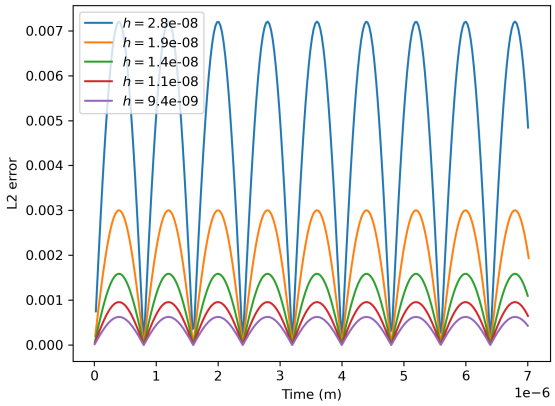


Figure 4.32: Evolution of the relative error $\mathfrak{E}_{J,rel}$ for increasing mesh refinement in the ideal case $\mathbf{E} = \tilde{\mathbf{E}}$.

4.4.3 Maxwell-Drift-Diffusion at optical frequencies

With reference to the test case of Section 4.2.3 and its physically meaningful setting, we report the results of a simulation of the MDD system with recombination, generation (4.47), DC components. Suitable attenuation factors are considered, namely 10^{-19} for $G - \tilde{G}$, 10^{-10} for n_{DC} and p_{DC}.
The computational domain is the one in Figure 4.7, where $h \in [0.05, 0.1]$ μm. An explicit RK2 scheme is considered, with two different time steps for Maxwell's equations and the semiconductor ones according to the algorithm mentioned in Section (3.1.4.4). Here, in particular $\Delta t_{EM} = 3.3 \times 10^{-5}$ ps has been calculated with the aid of (3.102); then, $\Delta t_{np} = 4\Delta t_{EM}$ has been determined empirically. Space discretization is based on linear Lagrange polynomials.
Following Section 3.1.1.5, the fully upwind flux (3.33) is selected for Maxwell's equations, the local Lax-Friedrichs flux defined in (3.49)-(3.50)-(3.51) for drift currents, and the alternate flux (3.43) for the diffusive part of the equations, with $b = 1$ and $\mathbf{w} = [1, 1]$.

Similarly to Section 4.4.1 a first stability check consists in keeping track of $\mathfrak{W}_{EM}$ and $\mathfrak{W}_n$ defined in (4.79) and (4.80) respectively. Referring to the plots in Figure 4.33, we notice that $\mathfrak{W}_{EM}$ displays the pulsed behavior of $\tilde{\mathbf{E}}, \tilde{\mathbf{H}}$ as expected. Also, p evolves slowly compared to n, as a result of the realistic physical setting in which $D_n = 20D_p$.

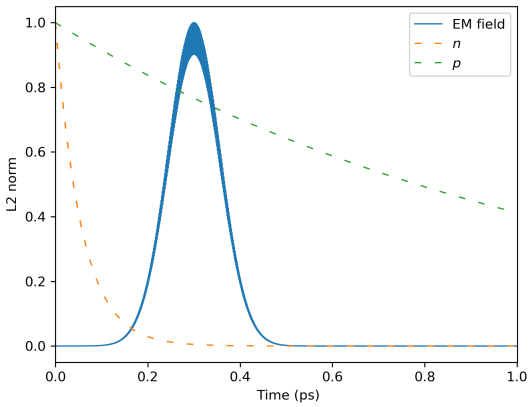


Figure 4.33: Evolution of the electromagnetic energy $\mathfrak{W}_{EM}$ stored in the computational domain along with the L^2 norms of n and p. The three quantities are normalized to their respective peaks.

The numerical solution is then observed by means of a probe placed at the center of the semiconductor. For example, $\boldsymbol{\alpha} \cdot \hat{\mathbf{x}}$, $\boldsymbol{\alpha} \cdot \hat{\mathbf{y}}$ and n are compared to their exact counterparts throughout the simulation (Figures 4.34 and 4.35).[11]

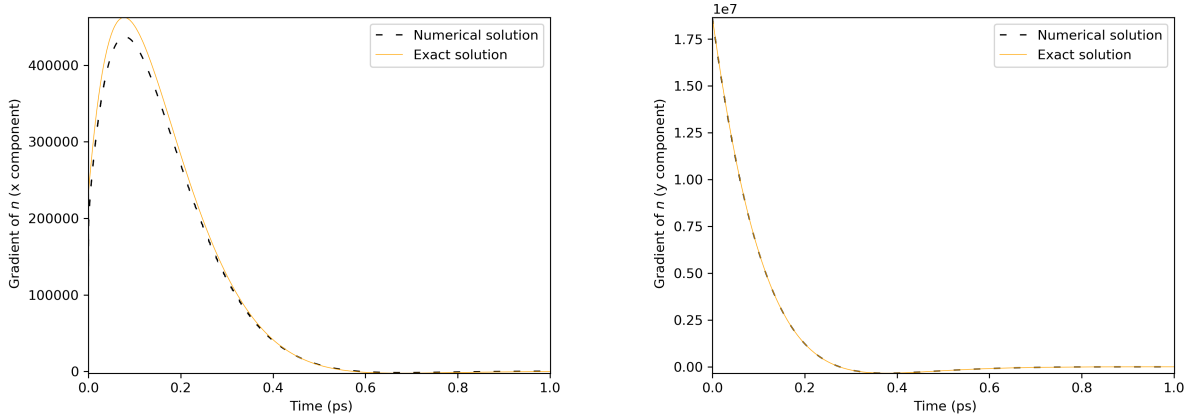


Figure 4.34: The evolution of $\boldsymbol{\alpha} \cdot \hat{\mathbf{x}}$, $\boldsymbol{\alpha} \cdot \hat{\mathbf{y}}$ at the center of the semiconductor, compared to the exact solution.

[11]The very large values of the components of the gradient stem from considering (4.46) with a,$b \sim 10^{-9}$ m. In retrospect, this is an important point to consider when coding the MDD equations, especially if this entails integrating a semiconductor module into an existing electromagnetic one. Indeed, the electromagnetic part can be smoothly tested at optical frequencies if space is measured in meters, whereas for the transport part microns would be preferable.

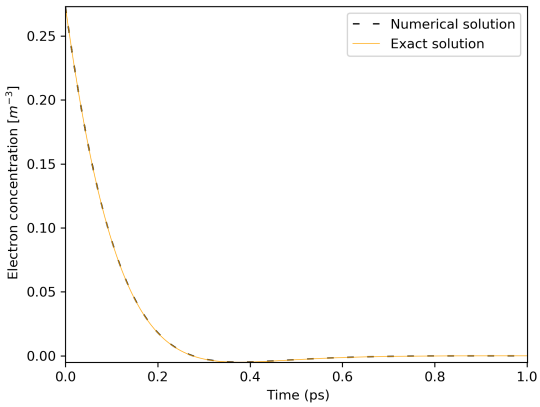


Figure 4.35: The evolution of n at the center of the semiconductor compared to the exact solution.

Moreover, the errors $\mathfrak{E}_n$ and $\mathfrak{E}_{EM}$ are plotted in Figures 4.36 and 4.37, and in the present analysis we wish to also highlight the behavior of

$$\mathfrak{E}_\alpha = ||\boldsymbol{\alpha} - \tilde{\boldsymbol{\alpha}}||_{L^2(\Omega)}, \tag{4.83}$$

which is plotted in Figure 4.38.
We notice that $\mathfrak{E}_{EM}$ inherits the pulsed behavior of the exact solution. $\mathfrak{E}_n$ initially grows but suggests settlement to a constant value. The convergence of the drift-diffusion part of the system is more complex.
$\mathfrak{E}_n$ starts at a very small value, it sharply increases to 1.3×10^{-10}, then increases to 3×10^{-10} at a much slower rate. $\mathfrak{E}_\alpha$ behaves the opposite way – it starts at about 0.13 and sharply reduces by 50%, to then very slowly relax towards constant value.

The two behaviors just depicted have been regularly found throughout the present work, at different space and time scales. In practice, $\boldsymbol{\alpha}$, which is calculated at the end of each iteration after n, seems to be much more sensitive to numerical error than n itself. This is not unreasonable if complemented by a further element: in test simulations on the sole drift-diffusion equations with $\mathbf{E} = \tilde{\mathbf{E}}$, it has been noticed that the diffusion part dominates in the CFL condition (3.102) and determines the maximum stable time step.

To conclude, it is pointed out that the monotone decreasing behavior of $\mathfrak{E}_\alpha$ observed in the present example has been identified, over the course of the present work, as a distinctive trait of convergence of the electron drift-diffusion equation. Moreover, all the above considerations can be extended to the hole equation.

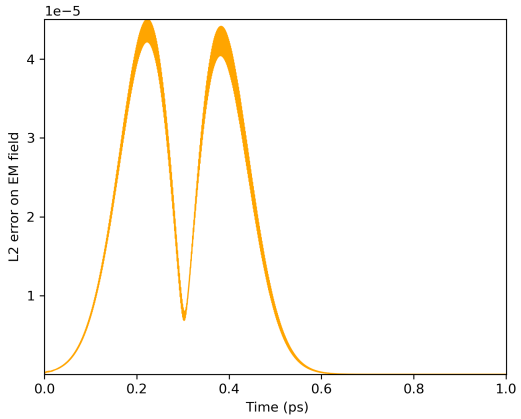


Figure 4.36: Evolution of the L^2 error on the electromagnetic field $\mathfrak{E}_{EM}$.

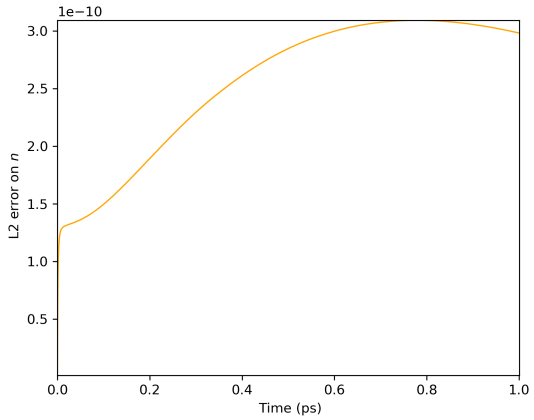


Figure 4.37: Evolution of the L^2 error on electron concentration $\mathfrak{E}_n$.

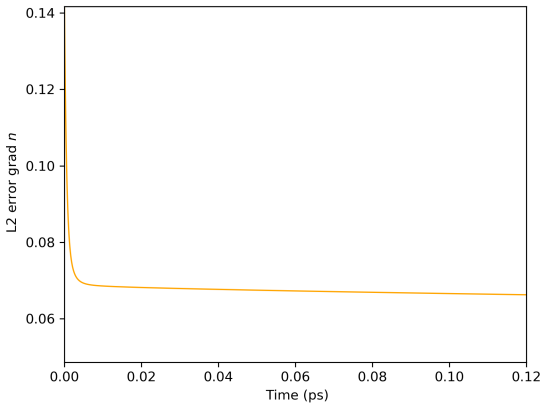


Figure 4.38: Evolution of the L^2 error on the gradient of electron concentration $\mathfrak{E}_\alpha$.

Chapter 5

Device simulation

Successful tests make the ground ready for real device simulation, up to a few preliminary considerations. Fictitious sources are removed, so that the scattered field is not canceled anymore. The practical implementation of (2.58-2.59) requires special attention – the topic is discussed in Section 5.1. The key task of calculating the DC operating point is confronted in Section 5.2. Section 5.3 introduces the devices considered in the work. The rest of the chapter is devoted to simulation results.

5.1 Field injection and absorption

The initial conditions from (2.58-2.59) require the incident field to not hit the semiconductor until some $t > 0$. When excitation is pulsed in time, it can be made practically null on Ω at $t = 0$ by retarding it of a suitable t_0 – an idea that can be recognized in the definition of $\tilde{\mathbf{E}}$ in Section 4.2.3. In the new setting with no manufactured solution or fictitious sources, a first-order Silver-Muller ABC is imposed on Γ_e:

$$\hat{\mathbf{n}} \times \mathbf{E} + \hat{\mathbf{n}} \times \zeta(\hat{\mathbf{n}} \times \mathbf{H}) = \hat{\mathbf{n}} \times \mathbf{E}^i + \hat{\mathbf{n}} \times \zeta(\hat{\mathbf{n}} \times \mathbf{H}^i) \quad \text{on } \Gamma_e \times [0, +\infty[. \tag{5.1}$$

Note the difference with (4.35) – here, (5.1) has the two-fold role of injecting the incident field and absorbing the scattered one. Since in all applications of our interest, vacuum surrounds an illuminated device, the impedance appearing in (5.1) should be $\zeta = \zeta_0 = \sqrt{\mu_0/\epsilon_0}$.

It can be shown that such a domain truncation technique offers practically perfect absorption for normal incidence ($|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}| = 1$) while sensible performance reduction is observed when $|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|$ falls away from 1 [AAP+10]. As a compensation strategy, a sufficient distance must be allowed between the absorbing boundary and the scatterer. Other techniques such as Perfectly Matched Layers (PMLs) are more attractive in terms of absorption performance and hence computational domain size, but the price to pay is an increase in implementation effort and computational cost [San07].

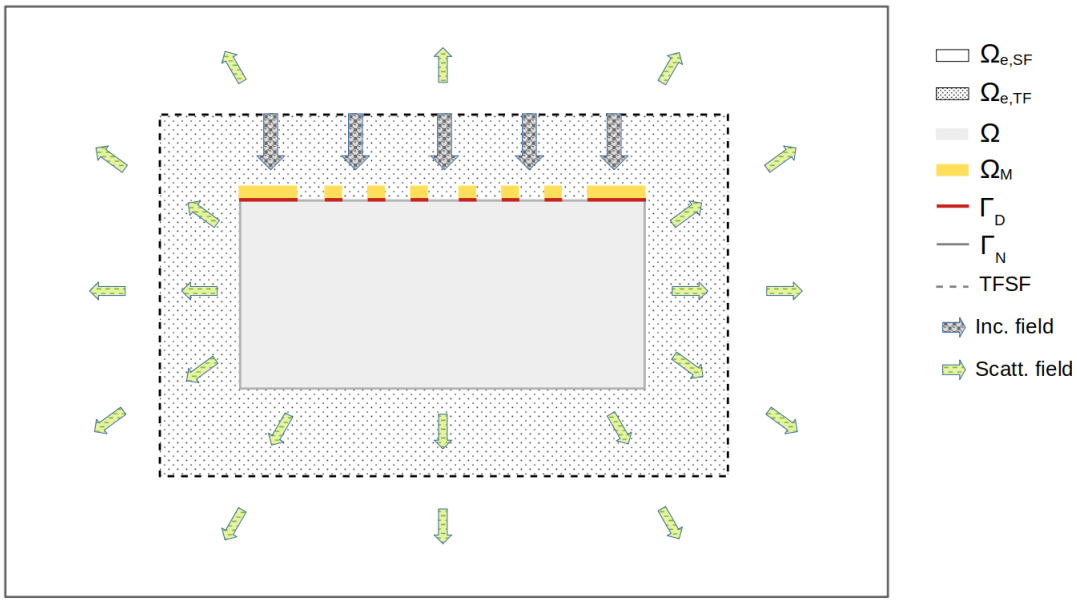


Figure 5.1: An illustration of the TFSF mechanism. Two disjoint regions of the computational domain are defined: $\Omega_{e,TF}$, in which both the incident and the scattered field exist; $\Omega_{e,SF}$, in which only the scattered field is present. From the interface between the two, called the *TFSF surface*, the incident field is injected and reaches the scatterer, which is contained in $\Omega_{e,TF}$. The resulting scattered field travels in both $\Omega_{e,TF}$ and $\Omega_{e,SF}$, whereas the incident field exists in the former only.

In the presence of PMLs, the computational domain can still be terminated with ABCs, but these cannot be used anymore to inject the incident field - it would be significantly attenuated before getting to the scatterer - and the task is accomplished with a *Total-Field-Scattered-Field* (TFSF) apparatus (5.1) as follows. A *Total-Field* (TF) region encloses the scatterer and the incident field is injected from its boundary, called the *TFSF surface*; outside, in the *Scattered-Field* (SF) region, only the scattered field exists, as the name suggests. With respect to the approach in Figure 4.6, in which the whole Ω_e can be considered TF, some computational cost is saved as the distance traveled by the injected wave to reach the semiconductor is reduced. Aside from that and more importantly, it becomes easier to visualize the scattered field around the device and construct near-to-far-field contours [Sil].
The possibility of adding PMLs to the two-dimensional code has not been considered; nonetheless, TF and SF regions have been implemented, having in mind either future PML implementation or the integration of the MDD solver into the DIOGENeS suite.
In the context of DG methods, suitable modification of the numerical flux is required on the faces composing the TFSF surface:

$$\begin{aligned} \mathbf{E}_* &= \mathcal{L}_E(\mathbf{E}^-, \mathbf{H}^-, \mathbf{E}^+, \mathbf{H}^+), \\ \mathbf{H}_* &= \mathcal{L}_H(\mathbf{E}^-, \mathbf{H}^-, \mathbf{E}^+, \mathbf{H}^+), \end{aligned} \tag{5.2}$$

where $-$ $(+)$ denotes the internal (external) page of a face and $\mathcal{L}_E, \mathcal{L}_H$ are for example the operators in (3.33). With reference to the interface a_{il} between K_i (TF) and K_l (SF), one operates as follows.

- In the local formulation on K_l, the incident field is subtracted from i part of the flux:

$$\begin{aligned}\mathbf{E}_* &= \mathcal{L}_E(\mathbf{E}_l, \mathbf{H}_l, \mathbf{E}_i - \mathbf{E}^{inc}, \mathbf{H}_i - \mathbf{H}^{inc}),\\ \mathbf{H}_* &= \mathcal{L}_H(\mathbf{E}_l, \mathbf{H}_l, \mathbf{E}_i - \mathbf{E}^{inc}, \mathbf{H}_i - \mathbf{H}^{inc}).\end{aligned} \tag{5.3}$$

- In the local formulation on K_i, the incident field is added to the l part of the flux:

$$\begin{aligned}\mathbf{E}_* &= \mathcal{L}_E(\mathbf{E}_i, \mathbf{H}_i, \mathbf{E}_l + \mathbf{E}^{inc}, \mathbf{H}_l + \mathbf{H}^{inc}),\\ \mathbf{H}_* &= \mathcal{L}_H(\mathbf{E}_i, \mathbf{H}_i, \mathbf{E}_l + \mathbf{E}^{inc}, \mathbf{H}_l + \mathbf{H}^{inc}).\end{aligned} \tag{5.4}$$

The computational domain is still closed with (5.1), except that now the condition should be homogeneous as only the scattered field exists on Γ_e:

$$\hat{\mathbf{n}} \times \mathbf{E} + \hat{\mathbf{n}} \times \sqrt{\frac{\mu_0}{\epsilon_0}}(\hat{\mathbf{n}} \times \mathbf{H}) = \mathbf{0} \quad \text{on } \Gamma_e \times [0, +\infty[. \tag{5.5}$$

5.2 Calculating the DC operating point

The knowledge of the steady state of the semiconductor has been highlighted as a prerequisite to solving the MDD system, but so far no route to its acquisition has been mentioned. In fact, the PhD project was conceived with the idea of importing DC quantities from Silvaco Atlas.

