Maryam Khaksar Ghalati

NUMERICAL ANALYSIS AND SIMULATION OF DISCONTINUOUS GALERKIN METHODS FOR TIME-DOMAIN MAXWELL'S EQUATIONS

Tese de Doutoramento do Programa Inter-Universitário de Doutoramento em Matemática, orientada pelo Professor Doutor Adérito Araújo e pela Professora Doutora Sílvia Barbeiro e apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Setembro 2016

· U C ·

UNIVERSIDADE DE COIMBRA

Numerical Analysis and Simulation of Discontinuous Galerkin Methods for Time-Domain Maxwell's Equations

Maryam Khaksar Ghalati

UC|UP Joint PhD Program in Mathematics

Programa Inter-Universitário de Doutoramento em Matemática

PhD Thesis | Tese de Doutoramento

September 2016

Acknowledgements

I would like to express my utmost gratitude to my supervisors, Professor Adérito Araújo and Professor Sílvia Barbeiro for their invaluable advices, enthusiasm and continuous guidance over the past four years. I sincerely appreciate all their contribution in form of time, ideas, emotional support. I am extremely grateful for all your concerns as well as your support to attend several conferences and schools during this period. This work would not be possible without your help.

I would like to thank the members and administrative staff of Department of Mathematics of University of Coimbra for all their helps and kindness with a particular thanks to Rute Andre, for her willingness to help in any situation.

I am glad to acknowledge the Portuguese Government through FCT/MEC that provide financial support for this PhD under the BD grant SFRH/BD/51860/2012. This thesis work started with an initial grant under the Project PTDC/SAUBEB/103151/2008 at AIBILI. I am thankful to all project members specially Professor Rui Bernardes.

I would like to thank all professors and past and present steering committee of the UC|UP PhD program. I would like to extend my thanks to my past and present colleagues in UC|UP PhD program for all friendly and cheerful moments we had specially during the first year of courses in Coimbra and Porto. Among them all, special thanks to Luis Pinto and Somayeh Gholami for fruitful discussions at early stages of my thesis.

I have been blessed with great friends scattered across the globe who I met mostly during this PhD. Many thanks to them and special thanks to Iranian circle here with whom I didn't fell myself far from home. Particular thanks go to Fatemeh, she is more than a housemate or even a friend here in Coimbra for me.

My special thanks go to my siblings, for all their emotional support, love, encouragement and their belief in little sister over all these years. Thanks Ashraf, Bahram, Azam and Hossein. You and your families are the reason of my smiles and my happiness. I am, and will always be, grateful.

My sincere appreciation goes to my beloved parents, who raised me with love, gave me life's real education, have thought me to follow my dreams and stand for my values and goals and be optimism to any decision I took. Your calm voice, faithful words and lovely looks in our daily Skype conversation gave me the strengths of living thousands of kilometres far from you. No matter what degree I gain, I still should learn many lessons in humanity from you. Thank you Maman and Baba for all you did and still doing for me.

... to my parents

Abdolkhalegh and Kobra

“ The light which shines in the eye is really the light of the heart ... ”

– *Rumi*

Abstract

In this thesis we present a detailed analysis of a fully explicit leap-frog type discontinuous Galerkin (DG) method for the numerical discretization of the time-dependent Maxwell's equations. The study comprehends models capable to deal with anisotropic materials and different types of boundary conditions. Despite the practical relevance of the anisotropic case, most of the numerical analysis present in the literature is restricted to isotropic materials. Motivated by a real application, in the present dissertation we consider a model which encompasses heterogeneous anisotropy, extending the existing theoretical results.

The DG formulation for the spatial discretization is developed in a general framework which unifies the study for different flux evaluation schemes. The leap-frog time integrator is applied to the semi-discrete DG formulation yielding to a fully explicit scheme. The main contribution of this thesis is to provide a rigorous proof of conditional stability and convergence of the scheme taking into account typical boundary conditions, either perfect electric, perfect magnetic or first order Silver-Müller absorbing boundary conditions and for different choices of numerical fluxes. The bounds of the stability region point out not only the influence of the mesh size but also the dependence on the choice of the numerical flux and the degree of the polynomials used in the construction of the finite element space, making possible to balance accuracy and computational efficiency. Under the stability condition, we prove that the scheme is convergent being of arbitrary high-order in space and second order in time. When Silver-Müller boundary conditions are considered we observe only first order convergence in time. To overcome this order reduction we propose a predictor-corrector time integrator which is also analyzed in this dissertation.