5.2.1 Electrostatic simulation in Silvaco Atlas

Atlas is a two- and three-dimensional simulator of semiconductor devices, supporting DC, AC small-signal analysis and full time-dependency. Light-semiconductor interaction is simulated along the steps described in Section 2.3 – Maxwell's equations are solved on a Cartesian grid with the FDTD method, and the same grid is further decomposed in triangles to solve the PDD model with the Box Integration Method (BIM). A plethora of physical models is available for a more refined description of the semiconductor, regarding (among many others) mobility, recombination, doping and contacts.
The standard procedure to simulate a device relies on writing a sequence of statements in the DeckBuild environment, which acts as a common interface between different tools in the Silvaco TCAD suite. A sequence of interest is illustrated below.

- GO Atlas - Set Atlas as the solver.
- MESH - Control the refinement of the structured Cartesian mesh along x and y.
- REGION - Define dielectric and semiconductor regions, i.e. set boundaries, assign a tag and specify a material.
- ELECTRODE - Introduce conductors in the mesh.
- DOPING - Assign a doping profile to the semiconductor.

- CONTACT - Set contacts to the semiconductor, e.g. Ohmic or Schottky, implement electrically floating conductors.
- MODELS - Call physical models, e.g. SRH and Auger models for recombination.
- MOBILITY - Set a particular mobility model for each semiconductor region and assign its parameters, e.g. the Caughey-Thomas model for field-dependent mobility.
- MATERIAL - Assign user-defined parameters to the called physical models, e.g. the trap level and the average lifetimes in the SRH function. Create user-defined materials.
- METHOD - Set a non-linear iteration strategy: Gummel, Newton, and block-Newton.
- LOG - Create a file for one-dimensional plots (e.g. light versus current in a photo-diode).
- SOLVE - Start the solver with the selected strategy and with a particular input. Multiple SOLVE statements are typically required to calculate the target state of the PDD system. The idea is to sweep on each independent source (e.g. voltage applied at a terminal) and use the result of the previous SOLVE statement as an initial guess for the next one. Hence, a first SOLVE statement calculates the steady state at thermodynamic equilibrium; this is the initial guess for the next steady state in which a small increment has been applied to turn on an independent source; the new result becomes the initial guess for a new iteration algorithm in which the independent source is slightly higher; and so on.
- SAVE - Create a file for scalar and vector two-dimensional plots, e.g. electron/hole concentration profile inside the semiconductor, arrow plots of electron/hole drift/diffusion currents.
- EXTRACT - Extract specific quantities for post-processing, e.g. the values of $\mathbf{E}, n, p$ at mesh nodes, so that they can be imported into the DGTD solver and become $\mathbf{E}_{DC}, n_{DC}, p_{DC}$.
- TONYPLOT - Launch Tonyplot visualization tool.

The SAVE statement is meant to create a *structure* file (.str) containing all the necessary data to generate two-dimensional plots, including the mesh itself, in the following format.

```
v ATLAS
j 4 1 2 94 95
k 2 2 0.6721053776586514 1.42
c 1 -2.1 -0.08 0
c 2 -2.1 -0.07 0
c 3 -2.1 -0.06 0
...
e 1 2177 2176 0
e 2 2176 2205 0
e 3 2178 2177 0
...
r 2 8
x 2 63 65
w 0 0
b 557
```

```
b 558
b 559
...
t 1 3 1 30 2 2 -1024 -1024
t 2 3 30 31 2 3 1 57
t 3 3 2 31 3 4 -1024 2
...
s 31 513 115 149 71 72 69 70 100 103 120 121 106 111 211
  220 224 107 112 210 221 225 108 524 525 215 222 226 419 420 118 124
n 0 91 3 0.0000000000000000e+00 0.0000000000000000e+00
  0.0000000000000000e+00 0.0000000000000000e+00
  0 0 -5 0 0 0 0 5 0 0 0 0 5 0 0 0 0 0 0 0 0 0 0 0 0 300
n 1 91 3 0.0000000000000000e+00 0.0000000000000000e+00
  0.0000000000000000e+00 0.0000000000000000e+00
  0 0 -5 0 0 0 0 5 0 0 0 0 5 0 0 0 0 0 0 0 0 0 0 0 0 300
...
n 12208 8 2 1.2999999999999998e+16 1.2999999999999998e+16
  1.2999999999999998e+16 0.0000000000000000e+00 0 0 5.581255187932605
  1.3707348 -0.30449534 1.33648662 1.29999972e+16 -4.99990076 21.1657433
  -5.09617948 20.543068 0.000381355941 -4.99990076 3.10450511e-20
  -7.47487406e-21 3.0131735e-20 21.1657433 -5.09617948 20.543068
  21.1657433 -5.09617948 20.543068 3e-13 4e-13 0.00179348178 300
...
p 6 1000000 1001000 1004000 1000001 1001001 1004001
d -5 -5 -7.86706123e-05 5 5 7.86706123e-05
```

Not all lines are essential to our purposes. Lines beginning by *c*, *e*, *t* contain connectivity information – they define vertices, edges and triangles respectively. For each node, a number and three coordinates are stated; each edge is assigned a number and its two endpoints are specified in terms of vertex numbers; each triangle is characterized by an entity number, a region number, three vertices and neighbors (-1024 indicates the presence of a boundary edge). The patterns are highlighted below.

```
c number x y z
e number vertex1 vertex2 0
t number region vertex1 vertex2 vertex3 neighbor1 neighbor2 neighbor3
```

Lines beginning by n contain the nodal values of the unknowns of the PDD solver, to be manipulated to derive further observables (e.g. drift and diffusion currents); this happens when the *.str file is opened in TonyPlot and, if necessary, when executing the EXTRACT statement. The latter produces a list of the kind

```
-2.1 1.6940658945086e-17 0
-2.1 0.01 -0.0469182037
-2.1 0.02 -0.117004635
...
```

where each line contains the abscissa and the ordinate of a mesh node and the value of the selected observable at that node.

5.2.2 Importing the steady state into an unstructured mesh

In order to provide DC arrays to the DGTD code it is necessary to process the *.str file along with the extracted observables. The code supports standard (unstructured) *.mesh files, which look as follows.

```
MeshVersionFormatted 2
Dimension
3
Vertices
3373
              -5E-08                    1E-07                    0      21
               5E-08                    1E-07                    0      22
            -2.5E-07                    1E-07                    0      25
               ...
Edges
490
1 75 18
75 76 18
76 77 18
...
Triangles
6646
1464 1465 1115 1
1162 1394 745 1
965 1552 1401 1
...
```

Such information is translated into the following variables/arrays:

- `ns,na,nt` - number of vertices, edges, triangles;
- `coor(1:ns,1:2)` - array of vertex coordinates;
- `ar(1:na,1:2)` - array of edge endpoints;
- `logar(1:na)` - array of edge tags (used to implement boundary conditions);
- `nu(1:nt,1:3)` - array of triangle vertices;
- `logtr(1:nt)` - array of triangle tags (used to identify mesh physical regions).

The *.str file is processed as follows.

The first step consists in using the connectivity information in the .str file to create similar variables/arrays describing the structured mesh. We shall call these *primed* to distinguish them from those introduced earlier and created from the unstructured .mesh file.
In practice, the required new variables/arrays are:

- `nsP,ntP` - number of vertices and triangles in the semiconductor region;

- `coorP(1:nsP,1:2)` - array containing the coordinates of the vertices lying in the semiconductor region;
- `nuP(1:ntP,1:3)` - array containing the vertices of the triangles constituting the semiconductor region.

Note that their construction requires to read the *.str file selectively, in that entities lying outside the semiconductor region must not be counted/collected. In the filtering process, for each node saved into `coorP`, the original node number is stored as well (i.e. the number that the node has in the Atlas mesh).

As a second step, the solution to the PDD system extracted from the Atlas simulation is imported in the form of arrays of `nsP` elements referring to `coorP`. The tricky point in this process is to translate quantities calculated in the Cartesian mesh to suitable counterparts defined on the unstructured mesh used in the DG discretization.
The extracted data are based on the *.str file and hence need to be read selectively – lines referring to vertices lying outside the semiconductor region must be skipped. This is done with the aid of the map linking the imported node numbers `1,2,...,nsP` to the original ones.
We shall refer to the imported quantities as $\mathbf{E}'_{DC}, n'_{DC}, p'_{DC}$ to stress the fact that they "live" on the primed mesh and that they are not practically utilizable by the DGTD solver – not yet.

The third and last step interpolates $\mathbf{E}'_{DC}, n'_{DC}, p'_{DC}$ for a migration into the unstructured mesh. Given $u = E_{x,DC}, E_{y,DC}, n_{DC}, p_{DC}$ and $u' = E'_{x,DC}, E'_{y,DC}, n'_{DC}, p'_{DC}$, the ultimate goal is to calculate $u(x_d, y_d)$ for all degrees of freedom (x_d, y_d) in the unstructured mesh, based on the knowledge of the corresponding u'. This is done by the following algorithm, whose fundamental concept is illustrated in Figure 5.2.

- For each triangle `jt` in `nu`, locate the nodes $\{(x_j, y_j) \ : \ j = 1, 2, ..., p\}$ where p is the degree of interpolation of the local DG discretization.
- For each node (x_j, y_j) find the triangle `jtP` in `nuP` containing it, and retrieve the coordinates of the three vertices $(x'_1, y'_1), (x'_2, y'_2), (x'_3, y'_3)$ from `coorP`.
- Interpolate $u'(x'_1, y'_1), u'(x'_2, y'_2), u'(x'_3, y'_3)$ to calculate $u(x_j, y_j)$ as

$$u(x_j, y_j) = \sum_{k=1}^{3} u'(x'_k, y'_k) \ \phi_k(x_j, y_j), \tag{5.6}$$

  where $\{\phi_k : k = 1, 2, 3\}$ are the linear Lagrange polynomials defined on the triangle `jtP`.

Note that the above algorithm operates triangle-wise, hence every node lying on an internal edge of the mesh is processed twice, which is redundant if u' is single-valued (as it is in the case of Atlas). However, the idea behind the current implementation is to leave the door open for importing u' from another DG solver.

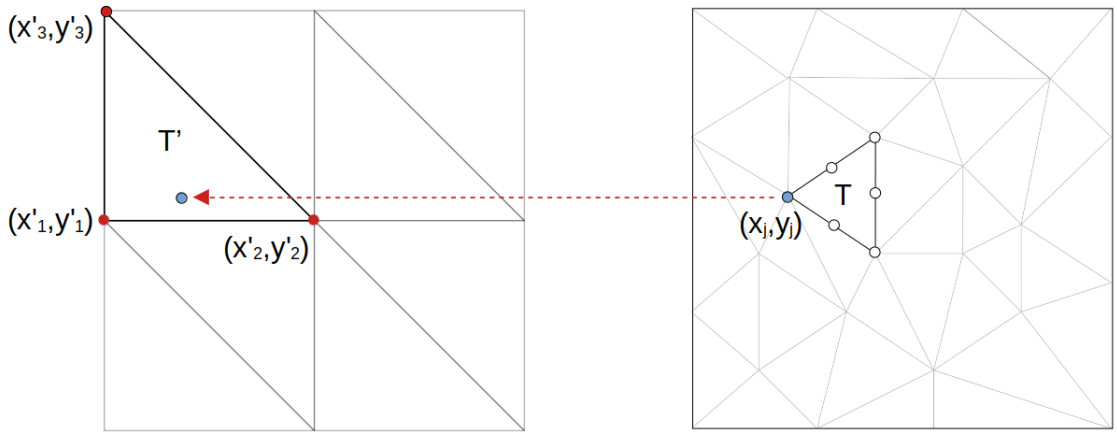


Figure 5.2: The degree of freedom (x_j, y_j) of the local DG scheme on T, an element of the unstructured mesh, is located in the Cartesian mesh. The element T' containing (x_j, y_j) is identified in the latter mesh, along with its vertices $(x'_1, y'_1), (x'_2, y'_2), (x'_3, y'_3)$. To calculate the steady state at (x_j, y_j), the values of the Atlas solution at such vertices are interpolated.

5.3 Simulated devices

Following the literature on FDTD simulation of PCAs [MPG+14],[KYI09],[CCSH00], a reasonable starting point is the classical structure depicted in Figure 1.1, consisting of two electrodes on a semiconductor, with the in-between gap targeted by incident radiation. This is sometimes referred to as an *Auston switch*.[1]
With respect to such references, in which large devices (of the order of 100 μm) are considered, here the interest is in structures that approximately fit 10μm. This is the result of a trade-off between computational load, whose limitation is not the core topic of the present work, and realistic applications. Indeed, advancements in fabrication technology have led to large antennas with gaps of a few μm [YHBJ14],[LGK+17] or to assembling antennas of such a small size into large arrays [YJ17],[ANH+07].
In both cases the point is to exploit plasmon resonance occurring in subwavelength-sized electrodes and/or metal nanoparticles suitably deposited on the gap – this appears to be *the* route to acceptable optical-to-THz conversion efficiency [BES17].

The present work aims at analyzing isolated antennas, regarded as potential elementary bricks of arrays, and at comparing different geometrical and numerical settings. Developing a method to study and refine a promising isolated antenna can be a precious stepping-stone to future developments, such as array optimization and investigation of coupling effects between multiple radiating elements.

It should be noted that large PCAs are typically complemented with a lens that collects and focuses the radiated field. The literature on arrays is less explicit on this topic, but an insulating substrate is typically present below the semiconductor. In our analysis the substrate is not essential – it would be the same for all the compared devices – hence it is omitted. Nevertheless, we look at its introduction as a reasonable future step towards large device simulation, and suggest two

[1] After the physicist D.W. Auston, who implemented it in the late '60s [Sie11], [Aug75].

configurations (Figure 5.3 and 5.4):

- Periodic array of small antennas
  The substrate has theoretically unlimited depth and is artificially truncated at the bottom of the computational domain. Both the semiconductor and the substrate span the entire width of the domain, which is terminated with absorbing (periodic) boundary conditions on horizontal (vertical) segments.

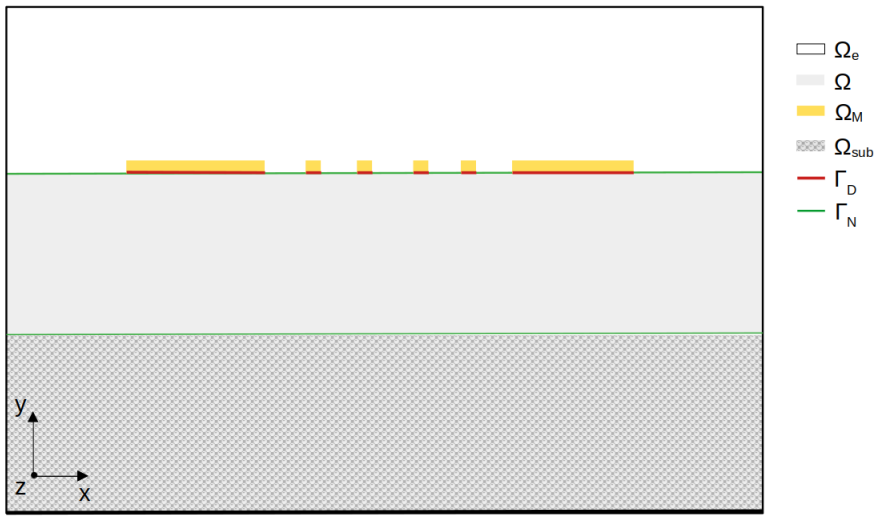


Figure 5.3: A possible computational setting for the element of a periodic array of small antennas. Infinite replicas of the computational domain are implied along x.

- Single large antenna with a small gap
  The idea is to simulate the illuminated portion of a large antenna – much larger than the gap. The semiconductor and the substrate are assumed have theoretically infinite width; the left (right) electrode is unlimited in the left (right) direction. Similarly to the previous case, the substrate has theoretically infinite depth. As a result, the computational domain is terminated with absorbing boundaries in all directions.[2]

[2]In this case, absorbing boundary conditions would have to be defined for electron and hole transport, e.g. the flow-based conditions seen in Section 4.2. A trickier point would be to calculate the stationary state in presence of semi-infinite electrodes.

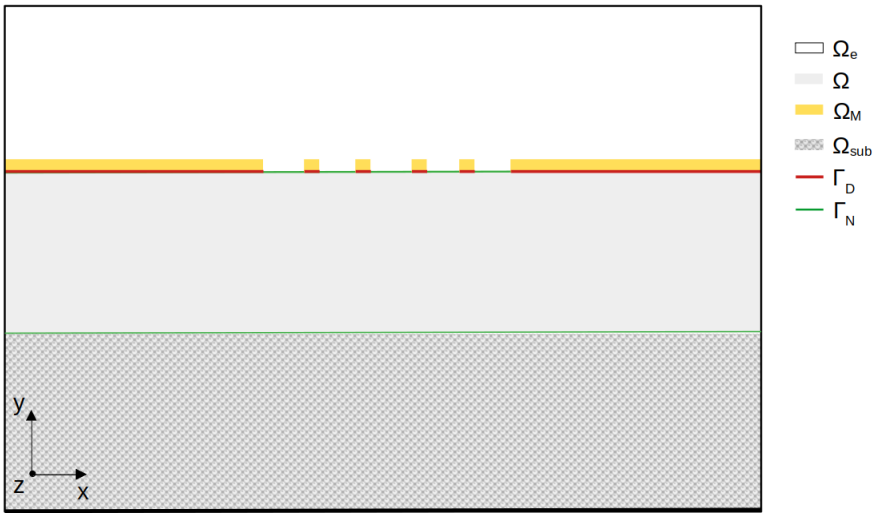


Figure 5.4: A possible computational setting for the illuminated portion of a large antenna. The structure is theoretically infinite along x.

For the validation of the developed DGTD solver, the following three topologies are considered and adapted into suitable two-dimensional representations; to choice is based on their relevance in applications and on the variety of challenges that they are expected to pose in the numerical framework.

- The classical Auston switch, downscaled to a few μm. As outlined in Section 1.4, low efficiency is its main pitfall; this leads to two main enhancements, that are considered in the next two structures.
- The initial PCA covered with an anti-reflective coating. This feature is implicit and ideal in the setup of [MPG$^+$14], where $\eta_e > 0.99$ is assumed in the predictive photo-generation function (2.20). Simulation and fabrication of realistic ARC-enhanced PCAs are discussed in [BFH$^+$19] and [BPK$^+$19]; a similar technique is used in solar cells as well [HYL$^+$20].
- A plasmon-enhanced PCA such as the one illustrated in Figure 1.2. The basic principle is to introduce a grating of metallic nanoparticles (*isles*) on the illuminated surface in order to produce a strong near-field distribution in the underlying semiconductor. This kind of geometry is simulated in [CB20a] and [BPK$^+$19]. Fabrication is discussed in [LGK$^+$17] and [BPK$^+$19].

5.3.1 Physical setting

The following physical setting is adopted in the simulations. The designated semiconductor is n-doped GaAs grown at 199°C, while electrodes/isles are made of gold. Experimental dispersion data for gold from Johnson and Christy [JC72] is fitted with a single-pole Drude model; a single-pole Lorentz model is considered to fit data from the Sopra S.A. database in the 600-1000 nm spectrum.[3]

[3]The database is available at http://www.sspectra.com/sopra.html

Both models stem from a simulated annealing algorithm [Viq18] implemented in the DIOGENeS software suite [Atl]), whose outputs are

$$\epsilon_\infty = 1.0, \qquad \omega_d = 1.39 \times 10^7 \text{GHz}, \qquad \gamma_d = 3.23 \times 10^4 \text{ GHz}, \tag{5.7}$$

for gold, and

$$\epsilon_\infty = 6.16, \qquad \omega_p = 1.96 \times 10^7 \text{GHz}, \qquad \omega_r = 4.89 \times 10^6 \text{ GHz}, \qquad \gamma = 0.98 \times 10^6 \text{ GHz}, \tag{5.8}$$

for GaAs. The static relative dielectric constant is necessary too, for the electrostatic problem. It could be drawn from experimental data, but since there is no definitive source – the literature shows values distributed between 10 and 13.2 [Ada99] – it seems reasonable to rather calculate it from the dispersion model:

$$\epsilon_r(\omega)|_{\omega=0} = \epsilon_\infty + \frac{\omega_p^2}{\omega_r^2} \approx 12.15, \tag{5.9}$$

the obtained value being consistent with the experimental range.

It is assumed that $n_i = 9 \times 10^6$ cm^{-3} and that doping yields a uniform donor concentration $N_D = 1.3 \times 10^{16}$ cm^{-3}. The recombination function is modeled with

$$\begin{aligned} n_1 = p_1 = n_i, \qquad & C_{An} = C_{Ap} = 7 \times 10^{-30} \text{ cm}^6 \text{ s}^{-1}, \\ \tau_n = 0.3 \text{ ps}, \qquad & \tau_p = 0.4 \text{ ps}, \end{aligned} \tag{5.10}$$

where $n_1 = p_1 = n_i$ implies that traps are at the Fermi energy level ($E_T = E_F$). For the Caughey-Thomas model of mobility, we take

$$\begin{aligned} \mu_{n0} &= 8000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}, & \mu_{p0} &= 400 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}, \\ D_{n0} &\approx 208 \text{ cm}^2 \text{ s}^{-1}, & D_{p0} &\approx 10 \text{ cm}^2 \text{ s}^{-1}, \\ \beta_n &= 1.82, & \beta_p &= 1.75, \\ v_{n,sat} &= 1.725 \times 10^7 \text{ cm s}^{-1}, & v_{p,sat} &= 0.9 \times 10^7 \text{ cm s}^{-1}. \end{aligned} \tag{5.11}$$

The generation function is (2.42) with $\eta_i = 1$, while the incident field is a plane wave pulsed in time and space:

$$\begin{aligned} \tau &= t - t_0 - \frac{\mathbf{k} \cdot \mathbf{r}}{\omega}, \\ \mathbf{k} &= \omega\sqrt{\epsilon\mu} \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \qquad \mathbf{r} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}, \\ \mathcal{G}(\tau, x) &= E_0 \exp\left(- \frac{\tau^2}{2\sigma_t^2} \right) \exp\left(- \frac{(x - x_0)^2}{2\sigma_x^2} \right), \\ \mathbf{E}^i(x, y, t) &= \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix} \mathcal{G}(\tau, x) \cos(\omega\tau), \\ \mathbf{H}^i(x, y, t) &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \frac{\mathcal{G}(\tau, x)}{\zeta_0} \cos(\omega\tau), \end{aligned} \tag{5.12}$$

with

$$\begin{aligned} E_0 &= 10^6 \text{ V cm}^{-1}, & \lambda &= \frac{2\pi}{\omega} = 800 \text{ nm}, \\ t_0 &= 0.33 \text{ ps}, & \sigma_t &= 0.08 \text{ ps}, \\ \sigma_x &= 1 \ \mu\text{m}, & \theta &= -\pi/2, \end{aligned} \tag{5.13}$$

where λ is the central wavelength in the spectrum. The laser spot is centered at x_0 and varies depending on specific requirements.
A voltage of 10 V is applied to all devices in a symmetrical way, with -5 V to the left contact (cathode) and 5 V to the right one (anode), as illustrated in Figure 5.5.

The geometry is defined with the idea of having a device characteristic size of a few microns and a gap larger than the laser spot size σ_x defined in (5.13). The other dimensions are inspired by the orders of magnitude and the proportions considered in references on fabrication and simulation of plasmon-enhanced PCAs (e.g. [Bur16] and [BPK$^+$19]). In detail, we consider isles measuring 100×80 nm with 100-nm center-to-center spacing, electrodes of 750×80 nm, and set the GaAs layer depth to 200 nm.

The device width is then supposed to be 4.2 μm, which allows a gap of 2.7 μm and a maximum of 13 isles for the device under test.

5.4 Simulation results

This section presents simulation results for the three selected antennas. The analysis starts with the illustration of the steady-state of the device, complemented by considerations on its influence on the subsequent photo-generated carrier dynamics.
For relevant observables, one-dimensional time-domain and frequency-domain probes are proposed, along with two-dimensional plots and sequences of snapshots. The simulated time interval is $[0, t_f]$ with $t_f = 4t_0 = 1.33$ ps in all cases.
Implementation-wise, the developed DGTD solver is coded in Fortran90 and executes sequentially, yelding computation times of the order of 10^5 seconds for the considered meshes and distributions of degrees of freedom. All the meshes have been created in Gmsh 4.8 using its Python API [GR09]. It should be noted that code optimization has not been regarded as a primary goal in the current work, due to the aim of migrating the developed DGTD solver into an existing optimized three-dimensional one (DIOGENeS, [Atl]). On the other hand, major attention has been devoted to performance-improving features of the DG method itself, such as local hp-refinement.

5.4.1 Classical PCA

The classical PCA implemented in Silvaco Atlas is shown in Figure 5.5. The computational domain, composed of the semiconductor, the two electrodes and the in-between air – no further surroundings – is discretized with a Cartesian mesh (omitted in the figures) whose size is 10 nm along x, whereas along y it starts from 10 nm at the electrodes to attain 20 nm at the bottom of the semiconductor. As outlined in Section 5.2, the static PDD solver is launched multiple times, with the bias voltage going from zero to the target value (-5 V and 5 V at the left and right contact respectively).

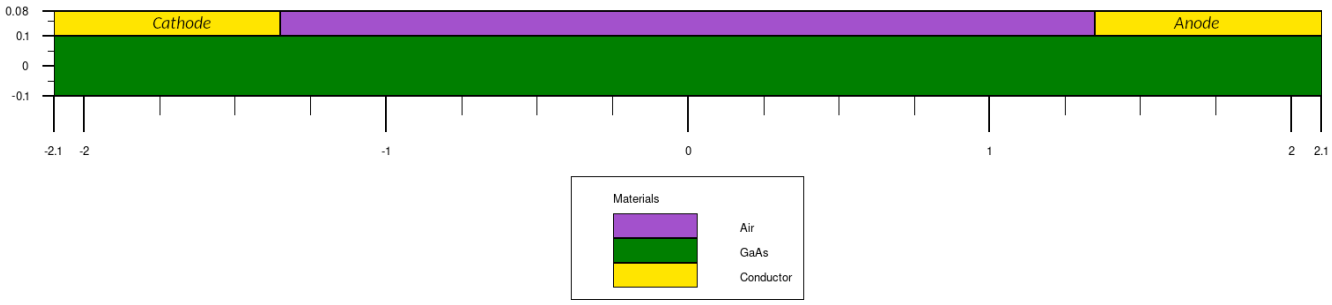


Figure 5.5: Representation in Atlas of the simulated PCA (axes are in μm). The electrodes have been labeled as according to the sign of the applied voltage: $-5V$ at the cathode, $+5V$ at the anode.

- The first computation yields the (uniform) thermodynamic equilibrium concentrations $n_{DC} \approx N_D = 1.3 \times 10^{16}$ cm^{-3} and $p_{DC} \approx n_i^2/N_D \approx 6.2 \times 10^{-3}$ cm^{-3}.
- The final one yields the steady state that will undergo illumination. The resulting electrostatic potential and carrier concentrations are reported in Figure 5.6. With respect to the thermodynamic equilibrium value, we notice that electron concentration experiences an increase by one order of magnitude close to the air-anode interface, counterbalanced by a comparable reduction in the underlying region.

Larger electron concentrations corresponds, as expected, to a stronger electrostatic field (Figure 5.7). Localizing the peak of the electrostatic field is important for optical-to-THz conversion efficiency – the stronger the field, the larger the amplitude of the photo-generated current pulse. It is convenient, then to also maximize photo-generation in the same region by pointing the laser spot closer to the anode. This topic will be investigated later on in the text.

Cutlines are employed to verify that the imported steady state is correctly represented on the unstructured mesh used for the DGTD simulation (introduced later, in Figure 5.11b). For example, $E'_{DC,x}$ and n'_{DC}, defined on the Cartesian mesh, are plotted against their interpolated counterparts (with quadratic Lagrange polynomials) in Figures 5.8 and 5.9.

The illustration of the steady state is completed with a brief insight into the static recombination function $R'(n_{DC}, p_{DC})$ introduced in (2.17) (Section 2.3).
According to Figure 5.10, strong thermal generation (negative recombination) permanently occurs in the bulk of the semiconductor, at the anode side. A comparison with Figure 5.6 suggests that the phenomenon takes place where electron concentration is reduced by one order of magnitude with respect to thermodynamic equilibrium, whereas hole concentration is slightly higher; as a result, overall $n_{DC}p_{DC} < n_i^2$, and $R' < 0$.

Having imported the DC quantities it is possible to set up the DGTD simulation of the MDD system. The computational domain consists in the PCA surrounded by air and is artificially truncated with ABCs at a distance of at least one wavelength from the device.
As highlighted in Section 3.1.4.3, the maximum stable time step is ultimately determined by the electromagnetic part of the system. Then, the mesh must be fine enough to resolve spatial variations of the unknowns. According to the Nyquist-Shannon theorem, within a distance equal to the

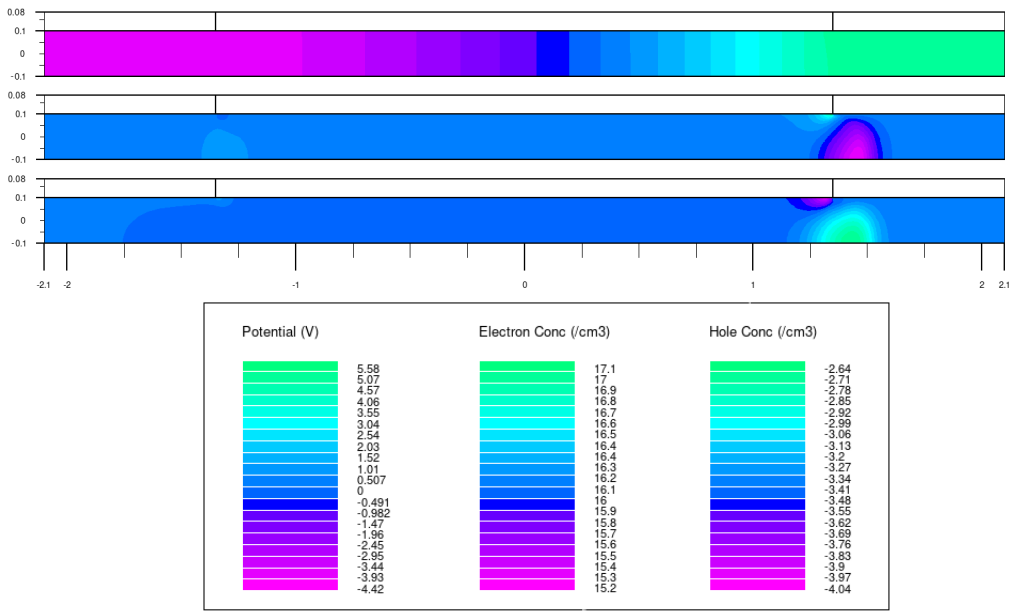


Figure 5.6: Two-dimensional plots of the steady state of the semiconductor. From the top to the bottom: electrostatic potential (linear scale), electron concentration (logarithmic scale), hole concentration (logarithmic scale).

minimum wavelength in the spectrum, there cannot be less than two samples (degrees of freedom) of the electromagnetic field [TH05]. The Nyquist-Shannon theorem holds for any space-varying function, including carrier concentrations; however, the fact that these obey non-linear advection-diffusion-reaction equations makes it unfeasible to define a wavelength – a different characteristic length should be sought.

In advection-diffusion problems it is common to keep track of how the two transport phenomena concur. A figure of merit is the local Péclet number

$$Pe = \frac{h}{2} \frac{||\text{Advection velocity}||}{\text{Diffusivity}}, \tag{5.14}$$

where h is the local mesh size and $||.||$ is the Euclidean norm. The condition $Pe > 1$ has been linked to spurious spatial oscillations [Qua13],[Smi80] and lower convergence rates [Led21] in studies on continuous Galerkin methods applied to simple cases of constant diffusivity and divergent-free advection velocity, in one dimension. These suggest to use mesh refinement as a counterweight whenever advection is expected to outnumber diffusion.

Our drift-diffusion model falls into this category – assuming $T = 300\text{K}$, $V_T \approx 0.026\text{V}$ and according to the Einstein relations ($D_n = V_T \mu_n$, $D_p = V_T \mu_p$) mobility is higher than diffusivity by two orders of magnitude. That said, diffusivity is (potentially strongly) space-varying, the divergence of electron (hole) drift velocity $\boldsymbol{v}_{n,drift} = \mu_n(\mathbf{E} + \mathbf{E_{DC}})$ ($\boldsymbol{v}_{p,drift} = \mu_p(\mathbf{E} + \mathbf{E_{DC}})$) cannot be neglected, and reaction plays a critical, possibly dominant role (as pointed out in the discussion on the generation function in Section 4.2). Such high complexity suggests that a suitable figure of merit should somehow take carrier concentration into account.

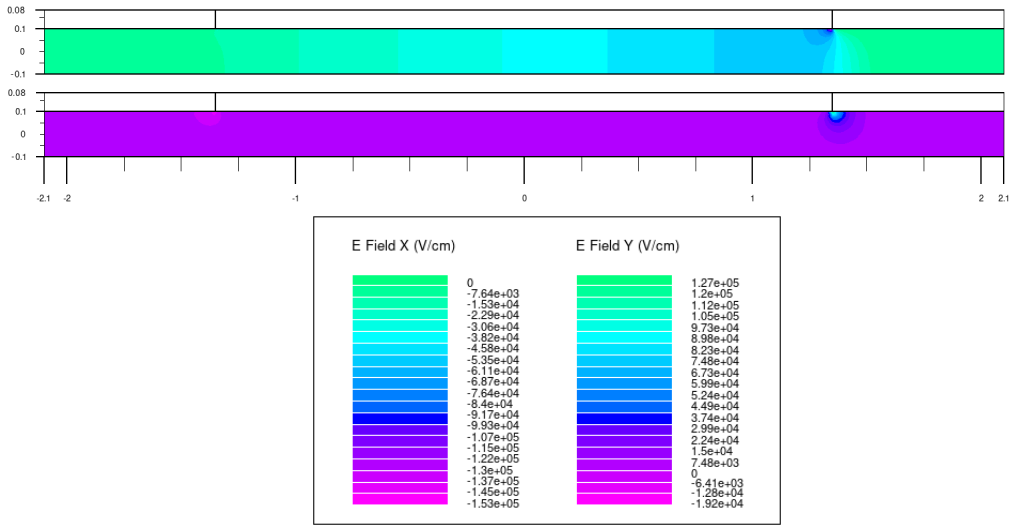


Figure 5.7: The x and y components of the electrostatic field inside the semiconductor.

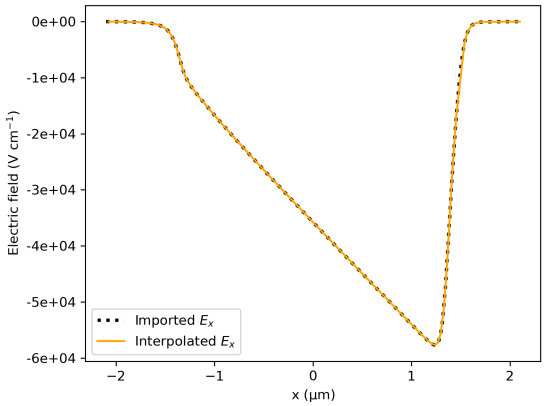


Figure 5.8: $E'_{DC,x}$ vs $E_{DC,x}$

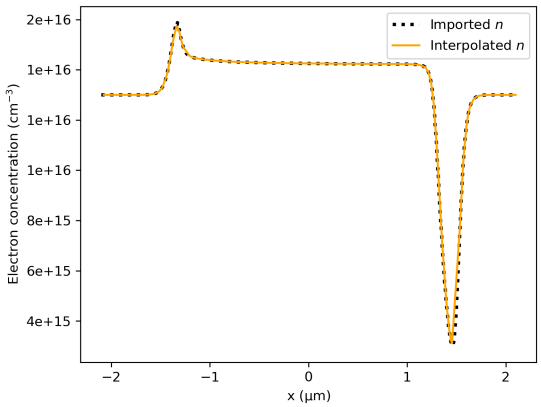


Figure 5.9: n'_{DC} vs n_{DC}

In [VG06], for example, doping concentration N_D is taken into account – the Debye length

$$L_D = \sqrt{\frac{\epsilon_0 \epsilon_r V_T}{q N_D}}, \tag{5.15}$$

is designated as the mesh size limiter in the FDTD resolution of the time-varying PDD model presented in Section 2.4. Although the same criterion is also found in the MDD context (FDTD simulation of PCAs in [KYI09], deriving a mathematical relation between N_D and n_{AC}, p_{AC} that justifies its usage seems not trivial.
Nevertheless, before going further, it is interesting to get a rough estimation of the mesh size required to comply with Péclet- and Debye-based criteria.
Assuming $\epsilon_r \sim 10$ and $N_D \sim 10^{16}$ cm^{-3}, the Debye length is $L_D \approx 40$ nm.

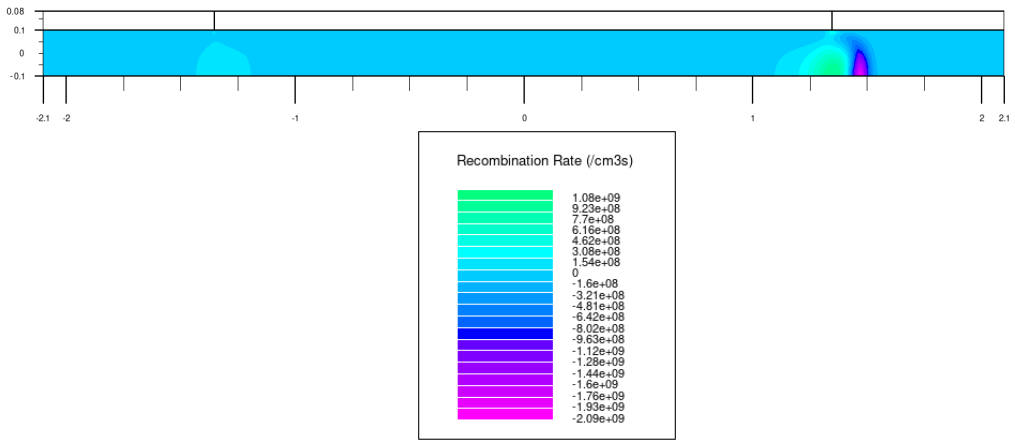


Figure 5.10: The spatial profile of the static recombination function $R'(n_{DC}, p_{DC})$ defined in Section 2.3 allows to recognize a region of steady-state thermal generation ($R' < 0$) below the anode, in the bulk of the semiconductor.