We illustrate the stability and convergence properties of the various schemes with numerical tests. The numerical results of our simulations support the theoretical analysis developed along the thesis.

Resumo

Nesta tese apresentamos uma análise detalhada de um método numérico totalmente explícito para as equações de Maxwell dependentes do tempo que combina um esquema de elementos finitos de Galerkin descontínuos para a discretização no espaço com um integrador do tipo *leap-frog* no tempo. O estudo apresentado permite considerar materiais anisotrópicos e diferentes tipos de condições de fronteira. Apesar da relevância prática do caso anisotrópico, a maioria dos trabalhos presentes na literatura restringe a análise numérica ao caso isotrópico. Motivados por uma aplicação real, nesta dissertação consideramos um modelo que compreende simultaneamente o caso anisotrópico e heterogéneo, generalizando os resultados teóricos existentes.

A formulação do método elementos finitos de Galerkin descontínuos para a discretização espacial é desenvolvida num contexto geral que unifica o estudo de esquemas com diferentes fluxos numéricos. O integrador temporal do tipo *leap-frog* aplicado à formulação semi-discreta de elementos finitos conduz a um esquema totalmente explícito. O principal contributo desta tese é a demonstração rigorosa da estabilidade condicional e da convergência do método numérico, tendo em conta as condições de fronteira mais usuais, que incluem as condições impostas no caso dos condutores perfeitos e as condições absorventes de Silver-Müller de primeira ordem e diferentes escolhas para os fluxos numéricos. Os limites da região de estabilidade evidenciam não só a influência do diâmetro da malha espacial, mas também a dependência da escolha do fluxo numérico e o grau dos polinómios usados na construção do espaço de elementos finitos, tornando possível estabelecer um compromisso entre precisão e eficiência computacional. Provamos que, sob a condição de estabilidade, o método é convergente podendo ser de ordem arbitrariamente elevada no espaço e de segunda ordem no tempo. No caso de serem consideradas as condições de fronteira de Silver-Müller, observamos apenas convergência de primeira ordem no tempo. Esta redução de ordem é ultrapassada pela definição de um método preditor-corretor temporal, que também é analisado nesta dissertação.

Ilustramos as propriedades de estabilidade e convergência dos vários esquemas considerados com testes numérico. Os resultados numéricos das simulações efetuadas suportam a análise desenvolvida ao longo da tese.

Table of contents

List of figures

List of tables

Introduction

Maxwell's equations form the complete set of laws governing electromagnetism and, when combined with constitutive relations, describe the effect of material media on the propagation of electromagnetic waves. A given set of boundary conditions along with prescribed initial conditions complete the model. Results about the existence and the uniqueness of solution can be found in *e.g.* [52].

In the last decades there has been a great interest in solving the Maxwell's equations accurately and efficiently in realistic applications because of their relevance in many different areas. The first – and most well known – method for solving Maxwell's equations is the so-called finite difference time-domain (FDTD) scheme proposed by Yee in 1966 [89]. This method uses a staggered grid both in space and time to obtain a second order convergent algorithm, fully explicit, very easy to implement, that has became widely popular [82]. Despite its success, FDTD, like all finite difference methods, is difficult to generalize to irregular domains and unstructured grids.

The most appealing methods which overcome the limitations of the Yee's scheme and other time-domain methods are the discontinuous Galerkin (DG) methods. These methods gather many desirable features such as being able to achieve high-order accuracy and easily handle complex geometries. Moreover, they are suitable for parallel implementation on modern multi-graphics processing units [48]. Local refinement strategies [20, 30] can also be incorporated due to the possibility of considering irregular meshes with hanging nodes and local solutions.

The main goal of this thesis is to study a fully discrete scheme for time-dependent Maxwell's equations that combines nodal DG methods for the spatial discretization [42] with an explicit leap-frog time integrator. One important ingredient for the efficiency of DG methods is the definition of the numerical fluxes, which generalizes ideas from finite volume methods [79]. Two main choices of fluxes are central fluxes and upwind fluxes. In our investigation we consider a general unwinding flux which contains both of these fluxes.