As to the Péclet number, reasoning for example on electrons, using the Einstein relation with $V_T \approx 0.026$V and the approximation $||\mathbf{E}_{AC}|| << ||\mathbf{E}_{DC}||$, we write

$$\begin{aligned} \frac{Pe}{h} = \frac{||\boldsymbol{v}_{n,drift}||}{2D_n} = \frac{||\mu_n \ (\mathbf{E}_{AC} + \mathbf{E}_{DC})||}{2D_n} = \frac{\mu_n \ ||\mathbf{E}_{AC} + \mathbf{E}_{DC}||}{2D_n} \\ = \frac{||\mathbf{E}_{AC} + \mathbf{E}_{DC}||}{2V_T} \approx \frac{||\mathbf{E}_{DC}||}{2V_T} \approx 20 \ ||\mathbf{E}_{DC}||. \end{aligned} \tag{5.16}$$

Note that reasoning on holes would yield the same result since $V_T = D_n/\mu_n = D_p/\mu_p$. We conclude that

$$Pe < 1 \iff h < \left(20 \ ||\mathbf{E}_{DC}||\right)^{-1}. \tag{5.17}$$

So, if $||\mathbf{E}_{DC}|| = 10^6$ V m^{-1}, $h_{max} = 50$ nm, whereas if $||\mathbf{E}_{DC}|| = 10^7$ V m^{-1}, the constraint becomes much more severe: $h_{max} = 5$ nm.

In the end, given

- the ultimate goal of every estimation – making the spatial discretization robust enough to resolve sharp spatial variations of n_{AC} and p_{AC},
- the difficulty to derive a rigorous formula that quantifies the maximum the mesh size (apart from the Nyquist-Shannon theorem),

over the course of the present work, it has appeared sensible to opt for a semi-empirical approach:

- tailoring the mesh for the electromagnetic field with a certain number of samples per wavelength;
- identifying regions where spatial variations of carrier concentrations are likely to be sharpest (e.g. on the basis of the steady state and the illumination type);

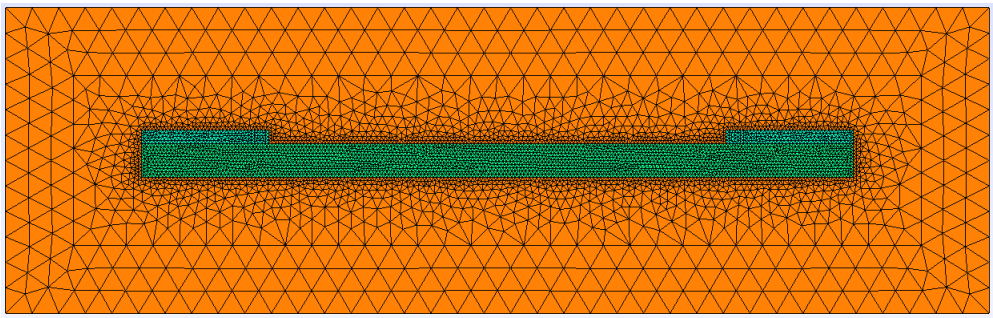
- locally refining the mesh, if such variations are not correctly resolved.

The effect of refining the mesh is to reduce the distance between two samples of the solution; one alternative consists in leaving the mesh unchanged while increasing the interpolation degree. The point can be better appreciated by checking tracing a cutline intercepting the region where photo-generation has its spatial peak, i.e. where carrier concentration gradients are expected to be highest.
If $x_0 = 0$ (centered laser spot), a good candidate is the neighborhood of $(x, y) = (0, 0)$, as will be confirmed a posteriori by the two-dimensional plots proposed later on. To take the worst case, the observation should take place when carrier concentration has reached its highest value, which is expected to happen slightly after the illumination peak $t = t_0$. Simulations have proven that a suitable choice is $t = \bar{t} := 1.3\ t_0$.

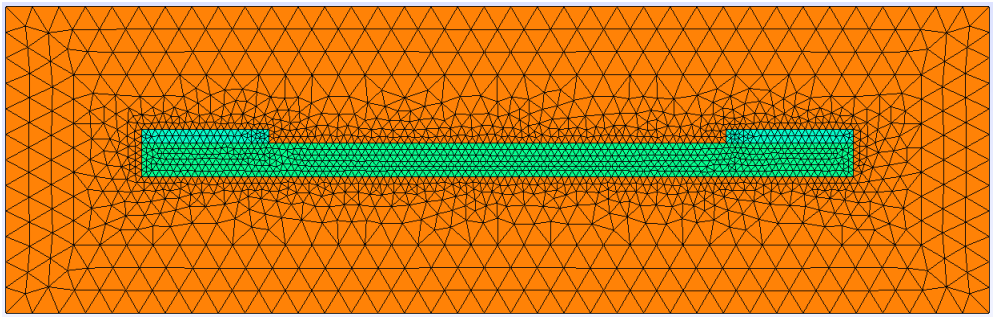
For the sake of completeness, the analysis should also ponder the effectiveness of Péclet- and Debye-based criteria. Since h required by $Pe < 1$ is the same for electrons and holes, we can begin by considering the latter. With reference to the coordinates of Figure 5.5, the horizontal cutline

$$\{p(x, y, t) \ : y = 0, t = \bar{t}\},$$

is observed in the following two scenarios.



(a) The mesh used in (a), with $h = 21$ nm inside the semiconductor and the electrodes (green) and an increasing size until $h = 157$ nm in the rest of the computational domain (orange).



(b) The coarser mesh used in (b), with $h = 42$ nm inside the semiconductor and the electrodes.

Figure 5.11: The two levels of refinement considered in the present test.

(a) The mesh size is $h = 21$ nm inside the semiconductor and in the vicinity of inner electrode wedges; it then increases to reach $h = 157$nm at the boundary of the computational domain. Linear Lagrange interpolation is used. The maximum stable time step is $\Delta t = 0.33 \times 10^{-5}$ ps (Figure 5.11a).
The triangulation is such that at least 10 samples per wavelength are ensured inside the semiconductor and at least 5 at the boundary of the computational domain.[4]

(b) The minimum mesh size is doubled ($h = 42$ nm) but quadratic Lagrange interpolation is used. The maximum stable time step is now $\Delta t = 10^{-5}$ ps (Figure 5.11b).

The cuts are juxtaposed in Figure 5.12. In (a) with $h = 21$ nm and linear interpolation, spatial oscillations arise. Case (b) shows a doubling the interpolation degree has a stabilizing effect, in spite of the coarser mesh, and allows a larger time step as a plus.

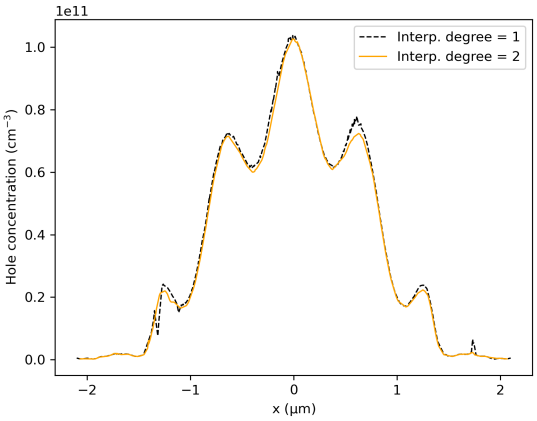


Figure 5.12: Superposition of two plots of $\{p(x, y, t) \ : y = 0, t = \bar{t}\}$ in the cases (a), (b).

[4]The maximum size is chosen with the idea of representing the minimum wavelength in the incident spectrum carrying 10% of the power, which is $\lambda_{min} \approx 788$nm. Then, this reduces to about 211 nm inside GaAs by virtue of $\epsilon_{GaAs}(\lambda_{min}) \approx 13.9 + i1.0$ (as per the dispersion data used in this work).

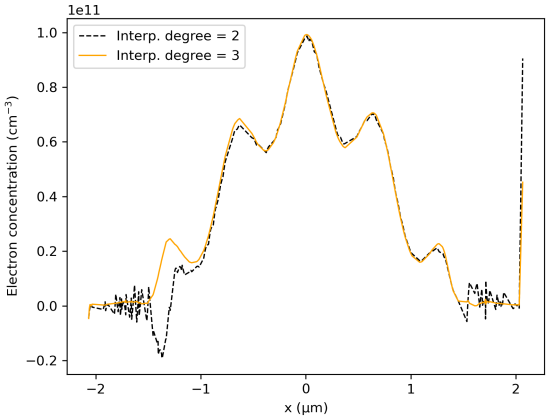


Figure 5.13: The line $\{n(x, y, t) \ : y = 0, t = \bar{t}\}$ calculated with quadratic and cubic interpolation, revealing the inadequacy of the former.

Let us see where the two settings stand with respect to Debye- and Péclet-based criteria. We have already noted that $L_D \sim 40$ nm with our physical input data. On the other hand, from Figures 5.7 and 5.8 we estimate $\max ||\mathbf{E}_{DC}|| \approx 1.6 \times 10^4$ V cm^{-1}, which imposes $h < 30$ nm. Now, we need to estimate the average distance between two solution samples in (b). Since $h = 42$ nm and the interpolation degree is 2, we can roughly assume that such a distance is $h/2 = 21$ nm. The effective value is perhaps lower, as 21 nm is the mesh size of (a), which (b) outperforms. Regardless, it can be concluded that both configurations are such that $h < L_D$ and $Pe < 1$.

Since the two criteria are the same for electron and hole transport equations, the above discussion might seem sufficient to designate (b) as an appropriate numerical setting. Yet, contrarily to expectations, a check on electron concentration calculated in (b), along the same cutline traced earlier, exhibits abrupt oscillations (Figure 5.13). We must conclude that $h = 42$ nm and quadratic interpolation do not constitute an appropriate setting. The verdict is confirmed by the second cutline in the figure; this is obtained by increasing the interpolation order, which yields appreciable stabilization.

The drastically different behavior of n with respect to p can be understood as follows.[5] Recollecting the expressions of the (time-varying) drift currents

$$\begin{aligned} \mathbf{J}_{n,drift} &= q\ n\ \mu_n\ (\mathbf{E}_{DC} + \mathbf{E}) + q\ n_{DC}\ \mu_n\ \mathbf{E}, \\ \mathbf{J}_{p,drift} &= q\ p\ \mu_p\ (\mathbf{E}_{DC} + \mathbf{E}) + q\ p_{DC}\ \mu_p\ \mathbf{E}, \end{aligned} \tag{5.18}$$

it seems reasonable to expect

$$||\nabla \cdot (q\ n_{DC}\ \mu_n\ \mathbf{E})|| >> ||\nabla \cdot (q\ p_{DC}\ \mu_p\ \mathbf{E})||, \tag{5.19}$$

because, as depicted in Figure 5.6, $n_{DC} >> p_{DC}$ while their spatial variations are similar (as they stem from the same source – the electrostatic field); so, electrons are solicited by an enormously larger source term and thus demand special surveillance in the numerical framework.

[5] Up to a few exceptions where the time-varying character of a particular quantity needs to be stressed, we shall keep the convention of implying the AC subscript, so that static quantities are explicitly marked as DC, while those with no subscript are time-varying.

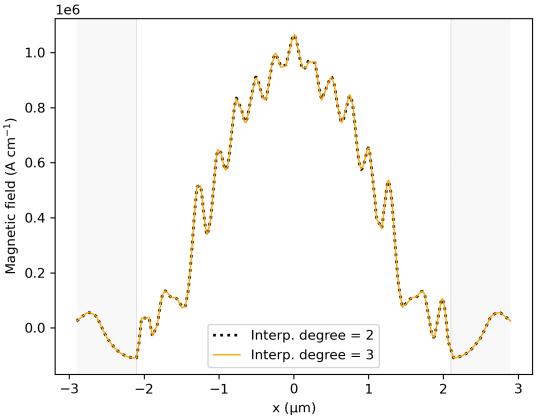


Figure 5.14: The magnetic field at the illumination peak ($t = t_0$) observed along the line $y = 0$. In the grey areas, which lie outside the device, quadratic interpolation is used. Inside the device, quadratic and cubic interpolation are compared.

A further important remark is necessary. The abrupt character of n that is recognized in Figure 5.13 at the endpoints of the device (the vertical walls of the semiconductor), which might be qualified at first glance as a byproduct of numerical instability, is rather a mathematically expected behavior.
First, the laser is polarized along x, so inside the semiconductor $|E_x| >> |E_y|$, implying also $|q\ n_{DC}\ \mu_n\ E_x| >> |q\ n_{DC}\ \mu_n\ E_y|$. Then, the insulating boundary condition stated in (2.59) becomes, at vertical walls, $\hat{\mathbf{n}} \cdot \mathbf{J}_n = \hat{\mathbf{x}} \cdot \mathbf{J}_n = 0$. We conclude that strong diffusion and a consequently large gradient ($\hat{\mathbf{x}} \cdot \mathbf{J}_{n,diff} = qD_n \hat{\mathbf{x}} \cdot \nabla n$) must arise along x to keep the total current tangential to the walls.
We recognize at this point a crossroad – we can try to fully capture such space variations with a tighter discretization, or move forward and only reconsider the first way if such higher-order effects come to majorly affect device simulation. In the present work, the latter option has been chosen.

It is natural to wonder whether the increase in the interpolation degree in (b) is required by the field too, since the wavelength (significantly) shortens inside the semiconductor. Figure 5.14 answers the question with a cutline on H_z at the illumination peak $t = t_0$:

$$\{H_z(x, y, t) \ : y = 0, t = t_0\}.$$

Outside the device (shadowed areas), the original (quadratic) interpolation is retained, whereas inside, quadratic and cubic approximations are compared. The juxtaposition manifests no major issues with the lower order, contrarily to Figure 5.13. This is an interesting result – it suggests to orient future research towards a local DG formulation with different interpolation degrees for light propagation and charge transport.

An important remark concerns field approximation inside the electrodes – conductors are not perfect – and in the proximity of metallic wedges, where higher gradients are expected with respect to the shadowed areas of Figure 5.14. This is the reason why the minimum mesh size is imposed

in the semiconductor *and* in such regions (Figure 5.11b), as a mean of support to quadratic interpolation.

In the end, the designated numerical setting consists in a locally adaptive DG scheme in the mesh of Figure 5.11b (3246 elements) with

- Lagrange interpolation of order 3 and $h = 42$ nm inside the semiconductor (985 elements),
- Lagrange interpolation of order 2 and $h = 42$ nm inside the electrodes (172 elements),
- Lagrange interpolation of order 2 and $42 < h < 157$ nm in the rest of the computational domain (2089 triangles),
- explicit LSRK2 time stepping with $\Delta t = 10^{-5}$ ps.

The ground is ready at this stage to proceed and analyze the physics of the transient simulation.

We start with a sequence of parallel snapshots (Figure 5.15) of the photo-generation function (2.42) and electron concentration n in the illuminated box $(x, y) \in [-1.35, 1.35] \times [-0.1, 0.1]$ μm, throughout a relevant time interval, namely $[0.21, 1.04]$ ps. [6][7]

The highest photo-generation occurs around $(x, y) = (0, 0)$, confirming the relevance of the cutlines traced earlier. The peak time, around $t = t_0 = 0.33$ ps, triggers a sharp rise of electron concentration, which attains its own plateau about 10 ps later, at $t = \bar{t} = 42$ ps. At $t = 50$ ps the light pulse is on its way out and recombination takes the lead, drastically reducing n. Reduction occurs slightly faster on the left side, as this is depleted not only by recombination but also by the electric field, which accelerates negative charges to the right.
As a result of centered illumination, the majority of photo-generated carriers recombines on its way to the electrode, so that the latter won't be able to significantly contribute to THz radiation. This point will be further discussed at the end of this section. The full evolution of $n(0, 0, t)$, $t \in [0, t_f]$, is shown in 5.16, with $G(0, 0, t)$ on the background.

The emergence of a THz pulse can be appreciated by means of a probe on the total electric current

$$\begin{aligned} \mathbf{J} &= \mathbf{J}_{drift} + \mathbf{J}_{diff}, \\ \mathbf{J}_{drift} &= q\big[n\mu_n(\mathbf{E}_{DC}) + p\mu_p(\mathbf{E}_{DC})\big](\mathbf{E}_{DC} + \mathbf{E}) + q\big[n_{DC}\mu_n(\mathbf{E}_{DC}) + p_{DC}\mu_p(\mathbf{E}_{DC})\big]\mathbf{E}, \\ \mathbf{J}_{diff} &= qD_n\nabla n - qD_p\nabla p. \end{aligned} \tag{5.20}$$

As just highlighted, photo-generated carriers are mostly accelerated along x. This is owed to at least two facts: the incident field is polarized along x; comparing Figures 5.15 and Figure 5.7, $E_{DC,x} >> E_{DC,y}$ where G is highest. Therefore, THz emission is mostly ascribed to J_x.

The probe $J_x(0, 0, t)$, $t \in [0, t_f]$ is proposed in Figure 5.17, in two variants. One is (5.20), which essentially displays the character of the optical pulse due to the large term $\propto n_{DC}\,\mathbf{E}$. The other is a much slower pulse, obtained by cutting off the optical spectral components of (5.20) with a Hamming window. The relevant portion of the resulting spectrum is shown in Figure 5.18, where it a -3 dB cut-off frequency is marked to highlight that at least 70% of the radiated power lies in the range $f < 10$ THz.

[6] The regions beneath the electrodes are excluded for reasons of page layout and visibility, but also because they are shielded from light if compared to the central aperture and hence inessential to the analysis.

[7] For a better appreciation of spatial profiles the amplitude axes have been capped at slightly lower values than the actual peaks, which were measured to be $7.8 \times 10^{15}\text{cm}^{-3}\text{s}^{-1}$ for G and $1.8 \times 10^{11}\text{cm}^{-3}\text{s}^{-1}$ for n. This scale will be retained when comparing this antenna to its plasmon-enhanced version.

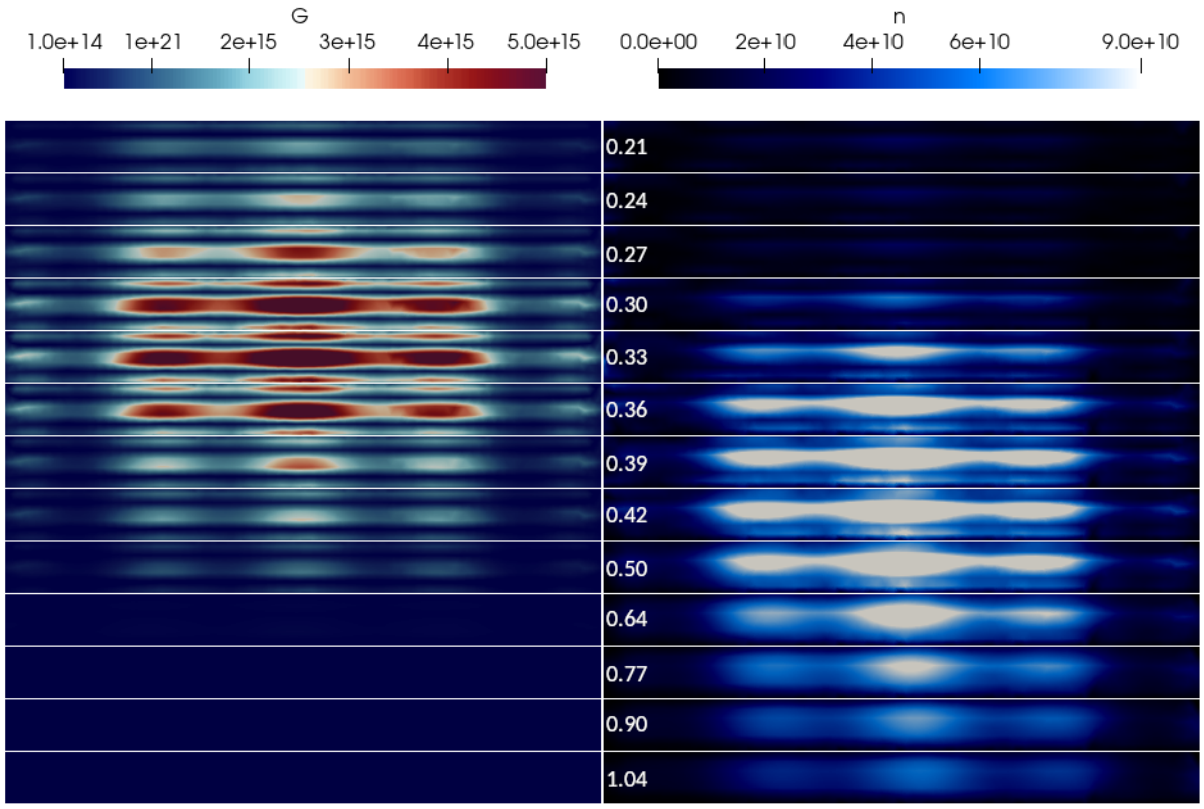


Figure 5.15: Concurrent snapshots of photo-generation (G, cm^{-3} s^{-1}) and electron concentration (n, cm^{-3}) at relevant times in $[0.21, 1.04]$ ps, in the illuminated region of the semiconductor. Times are marked in ps, in ascending order, from the top to the bottom.

At this stage, a remark on the emitted electromagnetic field is necessary. The literature on FDTD simulation of PCAs typically relies on a near-to-far-field transformation to observe the radiated THz pulse [MPG$^+$14], [KYI09], [CCSH00], [JJK96].
A simple but well-established model, based on arguments of classical electrodynamics, prescribes [Lee08], [ST05]

$$\mathbf{E} \propto \frac{\partial \mathbf{J}}{\partial t}, \tag{5.21}$$

suggesting that THz spectral components in the scattered field stem from $\mathbf{J}$, and making the latter a legitimate and convenient observable. For instance, in [MPG$^+$14] the spectrum and the time-domain shape of the current are observed along with the scattered field.

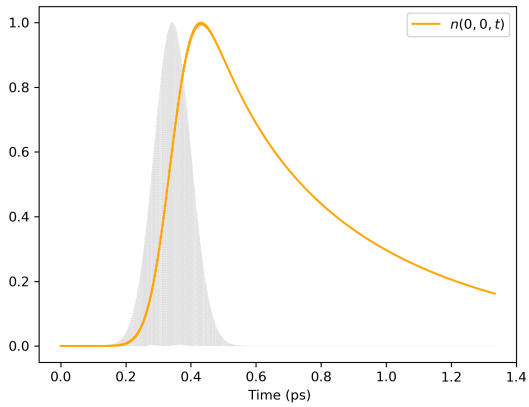


Figure 5.16: The evolution of generation and electron concentration throughout the simulation window $t \in [0, t_f]$, observed at the center of the semiconductor $(x, y) = (0, 0)$.

To conclude the analysis, we get back to the suggestion of intelligently pointing the laser so that photo-generation and the bias field are maximum in the same region.
Figure 5.19 shows two-dimensional plots of the magnitude of the total current (5.20) at the illumination peak time $t = t_0$ for different values of x_0, from the center of the gap towards the anode. As x_0 increases, the current density becomes more localized in the region underlying the anode (and more intense).
It is important in this test to monitor $||\mathbf{J}||$ rather than just $|J_x|$; indeed, if $|E_{x,DC}| >> |E_{y,DC}|$ at the center of the device, $|E_{y,DC}|$ is comparable to $|E_{x,DC}|$ in the proximity of the anode. Thus, as the illuminated spot moves to the right, both components of $\mathbf{E}_{DC}$ shape the trajectory of photo-generated carriers.[8]

It has been experimentally observed that asymmetric excitation improves optical-to-THz conversion efficiency [UFB+07]. The current analysis provides one reason, but there is a second, less evident one – the increase in the probability that photo-generated electrons attain the metal contact before recombining, which allows to exploit the electrode as a radiating element [Bur16]. This effect is not visible in the current model, where charge flow across semiconductor-metal interfaces is not quantified (an interesting discussion on the topic can be found in [ZHA+15]).

[8]It should be noted that the represented values in Figure 5.19 actually refer to $\frac{\mathbf{J}}{c_0\ q} = \left[\text{A C m}^{-1}\ \text{s}^{-1}\right]$ as a result of a convenient choice of physical dimensions in the code.

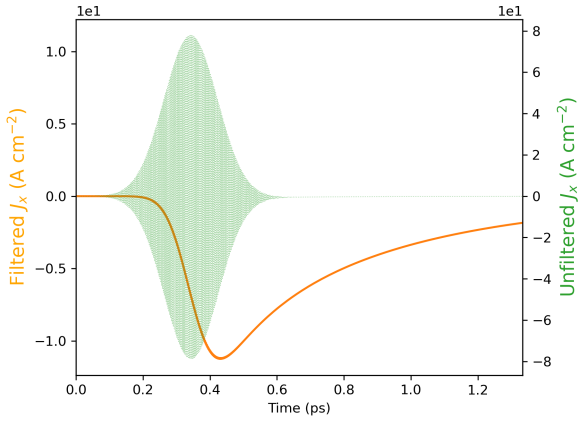


Figure 5.17: Filtered and unfiltered $J_x(0, 0, t)$ throughout the simulated interval $[0, t_f]$.

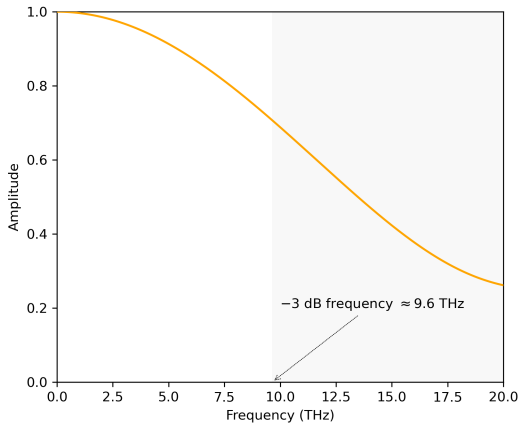


Figure 5.18: The magnitude of the Fourier transform of $J_x(0, 0, t)$ after filtering out the optical components.

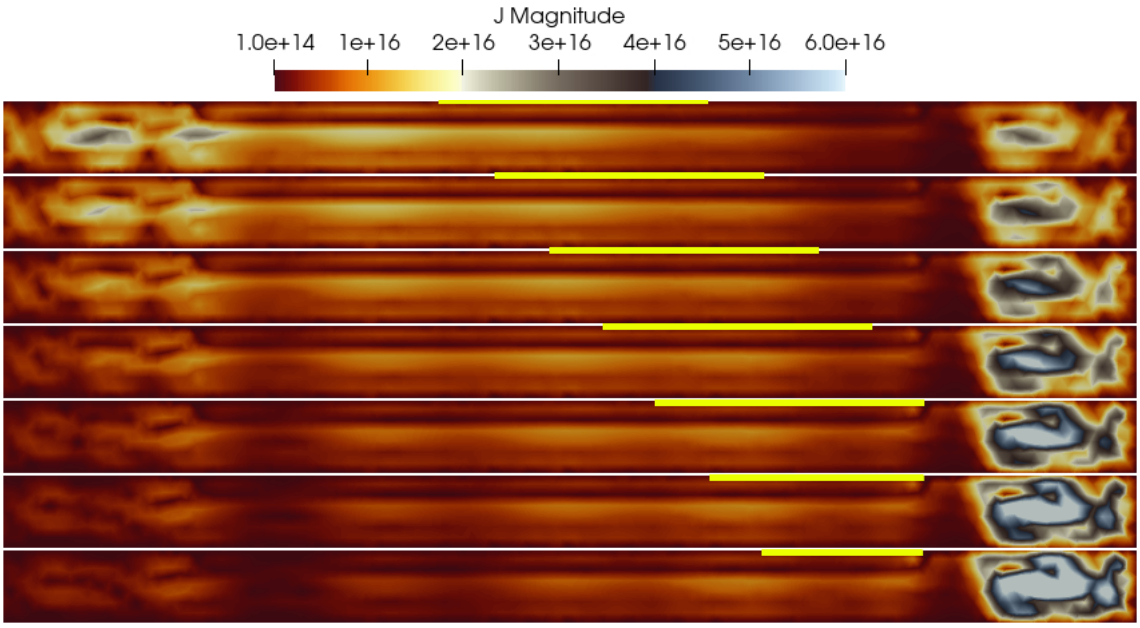


Figure 5.19: The effect of different illumination targets $x_0 \in \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2\}$ μm. As the laser spot (depicted in yellow) moves towards the anode, the current appears to be significantly enhanced. Interestingly, the effect is maximum when the beam is diffracted by the anode ($x_0 \geq 0.8\mu$m) and hence does not reach the semiconductor.

5.4.2 Classical topology with anti-reflective coating

An important pitfall in the previous antenna is the poor power matching between the laser and the semiconductor. Indeed, under radiation at $\lambda = 800$ nm, the semiconductor exhibits a refractive index $n_{GaAs} \approx 3.71 + i\ 0.13$, so that the resulting Fresnel's reflection coefficient at the air-semiconductor interface is

$$\mathcal{R} = \frac{1.0 - n_{GaAs}}{1.0 + n_{GaAs}} \approx -0.58 - i\ 0.01, \tag{5.22}$$

implying that almost 60% of the incident power is reflected back to the source.
As a possible workaround is to grow an Anti-Reflective Coating (ARC) on top of the illuminated surface – a layer of intermediate refractive index interposed between air and GaAs. A natural option is silicon dioxide, since this is a standard material for integrated circuit passivation, and $n_{SiO_2} \approx 1.45$ at $\lambda = 800$ nm [Mal65]. As outlined earlier, such a practice is common in solar cells [HYL+20] and has been adopted in PCA fabrication as well [BFH+19].[9]

[9]Passivation is the process of covering an integrated circuit with dielectric layers (typically made of silicon dioxide and silicon nitride) that protect it from impurities, moisture and scratches [SN06],[RCN02].

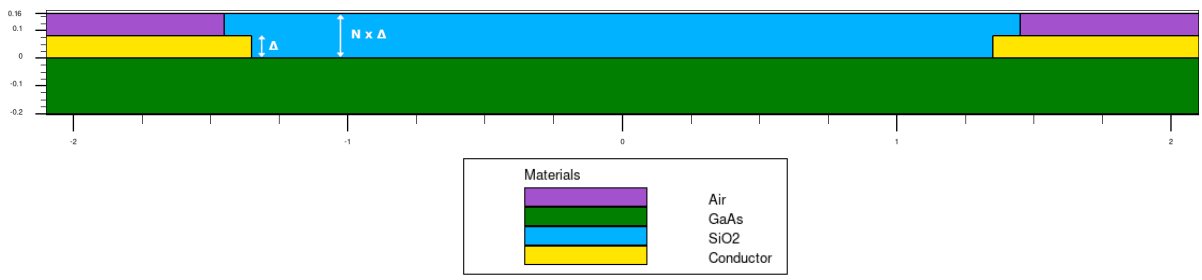


Figure 5.20: PCA with SiO2 Anti-Reflective Coating (ARC) of variable thickness $N \times \Delta$, with $\Delta = 0.08$ nm (i.e. the height of the electrodes)

The implemented ARC is shown in Figure 5.20. A simplifying assumption is to let the thickness $d = N \times \Delta$ be a multiple of that of the electrodes ($\Delta = 0.08$ nm), while the requirement on the width is to be large enough to cover the gap, e.g. with 50-nm margins.

The aim here is to investigate the effect of different design choices, such as $N = 0, 1, 2, 3, 4$. These are summarized in Figure 5.21, where the spatial profiles of the generation function G and electron concentration n are compared for each value of N. As usual, the two are monitored at their respective peak times, which occur around the incident pulse delay t_0 for G and around $\bar{t} \approx 1.1t_0$ for n. For each frame, the maximum value of the spatial distribution is recorded and collected in Table 5.1. The options $N = 1, 4$ appear promising at first glance, but unexpectedly that is the case of the initial configuration too ($N = 0$, i.e. no ARC). The table fosters $N = 4$ but does not make things much clearer and reveals quite interestingly, that the link between G and n is not trivial.

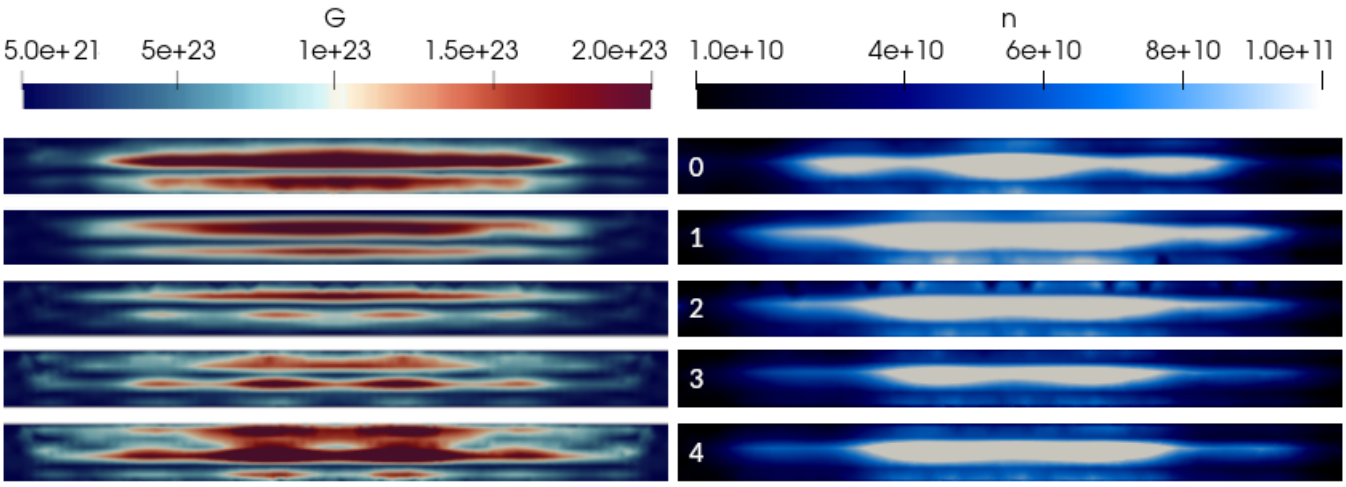


Figure 5.21: The spatial distribution of G ($\mathrm{cm}^{-3}\ \mathrm{s}^{-1}$) and n (cm^{-3}) at their respective peak times for $N = 0, 1, 2, 3, 4$.

N	$\max(G)$ $(\text{cm}^{-3}\ \text{s}^{-1})$	$\max(n)$ (cm^{-3})
0	4.7×10^{23}	1.0×10^{11}
1	2.5×10^{23}	2.0×10^{11}
2	2.3×10^{23}	1.6×10^{11}
3	2.6×10^{23}	1.4×10^{11}
4	4.6×10^{23}	1.8×10^{11}

Table 5.1: Peak values of photo-generation and electron density for different depths of the ARC.

In fact, the values are too close to infer a definitive conclusion from the observation of just one time sample. Indeed, another important datum in the picture is how long the generation peak persists, which depends on the multiple reflections that light undergoes inside the device and hence should be affected by the introduction of the ARC.
This additional element can be taken into account by looking at the evolution of the L^2 norm of n throughout the simulated interval $[0, t_f]$ (Figure 5.22).

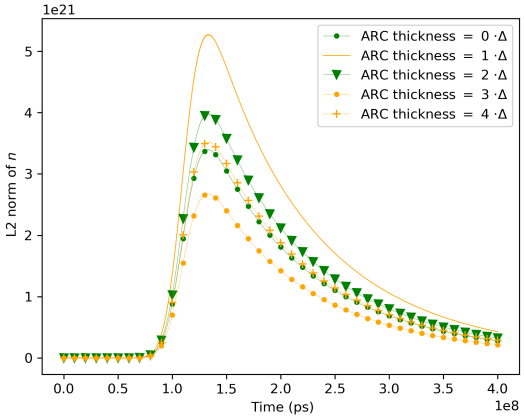


Figure 5.22: The evolution of $||n||_{L^2(\Omega)}$ throughout the simulations run for different values of the ARC thickness $N\Delta$, $N = 0, 1, 2, 3, 4$.

From norm analysis we infer that $N = 1$ is the best option. The resulting geometry is illustrated in Figure (5.23).

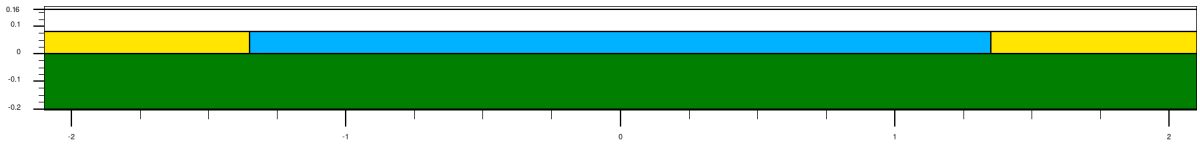


Figure 5.23: The designated ARC implementation with a SiO2 layer of the same thickness as the electrodes.

5.5 Plasmon-enhanced photo-conductive antenna

A more sophisticated enhancement to the basic antenna consists in fabricating electrically floating gold isles at the semiconductor-air interface between the two charged electrodes. Ohmic contacts are assumed at the interfaces between such new metal regions and the semiconductor [LGK+18]. The geometry is depicted in 5.24.

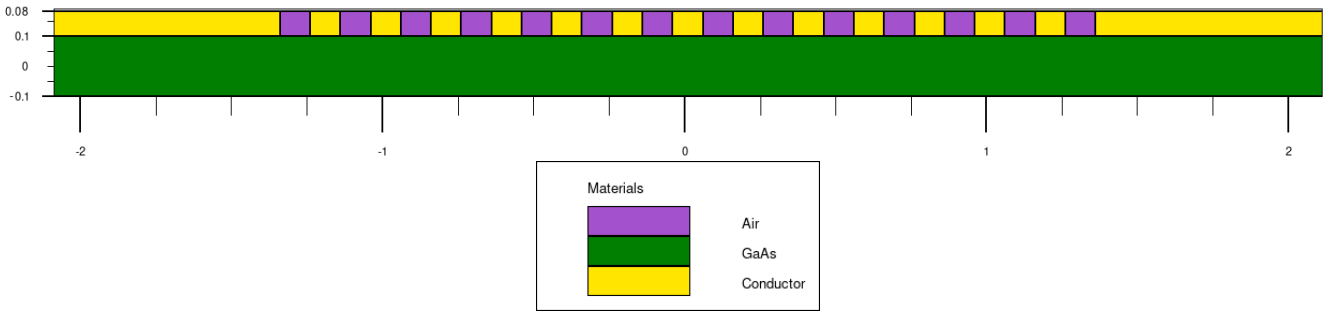


Figure 5.24: The computational domain for the electrostatic simulation of the considered plasmon-enhanced PCA.