The idea of using a leap-frog time integrator coupled with a DG method was already presented in the literature. In [32] a locally implicit scheme is defined with central fluxes and Silver-Müller absorbing boundary conditions (SM-ABC), and in [56] an implicit scheme is defined with upwind fluxes. Our derivation extends the results in [32] and [56] to a fully explicit in time method for both cases, central fluxes and upwind fluxes and the most typical boundary conditions: perfect electric conductor (PEC), perfect magnetic conductor (PMC) and first order SM-ABC. Moreover, we consider anisotropic material properties in all our analysis and simulations.

Motivated by our application of interest that will be described later, in this thesis we consider a model with a heterogeneous anisotropic permittivity tensor. Most of ocular tissues exhibit form birefringence that is due to their oriented cylindrical structures [91]. Anisotropy could play a role in

biological waveguides [58]. The anisotropy of biological waveguides and, in particular, of retinal optical photoreceptors has been studied and measured by several authors during the last decades by using various techniques in mammalians and also in the human eye [57, 80].

Even though the DG methods have capability to deal with different material properties, most of the formulation and analysis of the DG method is limited to isotropic cases [39, 42] and in some cases dispersive materials [60]. This excludes a wide class of anisotropic materials and related applications [11, 53, 90]. The treatment of anisotropic materials within a DG framework was discussed, for instance, in [16, 32] for central fluxes and in [2, 49] for upwind fluxes. In this thesis we consider dielectric anisotropy in a general upwinding flux formulation in two dimensions (2D) as well as an extension to a three dimensional problem with anisotropy in electric and magnetic material properties.

The outline of this thesis is as follows:

- **Chapter 1** is devoted to introductory concepts related to electromagnetism. We will start with the Maxwell's equations as a fundamental set of equations to formulate the electromagnetic wave propagation, their constitutive relations and the definition of the most common boundary conditions. We describe the electromagnetic wave propagation in dielectric anisotropic materials with a focus on form birefringence, which plays a role on biological anisotropy. In order to put our work into proper perspective, we review the most popular numerical methods for solving Maxwell's equations highlighting the advantages of using the DG method in computational electromagnetics in the last part of the chapter.

- In **Chapter 2** the numerical tools for solving the Maxwell's equations are introduced. We will consider the three dimensional Maxwell's equations with tensorial material properties (anisotropy in permittivity and permeability). Then we introduce the transverse electric mode which is a model in two dimensions. To complete the model we apply typical reflecting boundary conditions such as perfect electric or perfect magnetic and as an absorbing boundary condition, the first order Silver-Müller is applied. We present the DG method for Maxwell's equations in the conservation form taking into account the tensor material properties in the definition of the numerical flux. The semi-discrete DG scheme is coupled with a leap-frog time integrator which employs a centered approximation for time derivatives and the central flux terms and a backward approximation for the upwind flux terms yielding an explicit leap-frog DG scheme that is capable to deal with different flux evolutions, material properties and different boundary conditions.

- **Chapter 3** is devoted to stability analysis of the leap-frog DG solution of Maxwell's equations. We present a rigorous proof of the stability of 2D scheme showing the influence of the mesh size, the choice of the numerical flux and choice of the degree of the polynomials used in the construction of the finite element space as well as the boundary conditions, which can be either perfect electric, perfect magnetic or first order Silver-Müller. This analysis is further extended to three dimensions (3D) model.

- In **Chapter 4**, we assess the convergence properties of the leap-frog DG scheme in 2D and derive the error estimates. The analysis proves that the leap-frog DG scheme is arbitrary

high-order in space and second order in time in case of PEC and PMC boundary condition and first order in case of SM-ABC. In order to overcome the order reduction that occurs when Silver-Müller boundary conditions are considered, a fully explicit predictor-corrector time integrator is proposed and generalized to an iterative predictor-corrector. We will show that the iterative predictor-corrector scheme converges to a second order convergent in time implicit method. The stability analysis of the implicit method is presented.