In Silvaco Atlas, an electrically floating conductor in (Ohmic) contact with a semiconductor is implemented by including the CURRENT attribute the CONTACT statement. This enforces the following boundary condition on the total electric current of the static PDD system:

$$\sum_{(\bar{x},\bar{y})\in\mathcal{V}} \hat{\mathbf{n}}(\bar{x},\bar{y}) \cdot \mathbf{J}(\bar{x},\bar{y}) = 0, \tag{5.23}$$

where $\hat{\mathbf{n}}$ is the outward unit normal vector to the semiconductor's boundary and $\mathcal{V}$ is the set of the mesh vertices lying on the considered semiconductor-conductor interface. In essence, the net current density flux across the interface is null – to see it from a circuit perspective, the contact is *open*, i.e. no electric current flows across it.

The calculated steady state variables are plotted in Figure 5.25. With respect to Figure 5.6, we notice that the isles (and electric charge distribution inside them) yield small wells of accumulation and reduction of electrons (with respect to the thermodynamic equilibrium value of about 10^{16} cm^{-3}).
In the same regions, Figure 5.26 highlights local peaks of the electrostatic field. As discussed for the basic PCA, peaks in electron distribution and the electrostatic field play a relevant role in the transient operation of the device. For example, we notice that such peaks becomes more significant as x increases, suggesting that this device as well, like the previous ones, is best exploited when the laser spot is close to the anode. That said, we remark that the main motivation for this topology lies in plasmon resonance occurring in the isles during the illumination time window, which is responsible for boosting optical field transmission inside the device and thus optical absorption. The effect is shown in the results that follow.

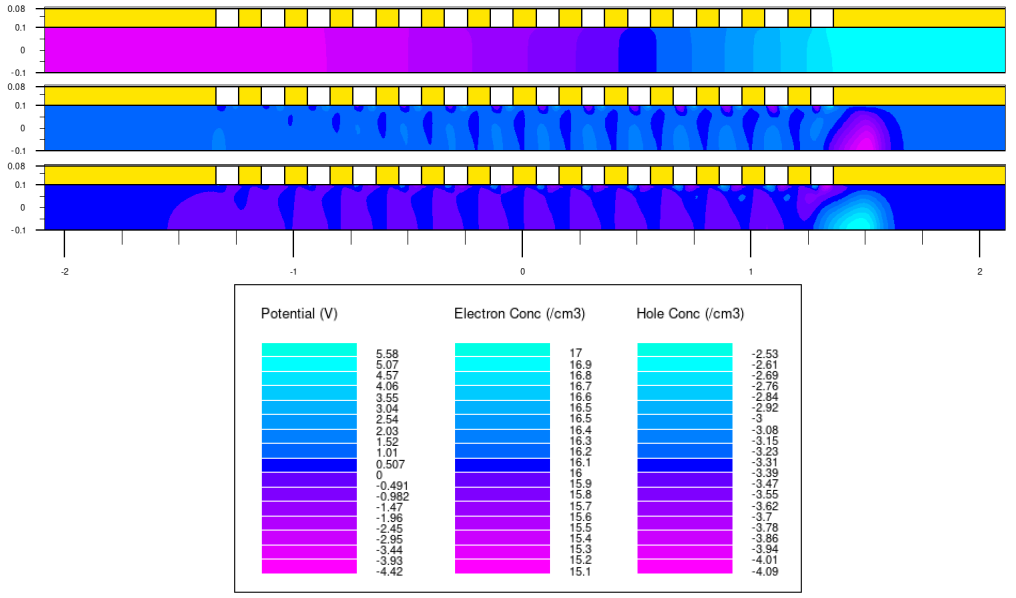


Figure 5.25: From the top to the bottom: the electrostatic potential (linear scale), electron concentration (logarithmic scale), hole concentration (logarithmic scale) calculated in Atlas.

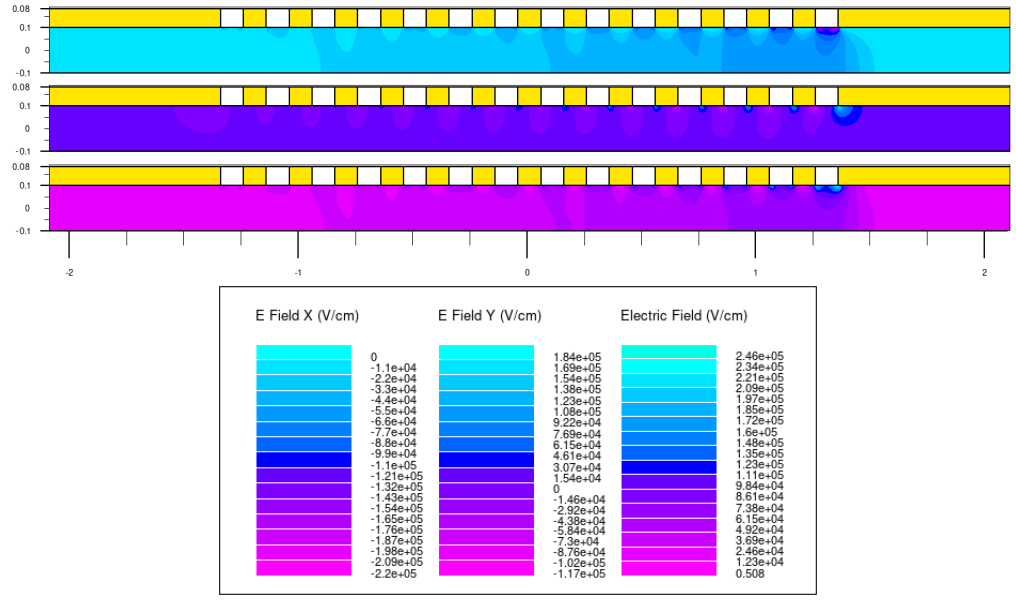


Figure 5.26: $E_{x,DC}$, $E_{y,DC}$ and $||\mathbf{E}_{DC}||$ deduced from the electrostatic potential.

From a numerical standpoint, introducing the isles also brings the requirement of a finer mesh inside and around them, where sharp field variations are expected. Refinement is necessary in the semiconductor too, due to the stronger photo-generation and the above-mentioned peaks in n_{DC}, $\mathbf{E}_{DC}$, that enhance $\mathbf{J}_{n,drift}$ near semiconductor-conductor interfaces. Therefore, at such interfaces and not only in the central part of the device we should expect sharp spatial variations. Beside mesh refinement, the possibility of locally adapting the interpolation degree is exploited as well. An analysis similar to the one carried out in Section 5.4.1 has led to the following setting, with reference to the mesh in Figure 5.27. The mesh size is $h = 27$ nm inside the semiconductor, metallic regions and their proximities, to increases up to $h = 157$ nm as the boundary of the computational domain is approached. Interpolation is based on Lagrange polynomials of

- order 4 inside the semiconductor (10510 elements),
- order 3 inside the isles and the electrodes (5436 elements),
- order 2 in the rest of the computational domain (9306 elements)

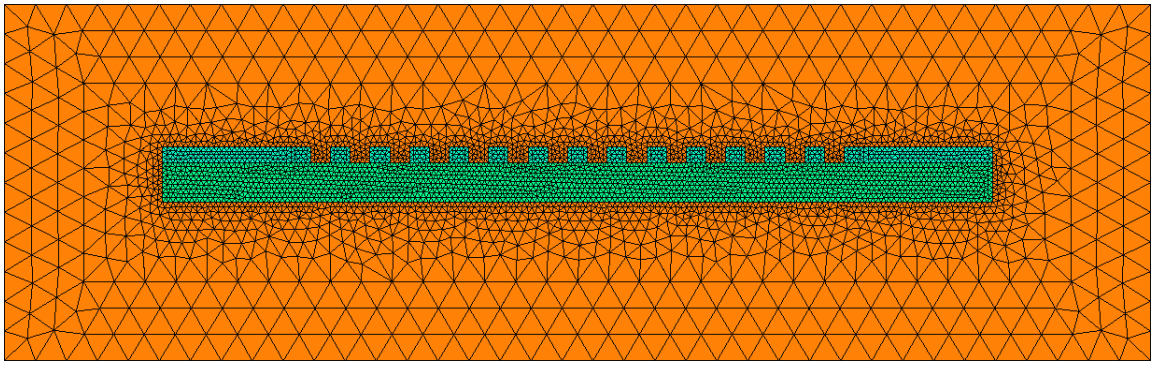


Figure 5.27: The mesh used in the simulation of the plasmon-enhanced PCA.

To appreciate the effect of plasmon resonance we propose in Figure 5.28 a sequence of snapshots along the lines of 5.15. In the scale tailored to the previous PCA, the plasmon-enhanced generation function now shows the presence of a strong near-field distribution in the proximity of the isles; as a result, electron concentration increases in such region. Farther from the isles, the peak value of the initial geometry can be recognized again, but the area in which this value is attained is larger. This phenomenon is consistent with what is observed in a similar structure in [CB20a].

As pointed out when introducing Figure 5.15, a slightly reduced amplitude axis is considered in the snapshots for optimized visualization. To quantify the enhancement due to the introduction of the nanoparticles it is convenient to compare the actual generation peaks (Figure 5.29). The basic PCA allows a photo-generation rate of 7.8×10^{15} cm^{-3} s^{-1} whereas the value reaches 1.2×10^{16} cm^{-3} s^{-1} in the enhanced structure. Hence, G is increased by a factor of about 1.5, which is consistent with the empirically measured enhancement (1.5 to 2) reported in [LGK$^+$17] for a similar structure.

We conclude the discussion with a final remark. Looking at such simulation results, performance comparison between the ARC-enhanced and the plasmon-enhanced antennas might be tempting. In fact, we have considered one possible geometry out of a huge variety that has been fabricated

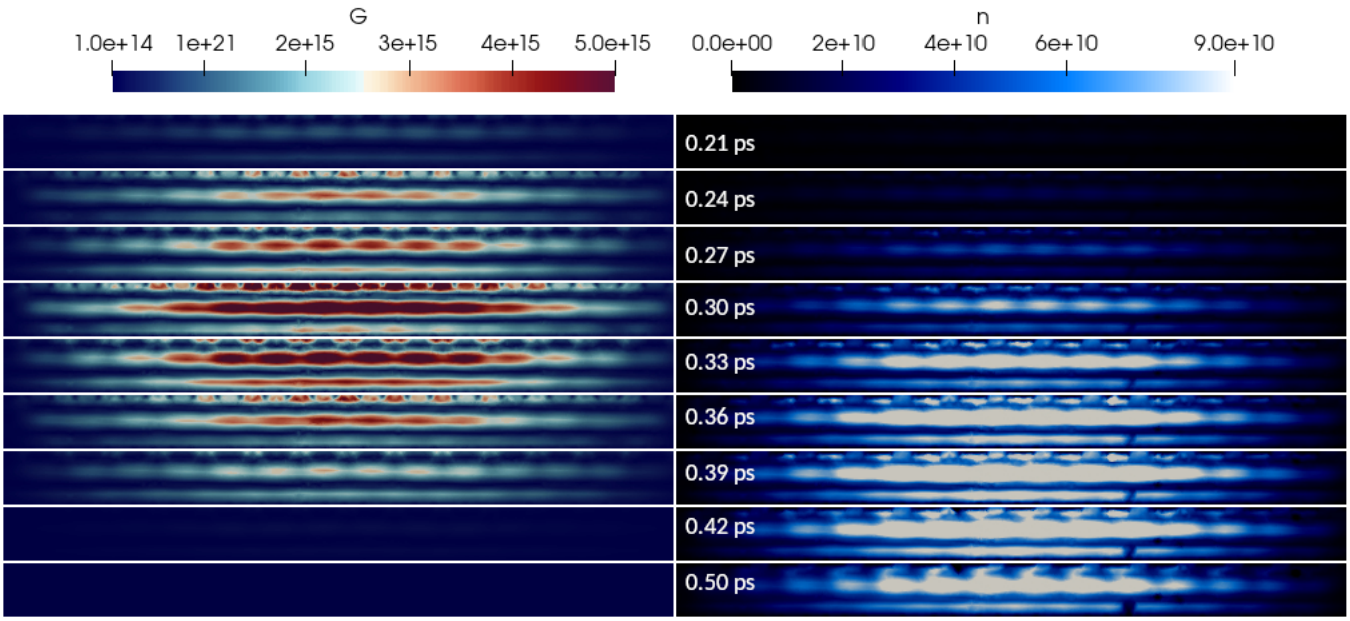


Figure 5.28: Concurrent snapshots of photo-generation (G, cm^{-3} s^{-1}) and electron concentration (n, cm^{-3}), at relevant times in $[0.21, 0.50]$ ps, in the illuminated region of the plasmon-enhanced antenna. Times are marked in ps, in ascending order, from the top to the bottom.

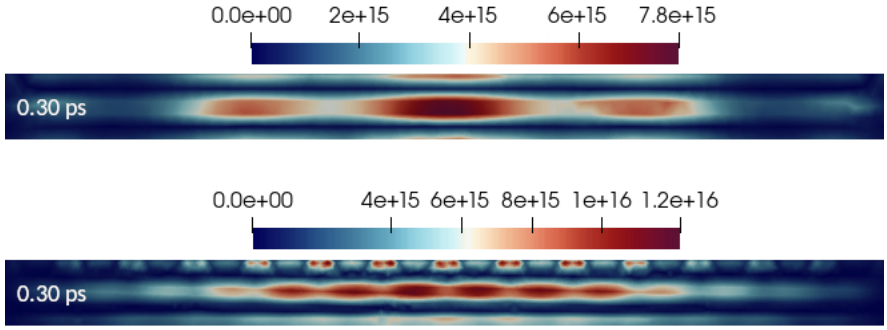


Figure 5.29: Comparison between the distribution of G in the classical (top) and the plasmon-enhanced (top) antenna, evaluated at its peak time.

and proposed in the literature (as documented in [LGK$^+$17]), whereas there is plenty of room for optimization of nanostructures compared to a simple dielectric layer. As a matter of fact, the two methods are compatible and it is not uncommon to find ARCs on the top of a nanoparticle arrays in solar cells [HYL$^+$20]. Regardless, the great deal of research on the subject suggests to consider plasmon resonance as the primary phenomenon to look at for THz PCA optimization.

Chapter 6

Conclusions and perspectives

In this dissertation we have illustrated the conception, the design and the verification of a DGTD solver for the MDD system.

We have started by formally defining the model in interior and exterior problems, detailing its physical meaning, with particular attention to the task of choosing an appropriate photo-generation function. We have coupled the initial system to Drude and Lorentz dispersion models to describe plasmon resonance in nanoparticle-enhanced PCAs and light absorption in the semiconductor. Also, the concept of device biasing has been outlined and has led to elaborating on static and quasi-static PDD models.

Some further mathematical aspects have been investigated, such as the possibility of a preliminary, a priori energy bound, based on existing results on the well-posedness and the asymptotic behavior of the MDD model. The latter has been cast in a form allowing an interesting connection with FV methods and a discussion on handling the numerical flux of the advective part of the equations.

The DG numerical formulation has been extensively developed, in its local/global weak form and matrix form. Major attention has been devoted to the concept of numerical flux and to the exploration of possible routes to define it, such as calculating the advective part along procedures peculiar to FV methods. The resulting semi-discrete formulation has been the object of explicit time integration with Runge-Kutta and Low-Storage Runge-Kutta schemes. To that purpose, a formula generalizing the CFL condition to advection-diffusion problems has been proposed – this kind of analysis has not been found in the literature on optoelectronic simulation.

We have presented an itinerary through the winding route of code verification (tailored to the MDD-Lorentz equations), in which we have confronted the difficulty of considering a physically meaningful setting in tests and provided simulation results that calculate the numerical order of convergence. This is a significantly different approach with respect to the literature on PCA simulation, which typically restricts the attention to design-related topics.

The implemented two-dimensional code has been used for the simulation of realistic geometries: the classical Auston switch and two possible ways to optimize it, namely a dielectric anti-reflective coating and a grating of gold nanoparticles. As a prerequisite, the steady state of a PCA has been calculated in Silvaco Atlas (by solving the static PDD system) and imported into an unstructured mesh for the subsequent DGTD simulation.

We have derived an appropriate numerical setting for PCA simulation, that applies to the MDD

model in general. The proposed procedure appears uncommon with respect to the literature on photo-conductive devices, and essentially relies on the two points: 1) a concrete verification of the applicability of usual characteristic lengths considered in the process of determining the maximum mesh size in a semiconductor; 2) the exploitation of local hp-refinement to tackle the different space scales of the underlying physics.

Finally, simulation results have been shown for the three selected PCAs; their behavior has been illustrated through sequences of snapshots and probes, an optimization study on the anti-reflective case, and the observation of the beneficial effects of plasmon resonance occurring in the nanograting. The latter has proved to significantly enhance photo-generation, consistently with results available in the literature.

Throughout the work we have regularly envisaged future research directions. These are summarized below in more detail, along with further propositions.

6.1 Independent interpolation of Maxwell's and semiconductor equations

When seeking a suitable numerical setting for PCA simulation in Chapter 5.4.1, it was apparent that refinement in the space discretization was ultimately driven by charge carrier dynamics – the same mesh, with quadratic polynomials, was found to suit the electromagnetic field whereas cubic polynomials were required to correctly resolve electron concentration. In consideration of such a difference in space scales, a natural way forward would be to seek different and independent interpolation orders in the local DG formulation, namely $\mathfrak{p}_{EM}$ for $(\mathbf{E}, \mathbf{H})$ and $\mathfrak{p}_{SC}$ for $(n, p, \boldsymbol{\alpha}, \boldsymbol{\beta})$ – an aspect that has not been considered in the literature on MDD equations so far.
Implementation-wise, the cost of the necessary re-interpolation at each time step would have to be evaluated. Indeed, the total current $\mathbf{J}_n + \mathbf{J}_p$ calculated in the $\mathfrak{p}_{SC}$-representation would have to be represented in $\mathfrak{p}_{EM}$ degrees of freedom before being injected into Maxwell's equations. The reverse process would be necessary to inject $\mathbf{E}$ and the polarization current density $\mathbf{J}_l$ (originally calculated in the $\mathfrak{p}_{EM}$-representation) into the DD equations.