- **Chapter 5** is dedicated to a detailed numerical study of the proposed leap-frog DG method on triangular meshes for solving two dimensional Maxwell's equations. We examine the stability condition and the error estimate of the method, through numerical experiments for 2D electromagnetic wave propagation in anisotropic media with PEC boundary condition and SM-ABC while considering central and upwind fluxes. We present numerical results which support the stability and convergence theoretical results. In our simulations, we include models with non diagonal and spatially-varying permittivity tensors. The sharpness of the stability region and the high-order convergence property of DG scheme are observed. Besides achieving the temporal order for different types of boundary conditions, the efficiency of the proposed predictor-corrector method to recover the temporal convergency in the case of SM-ABC is visualized. In the last part of this chapter we simulate light scattering in a 2D domain which aims to represent a simple example of light scattering in retina.

- In **Appendix A** we include some useful mathematical tools, namely inverse and trace inequalities, polynomial approximation properties and the discrete Gronwall's lemma that were used in the analysis of the DG methods

Motivation behind this work

The human retina is a complex structure in the eye that is responsible for the sense of vision. It is a part of the central nervous system and it is composed by several layers, namely the outer nuclear layer that comprises the cells bodies of light sensitive photoreceptors cells, rods and cones (see Figure 1).

For many diseases that affect the eye, the diagnosis is not straightforward. The sensitivity of this structure makes medical analysis particularly complicated. Most of the diagnoses are made either by direct observation, with the possible injection of dyes, to enhance certain parts of the organ, or by numbing the eye and directly measuring its inner pressure or thickness. There are a number of eye-related pathologies that can be identified by the detailed analysis of the retinal layers.

The Optical Coherence Tomography (OCT) technique became increasingly popular in the past decades and has been successfully used as a diagnostic tool in ophthalmology [34]. This low coherence interferometry noninvasive technique [78] allows the assessment of the human retina *in vivo* and has been shown to provide functional information from the ocular fundus [81] due to its sensitivity to small variations of the refractive index [1]. The OCT technology's working principle, which is schematically described by Figure 2, is analogous to ultrasound, but it uses light instead of sound to locate subtle differences in the tissue being analyzed [33]. Discontinuities in the refractive index of the tissue give rise to light scattering, with some light backscattered to the detector. Factors such as the shape and size of the scatterer, wavelength of the incident light and refractive index differences

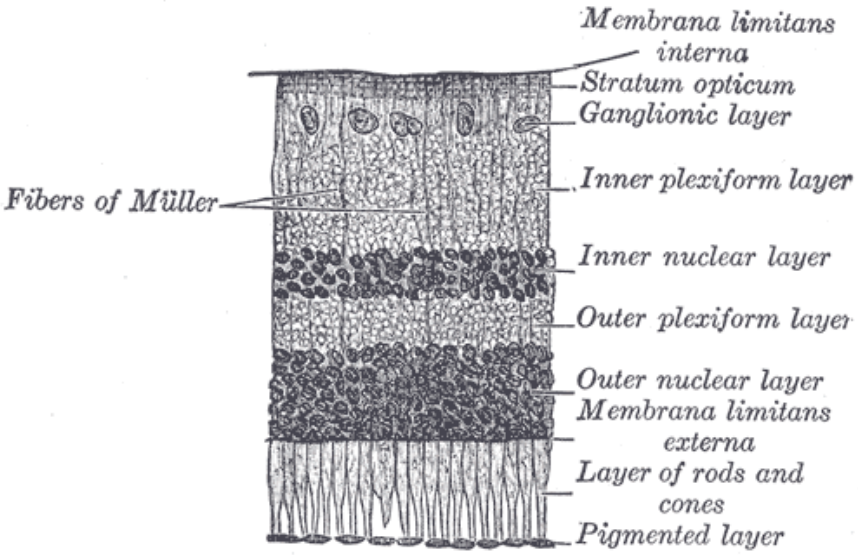


Fig. 1 Section of retina. Henry Gray. Anatomy of the Human Body. Philadelphia: Lea & Febiger, 1918. (in public domain at Bartleby.com)

have an impact on the amount of backscattered light. During a scan, the OCT machine directs a light beam into the retina and extracts, through interferometry, the backscattered light intensity of retinal structures and their depth location in an A-scan. By transversely moving the light beam, several A-scans can be collected into a cross-sectional image – a B-scan. Usually, several cross-sectional images are acquired by probing an azimuthal direction and combined into a volume, see Figure 3.