6.2 Independent time integration of Maxwell's and semiconductor equations

In the context of DGTD resolution of the MDD model, and particularly of PCAs, the authors in [CB20a] adopt two separate explicit time-integration schemes: Total-Variation-Diminishing (TVD) RK3 for the DD equations with step Δt_{SC} and LSRK45 for the electromagnetic field with $\Delta t_{SC} = N\Delta t_{EM}$ for some integer N. As mentioned at the end of Section 3.1.4, a similar setting was tested in the present work in a test case (Section 4.4.3) but with a RK2 scheme for the whole MDD system. The extension to real device simulation, in which no exact solution is available and the original equations (without fictitious sources) are solved, should require further investigation. The crucial point is how to handle the field-dependent terms $(\mathbf{J}_{n,drift}, \mathbf{J}_{p,drift}, G)$ in the DD equations.
In [CB20a], G is calculated at the rate Δt_{EM} and averaged over the skipped iterations to compensate for the missed updates. Drift terms are not averaged, which practically means that the electric field undergoes a low-pass filter before drifting electrons and holes. In principle, this might be consistent

with some higher-order effects described by hydro-dynamic semiconductor models [GDI+22], which show that n and p may lag behind $\mathbf{E}$, thereby confirming a fundamental difference in the two concerned time scales; however, assigning a value to N in a way that rigorously reproduces such physical behavior appears non-trivial.

6.3 Three-dimensional implementation

The verification process proposed in Section 4.2 can be easily extended to the three-dimensional framework. Integrating a DD module in the DIOGENeS suite would allow to exploit interesting features such as PMLs and near-to-far-field transformation to evaluate the emitted electromagnetic field as usually done in the FDTD literature on PCA simulation. On the other hand, calculating the DC operating point of a device would require a three-dimensional solver for the static PDD model. An interesting option would be to consider a full-wave formulation of the electromagnetic problem based on the wave equation for electric field and the electric potential, which would allow to calculate the steady state asymptotically. The approach is followed in [GSML19] to calculate the step response of a PN junction with continuous finite elements in one dimension.

6.4 Applications

It was pointed out in Section 5.3 that a typical PCA setup entails a large substrate, to be artificially truncated in simulations. This enhancement of the numerical setting would allow to consider arrays of antennas and deploy the developed DGTD solver in the field of optimization. The main obstacle would be to calculate the steady state in a theoretically periodic (infinite) physical domain.
Another interesting step forward would be to increase (e.g. double) the bias voltage to consider high-power scenarios in which the DD model ceases to be acceptable and higher-order moments of the Boltzmann equations need to be taken into account.
Aside from PCAs, solar cells are a natural application of the developed solver; these demand to conceive a broadband generation function G based on multiple Lorentz oscillators (as suggested for example in [CB21b]).
In all the above-mentioned cases, the code could be integrated in an optimization. For example, the nanostructure deposited on the illuminated surface of a PCA has several degrees of freedom – the number of isles, their height, but also the material, which for instance could be dielectric (see [BFH+19]).

6.5 Alternative DGTD schemes

As outlined in (5.18), the steady state of the device introduces terms in drift current that are proportional to the time-varying electric field, but do not depend on the time-varying charge concentration: $n_{DC}\mu_n\mathbf{E}_{AC}$, $p_{DC}\mu_p\mathbf{E}_{AC}$. These act as source terms in the DD equations and introduce optical oscillations in n_{AC}, p_{AC}, which may pull n_{AC} (p_{AC}) below zero in regions of low photogeneration. The is analyzed in the present section, referring with no loss of generality to a classical PCA – the conclusions hold for ARC- and plasmon-enhanced devices.
Figure 6.1 shows a snapshot of n_{AC} at its peak time $t = 0.42$ ps in PCA of Section 5.4.1. Negative values at segments of Γ_N (electrically insulating boundary) are revealed; here, $\hat{\mathbf{n}}\cdot\mathbf{J}_n = 0$ is required,

that is

$$n_{AC}\mu_n\hat{\mathbf{n}}\cdot(\mathbf{E}_{DC}+\mathbf{E}_{AC})+D_n\hat{\mathbf{n}}\cdot\nabla n_{AC}=n_{DC}\mu_n\hat{\mathbf{n}}\cdot\mathbf{E}_{AC}, \tag{6.1}$$

hence n_{AC} must oscillate in accordance with the right-hand side to fulfill the boundary condition. A comparison with Figure 5.6 suggests that the remaining points where $n_{AC}<0$ belong to areas where ∇n_{DC} is particularly large (near electrode wedges and below the anode). Note that restricting the represented range of values would attest that $n_{AC}>0$ in the illuminated region represented in Figure 5.15, thereby legitimating the positive range used therein.

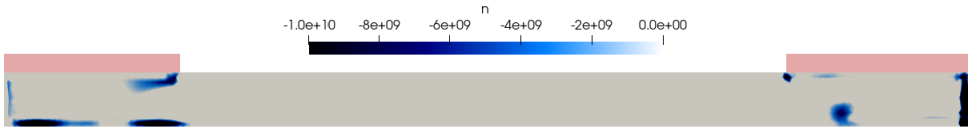


Figure 6.1: Snapshot of n_{AC} at $t=0.42$ ps. The condition $n_{AC}<0$ is observed outside the illuminated region, in particular at segments of electrically insulating interfaces, as well as in areas where spatial variations of n_{DC} are sharpest. Values are expressed in cm^{-3}.

Physically, n_{AC}, p_{AC} represent transient deviations from steady-state concentrations and it is therefore licit for them to be negative, as long as the total concentrations remain positive:

$$\begin{aligned} n_{AC}+n_{DC}&>0,\\ p_{AC}+p_{DC}&>0. \end{aligned} \tag{6.2}$$

It can be verified that $n_{AC}+n_{DC}>0$ is preserved all throughout the simulation ($n_{DC}>10^{15}$ cm^{-3}). On the other hand, with p_{DC} lying between 10^{-4} and 10^{-2} cm^{-3}, $p_{AC}+p_{DC}<0$ occurs. The phenomenon is shown in a snapshot at the peak time $t=0.42$ ps in Figure 6.2, which spots $p_{AC}<-\min_{\Omega} p_{DC}$ at Ohmic contacts. This condition is more restrictive than necessary but allows to clearly identify the interested areas. A more rigorous check on a per-node basis would show that ((6.2)) is actually violated at a subset of points of such interfaces (this will become apparent in the forthcoming Figure 6.4).

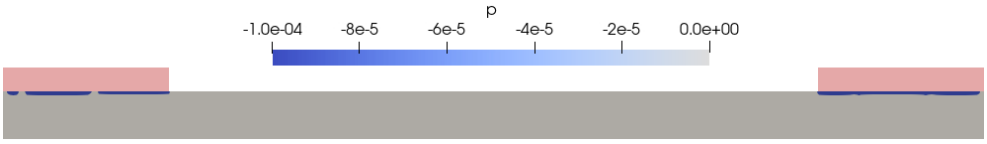


Figure 6.2: Snapshot of p_{AC} at $t=0.42$ ps. The scale is adapted to highlight regions where $p_{AC}<-\min_{\Omega} p_{DC}$. Values are expressed in cm^{-3}.

In practice, it doesn't seem unreasonable to neglect this effect:

- device behavior and performance are almost totally determined by electrons, for which $n_{AC}+n_{DC}>0$ is preserved;
- $p_{AC}<0$ occurs where G is lowest, hence in areas that do not significantly contribute to THz radiation, which results from conductivity modulation and hence from photo-generation.

Nevertheless, the phenomenon is interesting from a numerical standpoint and several tests have been conducted to understand it, with the aim of setting guidelines for future research. The three most relevant ones are summarized below.

- A different dynamic recombination function, R''

  The adopted model (2.18) is defined in [MPG$^+$14], in which a FDTD scheme for the MDD system simulates a conventional PCA in a particular setting: there is no incident field, and optical excitation is buried inside the predictive Gaussian generation function (2.20). As a consequence, no optical oscillation exists in the DD equations, and Maxwell's equations only describe the THz field.
  In this paper, the sign of $p_{AC} + p_{DC}$ is not monitored, but it seems plausible that if $\mathbf{E}_{AC}$ is purely THz and optical-to-THz efficiency is low enough, the term $\nabla \cdot (p_{DC}\mu_p \mathbf{E}_{AC})$ in the hole transport equation will be too weak to entail any violation of (6.2).
  In the first place, in the present work, the condition $p_{AC} + p_{DC} < 0$ was ascribed to the absence of n_{DC} and p_{DC} in (2.18), which led to considering the alternative model

  $$R'' = R(n_{AC} + n_{DC}, p_{AC} + p_{DC}) - R(n_{DC}, p_{DC}), \tag{6.3}$$

  with R from (2.13). However, the new R'' has not significantly changed the behavior of Figure 6.2.

- A purely electromagnetic simulation (Maxwell's equations with no DD equations)

  This has revealed a crucial effect: spikes in the electromagnetic field at $(x, y) = (\pm 1.35, 0.1)\ \mu m$, i.e. at the two points of the top semiconductor surface where air, metal and the semiconductor meet. Further investigation has led to conclude that three sources of instability exist at such points in MDD simulations:

  - the above-mentioned spikes in $\mathbf{E}, \mathbf{H}$;
  - strong spatial variations of $n_{DC}, p_{DC}, \mathbf{E}_{DC}$;
  - spikes in n_{AC}, p_{AC} due to the change in the boundary conditions – from $n_{AC} = 0$ ($p_{AC} = 0$) on Γ_D (semiconductor-metal interfaces) to $\hat{\mathbf{n}} \cdot \mathbf{J}_n = 0$ ($\hat{\mathbf{n}} \cdot \mathbf{J}_p = 0$) on Γ_N (air-semiconductor interfaces).

  For example, Figure 6.4 shows a cutline along the top surface of the semiconductor:

  $$\{p_{AC}(x, y, t) \ : y = 0.1\mu\text{m}, t = 0.42\text{ps}\}.$$

  In the scale tailored to the peak value (the same considered in Figure 5.12), the boundary condition $p_{AC} = 0$ at Ohmic contacts seems fulfilled. However, a zoom reveals that spikes exist at $(x, y) = (\pm 1.35, 0.1)\mu m$ as well as spatial oscillations around 0 along the Ohmic contacts. The key point is the following. The amplitude of such oscillations around 0 is negligible with respect to the values of p_{AC} along Γ_N, i.e. the air-semiconductor interface highlighted in transparent red; however, it largely outnumbers the orders of magnitude spanned by p_{DC}. This observation has inspired the stabilization test that follows.

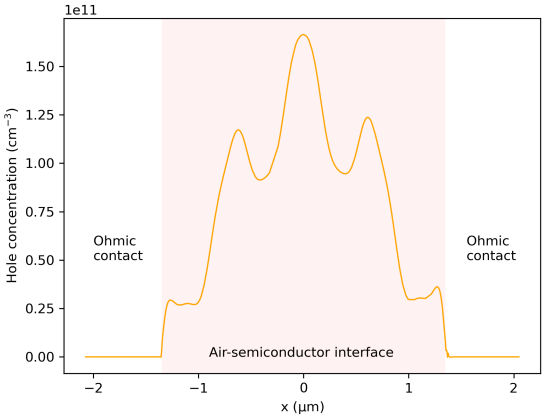


Figure 6.3: The cutline $\{p_{AC}(x, y, t) \ : \ y = 0.1 \ \mu\text{m}, t = 0.42 \text{ ps}\}$ in a scale tailored to the peak value. The boundary condition $p_{AC} = 0$ at Ohmic contacts seems satisfied.

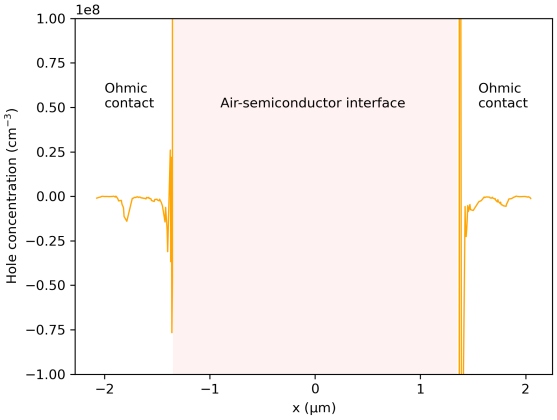


Figure 6.4: The same cutline, under a more suitable scale, shows spikes at $(x, y) = (\pm 1.35, 0.1) \ \mu m$ and oscillations along Ohmic contacts. The amplitude is negligible with respect to the values of p_{AC} in the highlighted region (air-semiconductor interface), but very large compared to p_{DC}.

- A different time-integration scheme: Total-Variation-Diminishing (TVD) Runge-Kutta 3 [SO88]. Known for its property of preventing spurious spatial oscillations, the scheme is considered in [CB20a] for the DGTD resolution of the MDD system, to simulate PCAs. The authors consider the same recombination model (2.18) as [MPG$^+$14], but contrarily to the latter reference the incident field is contemplated and G is calculated from it.
  In [CB20a], however, no formal justification is provided for using this kind of scheme. Also, the sign of n_{AC}, p_{AC} is not discussed. One possible reason is the potential neglect/miss of optical oscillations in n_{AC}, p_{AC} due to considering multi-rate time stepping, with the semiconductor updated every five time steps of the electromagnetic field, i.e. $\Delta t_{SC} = 5\Delta t_{EM}$.
  In the present work, the TVD RK3 option has not yielded significantly different results with respect to the initial LSRK2 implementation – the same spatial oscillations as Figure 6.4 were observed.

In conclusion, the above analysis should suggest to consider device simulation to be minorly affected by the undesired behavior of p_{AC} at Ohmic contacts, but it should also set guidelines for further research. For example, in Section 1.2 we outlined a basic principle behind notable FV schemes for the PDD model: the preservation of physical properties of the continuous problem. Similar criteria would deserve to be investigated in MDD-DG framework. To force (6.2) at all times, a bound-preserving time-integration scheme might be an interesting option. Alternatively, an attempt to reduce the spatial oscillations highlighted in Figure 6.4 might be done by means of an interior penalty or a slope limiter.

Bibliography

[AAP+10] J. Alvarez, L.D. Angulo, M. Fernandez Pantoja, A. Rubio Bretones, and S.G. Garcia. Source and boundary implementation in vector and scalar DGTD. *IEEE Transactions on Antennas and Propagation*, 58(6):1997–2003, 2010.

[ABCM02] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis for discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 5(39):1749–1779, 2002.

[ABL98] F. Assad, K. Banoo, and M. Lundstrom. The drift-diffusion equation revisited. *Solid-State Electronics*, 42(3):283–295, 1998.

[AD06] J.B. Apoung and B. Després. CFL condition and boundary conditions for DGM approximation of convection-diffusion. *SIAM J. Numer. Anal.*, 44(6):2245–2269, 2006.

[Ada99] S. Adachi. *Optical Constants of Crystalline and Amorphous Semiconductors: Numerical Data and Graphical Information*. Springer US, 1999.

[AM76] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.

[ANH+07] M. Awad, M. Nagel, H.Kurz, J.Herfort, and K.Ploog. Characterization of low temperature gaas antenna array terahertz emitters. *Applied Physics Letters*, 91(18):181124, 2007.

[ANS] ANSYS, Inc. *Ansys Lumerical DGTD*. `https://www.ansys.com/products/photonics/dgtd`.

[Atl] Atlantis team (Inria centre at Université Côte d'Azur, Sophia Antipolis, France). *Diogenes, a DG-based software suite for nano-optics*. `https://diogenes.inria.fr`.

[Aug23] P. Auger. Sur les rayons beta secondaires produits dans un gaz par des rayons x. *C.R.A.S.*, 177:169–171, 1923.

[Aug75] P. Auger. Picosecond optoelectronic switching and gating in silicon. *Appl. Phys. Lett.*, 26(101), 1975.

[Bal12] C.A. Balanis. *Advanced Engineering Electromagnetics, 2nd Edition*. Wiley, 2012.

[Bat21] J.I.S. Batista. *Computational Modeling of Black Phosphorus Terahertz Photoconductive Antennas using COMSOL Multiphysics with Photoconductive Antennas using COMSOL Multiphysics with Experimental Comparison against a Commercial LT-GaAs*

Emitter Experimental Comparison against a Commercial LT-GaAs. PhD thesis, University of Arkansas, Fayetteville, Jul 2021.

[BC12a] M. Bessemoulin-Chatard. *Développement et analyse de schémas volumes finis motivés par la préservation de comportements asymptotiques. Application à des modèles issus de la physique et de la biologie.* Theses, Université Blaise Pascal - Clermont-Ferrand II, November 2012.

[BC12b] M. Bessemoulin-Chatard. A finite volume scheme for convection–diffusion equations with nonlinear diffusion derived from the scharfetter–gummel scheme. *Numerische Mathematik*, 121(4):637–670, Aug 2012.

[BCCH19] M. Bessemoulin-Chatard and C. Chainais-Hillairet. Uniform-in-time bounds for approximate solutions of the drift–diffusion system. *Numerische Mathematik*, 141(4):881–916, Apr 2019.

[BES17] N. M. Burford and M. O. El-Shenawee. Review of terahertz photoconductive antenna technology. *Optical Engineering*, 56(1):010901, 2017.

[BFH+19] M. Bashirpour, M. Forouzmehr, S.E. Hosseininejad, M. Kolahdouz, and M. Neshat. Improvement of terahertz photoconductive antenna using optical antenna array of zno nanorods. *Scientific Reports*, 9(1414), 2019.

[BGP19] Arkabrata Bhattacharya, Dipa Ghindani, and S. S. Prabhu. Enhanced terahertz emission bandwidth from photoconductive antenna by manipulating carrier dynamics of semiconducting substrate with embedded plasmonic metasurface. *Opt. Express*, 27(21):30272–30279, 2019.

[Blo70] K. Blotekjaer. Transport equations for electrons in two-valley semiconductors. *IEEE Transactions on Electron Devices*, 17(1):38–47, 1970.

[BMM+05] F. Brezzi, L.D. Marini, S. Micheletti, P. Pietra, R. Sacco, and S. Wang. Discretization of semiconductor device problems (i). In *Numerical Methods in Electromagnetics*, volume 13 of *Handbook of Numerical Analysis*, pages 317–441. Elsevier, 2005.

[BMP89] F. Brezzi, L.D. Marini, and P.Pietra. Numerical simulation of semiconductor devices. *Computer Methods in Applied Mechanics and Engineering*, 75(1):493–514, 1989.

[BPK+19] M. Bashirpour, J. Poursafar, M. Kolahdouz, M. Hajari, M. Forouzmehr, M. Neshat, H. Hajihoseini, M. Fathipour, Z. Kolahdouz, and G. Zhang. Terahertz radiation enhancement in dipole photoconductive antenna on lt-gaas using a gold plasmonic nanodisk array. *Optics and Laser Technology*, 120:105726, 2019.

[BPW+22] J.P. Banon, P. Pelletier, C. Weisbuch, S. Mayboroda, and M. Filoche. Wigner-Weyl description of light absorption in disordered semiconductor alloys using the localization landscape theory. *Physical Review B*, 105(12), 2022.

[Brd12] S. Brdar. *A higher order locally adaptive discontinuous Galerkin approach for atmospheric simulations.* PhD thesis, Universität Freiburg im Breisgau, Dec 2012.

[BSL+07] V. Blickle, T. Speck, C. Lutz, U. Seifert, and C. Bechinger. Einstein relation generalized to nonequilibrium. *Phys. Rev. Lett.*, 98:210601, 2007.

[Bur16] N. M. Burford. *Design, Fabrication and Measurement of a Plasmonic Enhanced Terahertz Photoconductive Antenna*. PhD thesis, University of Arkansas, December 2016.

[But87] J.C. Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods*. Wiley, 1987.