By analyzing data acquired through OCT, several retinal pathologies, such as diabetic retinopathy, or macular edema, can be detected in their early stages, before noticeable morphologic alterations on the retina [78]. This approach for early diagnosis of retinal conditions is based on functional changes that modify the optical scattering properties of retina, prior to any structural alterations.

Diabetic macular edema (DME), is a major cause of visual loss in diabetic patients [18]. This disease is defined as an increase in retinal thickness due to fluid accumulation that can be intra- or extra-cellular [23]. In intra-cellular edema, cells have increased fluid intake, becoming enlarged. Extra-cellular edema, in contrast, results from fluid accumulation outside the cell, generally as a consequence of the breakdown of the blood-retinal barrier and subsequent leakage into the retinal space. Distinguishing which case is present or more prevalent in a patient's eye at an early stage is usually not straightforward. A common method to assess the progression of DME in patients is to monitor their retinal thickness, *e.g.* with OCT.

In [21] authors proposed methodology to identify and understand possible microscopic changes that lead to the differences in the OCT data between healthy and diseased cases, which are not possible to detect through direct observation. The proposed method combines a light scattering simulation using a Monte Carlo routine with a model of the outer nuclear layer (ONL). This layer was chosen as it consistently presents the characteristics of DME and because it can be adequately modeled by spherical scatterers, which helps to simplify the simulation. By varying the model's parameters, they expected to reproduce the data gathered from healthy and DME eyes and potentially infer which

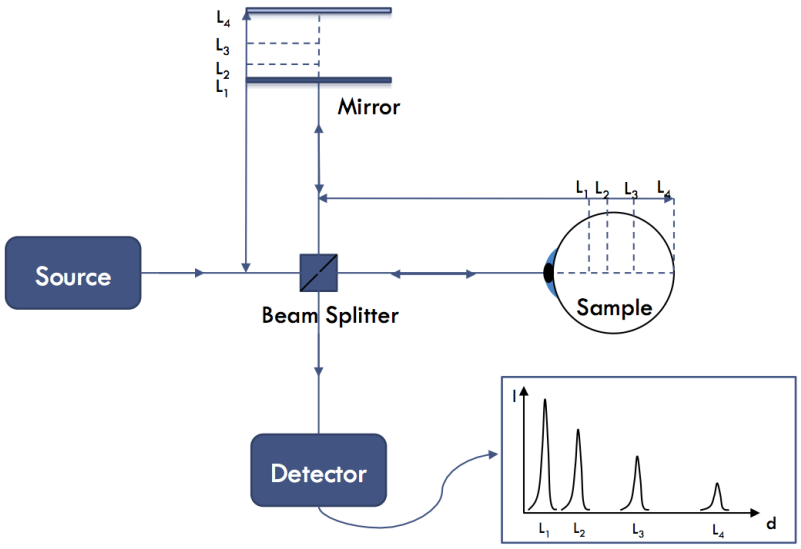


Fig. 2 Scheme for the principle of OCT [78].