[CB20a] L. Chen and H. Bagci. Multiphysics simulation of plasmonic photoconductive devices using discontinuous galerkin methods. *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, 5:188–200, 2020.

[CB20b] L. Chen and H. Bagci. Steady-state simulation of semiconductor devices using discontinuous galerkin methods. *IEEE Access*, 8:16203–16215, 2020.

[CB21a] L. Chen and H. Bagci. Analysis of screening effects on terahertz photoconductive devices using a fully-coupled multiphysics approach. *J. Lightwave Technol.*, 39(24):7876–7884, Dec 2021.

[CB21b] L. Chen and H. Bagci. A time-domain carrier generation rate model for optoelectronic device simulations. In *2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI)*, pages 1333–1334, 2021.

[CBV+09] A. Crocherie, P. Boulenc, J. Vaillant, F. Hirigoyen, D. Hérault, and C. Tavernier. From photons to electrons: a complete 3D simulation flow for CMOS image sensor. In *IEEE 2009 International Image Sensor Workshop (IISW)*, 2009.

[CC94] Z. Chen and B. Cockburn. Error estimates fora finite element method for the drift-diffusion semiconductor device equations. *SIAM J. Numer. Anal.*, 31(4):1062–1089, 1994.

[CCSH00] Y. Chung, C. Cheon, J.H. Son, and S. Hahn. Fdtd analysis of propagation characteristics of terahertz electromagnetic pulses. *IEEE Transactions on Magnetics*, 36:951–955, 2000.

[CDK87] L.O. Chua, C.A. Desoer, and E.S. Kuh. *Linear and Nonlinear Circuits*. McGraw-Hill, 1987.

[CHF07] C. Chainais-Hillairet and F. Filbet. Asymptotic behaviour of a finite-volume scheme for the transient drift-diffusion model. *IMA Journal of Numerical Analysis*, 27(4):689–716, 02 2007.

[CKS00] B. Cockburn, G.E. Karniadakis, and C.W. Shu. The development of discontinuous Galerkin methods. pages 3–50, 2000.

[COM] COMSOL, Inc. (Burlington, MA). *COMSOL Semiconductor Module User's Guide*.

[CS89] B. Cockburn and C.W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework. *Mathematics of Computation*, 52(186):411–435, 1989.

[CT67] D. Caughey and R. Thomas. Carrier mobilities in silicon empirically related to doping and field. *Proc. IEEE*, 52:2192–2193, 1967.

[CT92] B. Cockburn and I. Triandaf. Convergence of a finite element method for the drift-diffusion semiconductor device equations: the zero diffusion case. *Math. Comp.*, 59(200):383–401, 1992.

[D18] N. Huynh D. *Nonlinear optical phenomena within the discontinuous Galerkin time-domain framework*. PhD thesis, Humboldt-Universität zu Berlin, Jul 2018.

[DDL+13] S. Descombes, C. Durochat, S. Lanteri, L. Moya, C. Scheid, and J. Viquerat. Recent advances on a DGTD method for time-domain electromagnetics. *Photonics and Nanostructures - Fundamentals and Applications*, 11(4):291–302, 2013.

[Dem10] W. Demtröder. *Atoms, Molecules and Photons: An Introduction to Atomic-, Molecular- and Quantum Physics*. Springer Berlin Heidelberg, 2010.

[DFAA12] M.G. Deceglie, V.E. Ferry, A.P. Alivisatos, and H.A. Atwater. Design of nanostructured solar cells using coupled optical and electrical modeling. *Nano Lett.*, 12:2894−−2900, 2012.

[FGJJ95] D.K. Ferry, H.L. Grubin, C. Jacoboni, and A.P. Jauho. *Quantum Transport in Ultrasmall Devices: Proceedings of a NATO Advanced Study Institute on Quantum Transport in Ultrasmall Devices, Held July 17–30, 1994, in II Ciocco, Italy*. B: NATO ASI series. Springer US, 1995.

[FLLO89] L. Fezoui, S. Lanteri, B. Larrouturou, and C. Olivier. Resolution numerique des equations de Navier-Stokes pour un fluide compressible en maillage triangulaire. *INRIA, RR-1033*, inria-00075525, 1989.

[GAF10] G. Gilbert, A. Aspect, and C. Fabre. *Introduction to Quantum Optics, From the Semi-classical Approach to Quantized Light*. Cambridge University Press, 2010.

[GDI+22] A.C. Gungor, M. Doderer, H. Ibili, J. Smajic, and J. Leuthold. Coupled electromagnetic and hydrodynamic semiconductor modeling for terahertz generation. *IEEE Transactions on Magnetics*, 58(9):1–4, 2022.

[GG96] H. Gajewski and K. Gärtner. On the discretization of van roosbroeck's equations with magnetic field. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 76(5):247–264, 1996.

[GM01] P.A. Gremaud and J.V. Matthews. On the computation of steady hopper flows. *Journal of Computational Physics*, 166:63–83, 2001.

[Gob20] A. Gobé. *Méthodes Galerkin discontinues pour la simulation de problèmes multiéchelles en nanophotonique et applications au piégeage de la lumière dans des cellules solaires*. PhD thesis, Université Côte d'Azur, Feb 2020.

[GR91] E. Godlewski and P.A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Springer, 1991.

[GR09] C. Geuzaine and J. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009. `http://gmsh.info`.

[Gra16] Raphaël Grandin. Optique et physique des ondes (partie 4), 2016. `http://www.ipgp.fr/~grandin/Raphael_Grandin_personal_web_page/Teaching.html`.

[GSML19] A. Gungor, J. Smajic, F. Moro, and J. Leuthold. Time-domain coupled full maxwell- and drift-diffusion-solver for simulating scanning microwave microscopy of semiconductors. In *2019 Photonics & Electromagnetics Research Symposium - Spring (PIERS-Spring)*, pages 4071–4077, 2019.

[Hec12] E. Hecht. *Optics*. Pearson, 2012.

[Hes00] K. Hess. *Advanced Theory of Semiconductor Devices*. Wiley, 2000.

[HW02] J.S. Hesthaven and T. Warburton. Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations. *J. Comp. Phys.*, 181(1):186–221, 2002.

[HYL+20] W.J. Ho, H.Y. Yang, J.J. Liu, P.J. Lin, and C.H. Ho. Plasmonic effects of two-dimensional indium-nanoparticles embedded within sio2 anti-reflective coating on the performance of silicon solar cells. *Applied Surface Science*, 508:145275, 2020.

[ISKZ14] O. Isabella, H. Sai, M. Kondo, and M. Zeman. Full-wave optoelectrical modeling of optimized flattened light-scattering substrate for high efficiency thin-film silicon solar cells. *Prog. Photovolt. Res. Appl.*, 22:671–689, 2014.

[JC72] P.B. Johnson and R.W. Christy. Optical constants of the noble metals. *Phys. Rev. B*, 6:4370–4379, 1972.

[Jer96] J.W. Jerome. *Analysis of Charge Transport.* Springer, 1996.

[JJK96] P. Uhd Jepsen, R. H. Jacobsen, and S. R. Keiding. Generation and detection of terahertz pulses from biased semiconductor antennas. *J. Opt. Soc. Am. B*, 13(11):2424–2436, Nov 1996.

[Joc96] F. Jochmann. Existence of weak solutions of the drift diffusion model coupled with maxwell's equations. *Journal of mathematical analysis and applications*, 204(3):655–676, 1996.

[Joc98] F. Jochmann. Uniqueness and regularity for the two-dimensional drift-diffusion model for semiconductors coupled with Maxwell's equations. *J. Diff. Equ.*, 147(2):242–270, 1998.

[Joc99] F. Jochmann. Some Analytical Results Concerning the Drift Diffusion Equations for Semiconductor Devices Coupled with Maxwell's Equations. *International Series of Numerical Mathematics*, 130, 1999.

[Joc01] F. Jochmann. A quasistationary limit and convergence to equilibrium in the drift diffusion system for semiconductors coupled with maxwell's equations. *Differential and Integral Equations*, 14(4):427–474, 2001.

[Joc05] F. Jochmann. *Qualitative behavior of weak solutions of the drift diffusion model for semiconductor devices coupled with Maxwell's equations.* Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Mathematik, 2005.

[Jün95] A. Jüngel. Qualitative behavior of solutions of a degenerate nonlinear drift-diffusion model for semiconductors. *Mathematical Models and Methods in Applied Sciences*, 05(04):497–518, 1995.

[Jün09] A. Jüngel. *Transport Equations for Semiconductors.* Lecture Notes in Physics. Springer, 2009.

[Kan18] M. Kantner. *Modeling and simulation of electrically driven quantum dot based single-photon sources.* PhD thesis, Berlin Institute of Technology, September 2018.

[KCK06] S. Kasap, P. Capper, and C. Koughia. *Springer Handbook of Electronic and Photonic Materials.* Springer-Verlag, 2006.

[KD16] M. Khorshidi and G. Dadashzadeh. Plasmonic photoconductive antennas with rectangular and stepped rods: a theoretical analysis. *J. Opt. Soc. Am. B*, 33(12):2502–2511, Dec 2016.

[KHSB13] N. Khiabani, Y. Huang, Y.C. Shen, and S.J. Boyes. Theoretical modeling of a photoconductive antenna in a terahertz pulsed system. *IEEE Transactions on Antennas and Propagation*, 61(4):1538–1546, 2013.

[KMN21] T. Kürner, D.M. Mittleman, and T. Nagatsuma. *THz Communications: Paving the Way Towards Wireless Tbps.* Springer Series in Optical Sciences. Springer International Publishing, 2021.

[Koe11] M.C. Koenig. *Discontinuous Galerkin methods in nanophotonics.* PhD thesis, Karlsruhe Institute of Technology, July 2011.

[KYI09] P. Kirawanich, S.J. Yakura, and N.E. Islam. Study of high-power wideband terahertz-pulse generation using integrated high-speed photoconductive semiconductor switches. *IEEE Transactions on Plasma Science*, 37(1):219–228, 2009.

[Led21] P.L. Lederer. Finite element methods in computational fluid dynamics, 2021. `https://www.asc.tuwien.ac.at/~schoeberl/wiki/index.php/Philip_lederer`.

[Lee08] Y.S. Lee. *Principles of Terahertz Science and Technology.* Springer New York, 2008.

[LGK+17] S.I. Lepeshov, A. Gorodetsky, A. E. Krasnok, E.U. Rafailov, and P. A. Belov. Enhancement of terahertz photoconductive antennas and photomixers operation by optical nanoantennas. *Laser and Photonics Review*, 11(1), 2017.

[LGK+18] S. Lepeshov, A. Gorodetsky, A. Krasnok, N. Toropov, T. A. Vartanyan, P. Belov, A. Alú, and E. U. Rafailov. Boosting terahertz photoconductive antenna performance with optimised plasmonic nanostructures. *Scientific Reports*, 8(1), 2018.

[LNX+20] S.R. Lynch, N. Nama, Z. Xu, C.J. Arthurs, O. Sahni, and C.A. Figueroa. Numerical considerations for advection-diffusion problems in cardiovascular hemodynamics. *International journal for numerical methods in biomedical engineering*, 36(9):e3378, 2020.

[LS04] Y. Liu and C.W. Shu. Local discontinuous Galerkin methods for moment models in device simulations: formulation and one dimensional results. *Journal of Computational Electronics*, 3:263–267, 2004.

[LS07] Y. Liu and C.W. Shu. Local discontinuous Galerkin methods for moment models in device simulations: performance assessment and two-dimensional results. *Applied Numerical Mathematics*, 57:629–645, 2007.

[LS10a] Y. Liu and C.W. Shu. Error analysis of semi-discrete local discontinuous Galerkin method for semiconductor device simulation models. *Sci. China. Math.*, 53(12):3255–3278, 2010.

[LS10b] Y. Liu and C.W. Shu. Local discontinuous Galerkin methods for high-order time-dependent partial differential equations. *Communications in Computational Physics*, 7(1):1–46, 2010.

[LSV17] S. Lanteri, C. Scheid, and J. Viquerat. Analysis of a generalized dispersive model coupled to a dgtd method with application to nanophotonics. *SIAM Journal on Scientific Computing*, 39(3):A831–A859, 2017.

[Mai07] S.A. Maier. *Plasmonics: Fundamentals and Applications*. Springer US, 2007.

[Mal65] I. H. Malitson. Interspecimen comparison of the refractive index of fused silica∗,†. *J. Opt. Soc. Am.*, 55(10):1205–1209, Oct 1965.

[Map21] Maplesoft, a division of Waterloo Maple Inc.. Maple, 2021.

[Mar86] A. Markowich. *The Stationary Semi-conductor device Equations*. Springer, 1986.

[May13] I.D. Mayergoyz. *Plasmon Resonances in Nanoparticles*. World Scientific series in nanoscience and nanotechnology. World Scientific, 2013.

[MHH+19] A.H.K. Mahmoud, M. Hussein, M.F.O. Hameed, M. Abdel-Aziz, H.M. Hosny, and S.S.A. Obayya. Optoelectronic performance of a modified nanopyramid solar cell. *J. Opt. Soci. Am. B*, 36(2):357–365, 2019.

[MK02] R.S. Muller and T.I. Kamins. *Device Electronics for Integrated Circuits*. Wiley, 2002.

[MPG+14] E. Moreno, M.F. Pantoja, S.G. Garcia, A. Rubio Bretones, and R. Gómes Martin. Time-domain numerical modeling of THz photoconductive antennas. *IEEE Trans. Terah. Sci. Technol.*, 4(4):490–500, 2014.

[Nan91] G.A. Nanz. A critical study of boundary conditions in device simulation. In *Proc. of the 4th International Conference on Simulation of Semiconductor Devices and Processes, Zurich, Switzerland, 12-14 September 1991*, pages 321–328, 1991.

[NPB09] J. Niegemann, W. Pernice, and K. Busch. Simulation of optical resonators using dgtd and fdtd. *Journal of Optics A: Pure and Applied Optics*, 11(11):114015, 2009.

[QMPS00] R. Quaya, C. Moglestuea, V. Palankovskib, and S. Selberherrb. A temperature dependent model for the saturation velocity in semiconductor materials. *Materials Science in Semiconductor Processing*, 3:149–155, 2000.

[Qua13] A. Quarteroni. *Modellistica Numerica per Problemi Differenziali*. Springer Milan, 2013.

[RA15] S. Reif-Acherman. Augustus matthiessen and his contributions to chemistry. *Journal of Chemical Education*, 92(3):467–475, 2015.

[RCN02] J.M. Rabaey, A.P. Chandrakasan, and B. Nikolic. *Digital Integrated Circuits*. Prentice Hall electronics and VLSI series. Pearson Education, 2002.

[Riv08] B. Riviere. *Discontinuous Galerkin methods for solving elliptic and parabolic equations*. SIAM, 2008.

[RV04] E. Rosencher and B. Vinter. *Optoelectronics*. Cambridge University Press, 2004.

[San07] K. Sankaran. *Accurate domain truncation techniques for time-domain conformal methods*. PhD thesis, ETH Zurich, 2007.

[SB14] B. Streetman and S. Banerjee. *Solid State Electronic Devices*. Pearson Education, 2014.

[Sch18] N. Schmitt. *Méthodes Galerkin discontinues pour la simulation et la calibration de modèles de dispersion non-locaux en nanophotonique*. PhD thesis, University of Nice-Sophia Antipolis, Sep 2018.

[Sel84] S. Selberherr. *Analysis and Simulation of Semiconductor Devices*. Springer-Verlag, 1984.

[Sie11] P.H. Siegel. Terahertz pioneer: David h. auston. *IEEE Transactions on Terahertz Science and Technology*, 1(1):6–8, 2011.

[Sil] Silvaco, Inc. (Santa Clara). *Atlas User's Manual - Device simulation software*.

[Smi80] R.M. Smith. Finite element solutions of the energy equation at high peclet number. *Computers and Fluids*, 8(3):335–350, 1980.

[SN06] S.M. Sze and K.K. Ng. *Physics of Semiconductor Devices*. Wiley, 2006.

[SO88] C.W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, 1988.

[SR52] W. Shockley and W.T. Read. Statistics of the recombinations of holes and electrons. *Phys. Rev.*, 87:835–842, 1952.

[ST05] K. Sakai and M. Tani. *Terahertz Optoelectronics*. Springer-Verlag, 2005.

[ST07] B.E.A. Saleh and M.C. Teich. *Fundamentals of photonics; 2nd ed.* Wiley series in pure and applied optics. Wiley, New York, NY, 2007.

[SWM17] H.S. Sehmi, W.Langbein, and E.A. Muljarov. Optimizing the drude-lorentz model for material permittivity: Examples for semiconductors. In *2017 Progress In Electromagnetics Research Symposium - Spring (PIERS)*, pages 994–1000, 2017.

[TH05] A. Taflove and S.C. Hagness. *Computational Electrodynamics, The Finite-Difference Time-Domain Method*. Artech House, 2005.

[Tor09] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics, A Practical Introduction*. Springer Berlin, Heidelberg, 2009.

[UFB+07] P. C. Upadhya, W. Fan, A. Burnett, J. Cunningham, A.G. Davies, E.H. Linfield, J. Lloyd-Hughes, E. Castro-Camus, M.B. Johnston, and H.Beere. Excitation-density-dependent generation of broadband terahertz radiation in an asymmetrically excited photoconductive antenna. *Opt. Lett.*, 32(16):2297–2299, 2007.

[VG06] D. Vasileska and S.M. Goodnick. *Computational Electronics*. Morgan & Claypool Publishers, 2006.

[Viq15] J. Viquerat. *Simulation of electromagnetic waves propagation in nano-optics with a high-order discontinuous Galerkin time-domain method*. PhD thesis, University of Nice-Sophia Antipolis, Dec 2015.

[Viq18] J. Viquerat. Fitting experimental dispersion data with a simulated annealing method for nano-optics applications. *Journal of Nanophotonics*, 12(3):036014, 2018.

[vR50] W.W. van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell Syst. Tech.*, 29:560–607, 1950.

[Wil80] J. Williamson. Low-storage Runge-Kutta schemes. *J. Comput. Phys.*, 35:48–56, 1980.

[YBG+15] J.C. Young, D. Boyd, Stephen D. Gedney, T. Suzuki, and J. Liu. A dgfetd port formulation for photoconductive antenna analysis. *IEEE Antennas and Wireless Propagation Letters*, 14:386–389, 2015.

[YHBJ14] S.H. Yang, M.R. Hashemi, C.W. Berry, and M. Jarrahi. 7.5photoconductive emitters with three-dimensional plasmonic contact electrodes. *IEEE Transactions on Terahertz Science and Technology*, 4(5):575–581, 2014.

[YJ17] N.T. Yardimci and M. Jarrahi. High sensitivity terahertz detection through large-area plasmonic nano-antenna arrays. *Scientific Reports*, 7(42667):2045–2322, 2017.

[ZHA+15] B.Y. Zheng, H.Zhao, A.Manjavacas, M. McClain, P. Nordlander, and N.J. Halas. Distinguishing between plasmon-induced and photoexcited carriers in a device geometry. *Nature Communications*, 6(7797), 2015.

[ZTL+18] J. Zhang, M. Tuo, M. Liang, X. Wang, and H. Xin. Contribution assessment of antenna structure and in-gap photocurrent in terahertz radiation of photoconductive antenna. *J. Appl. Phys.*, 124:053107, 2018.