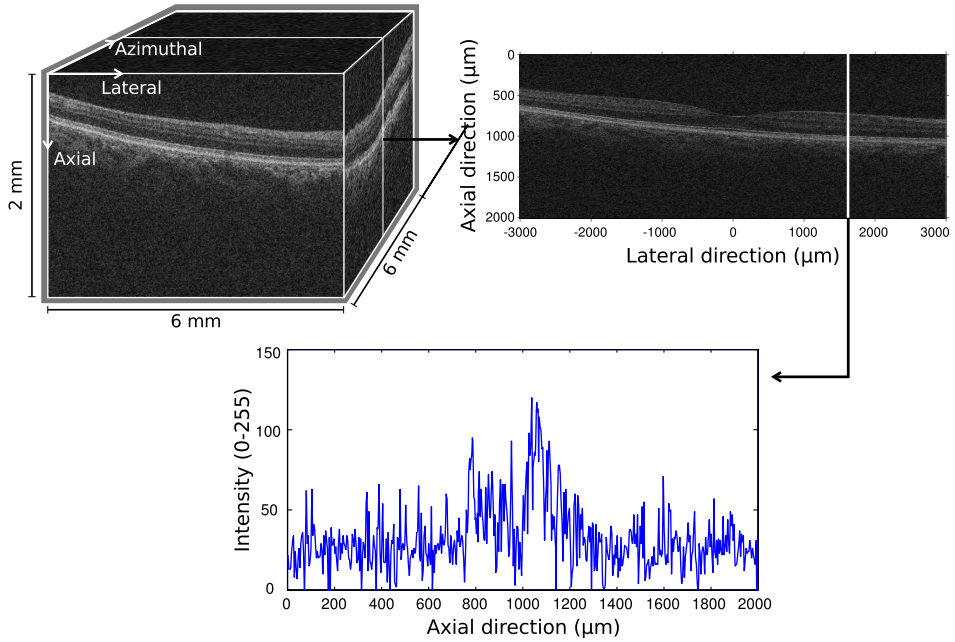


Fig. 3 Example of an OCT volume (top left), B-scan (top right) and A-scan (bottom) for a healthy retina [21].

changes at the cellular level are responsible for the differences in OCT data between the groups studied. As ONL can be modeled as an homogenous medium filled with spherical scatterers, the authors used the Mie theory to estimate the parameters describing the interaction of light with the medium. The results achieved were very promising once they seem to corroborate the existence of the two types of edema, cytotoxic (intra-cellular) and vasogenic (extra-cellular).

To expand the process presented from simulating a single layer to simulating the whole retina will allow the study of more complex retinal pathologies which have pronounced influences outside the ONL. In this context, more sophisticated models to obtain the scattered parameters at cellular level. Several different models have been developed to describe the interactions of the electromagnetic field with biological structures. The first models were based on single-scattering theory, which is restricted to superficial layers of highly scattering tissue in which only single scattering occurs. Accounting for such complexity, in particular to account for how the shape and internal inhomogeneities of each individual biological cell generates its own local optical electromagnetic field structure, requires a more accurate approach that can be achieved by numerically solving Maxwell's equations.

In [7, 73] discontinuous Galerkin numerical solution of Maxwell's equations was considered as an effective and accurate approach for simulation of the complexity of retina, especially the variation of the size and the shape of each layer, distance between them, and respective indexes. In particular, in [73] DG method is employed to solve the Maxwell's equations in retina's layer to obtain the scattered field and calculate the scattered cross-sectioning and anisotropy. This approach was the motivation of our work once it underlies the need of a rigorous analysis of the numerical methods used in the simulations as a first step to validate the results.

Chapter 1

Modeling of Electromagnetic Wave Propagation

Maxwell's equations are a fundamental set of partial differential equations which describe electromagnetic wave interactions with materials. This introductory chapter is started by recalling Maxwell's equations in time-domain and their constitutive relations. Then we discuss the transverse modes of Maxwell's equations in two-dimensions in the last part of Section 1.1. The behavior of the electromagnetic fields at interfaces and boundary conditions is discussed in Section 1.2.

In this thesis we consider electromagnetic wave propagation in dielectric anisotropic media. We shall see that in such media the electric vector of a propagating wave is not in general parallel to its polarization direction, defined by the direction of its electric displacement vector. The preliminaries of electromagnetic wave propagation in anisotropic media with a focus on biological anisotropy are discussed in Section 1.3. We review the existing computational methods for time-domain solution of Maxwell's equations in Section 1.4. We attempt to emphasize on strengths and limitations of some numerical methods when compared with discontinuous Galerkin methods.

1.1 Maxwell's equations

The actual equations that describe electromagnetic phenomena were first completely formulated in 1873 by James Clerk Maxwell. The electromagnetic fields in space is classically described by two field vectors, E and H, called respectively electric field and magnetic field. It is necessary to introduce a second set of vectors, D and B, electric displacement and magnetic induction to include the effect of electromagnetic fields on matter. Maxwell's equations state that these electromagnetic fields are related by two pairs of coupled partial differential equations:

$$\frac{\partial B}{\partial t} = -\nabla \times E, \tag{1.1}$$

$$\frac{\partial D}{\partial t} = \nabla \times H - J, \tag{1.2}$$

$$\nabla \cdot D = \rho, \tag{1.3}$$

$$\nabla \cdot B = 0, \tag{1.4}$$

where ρ and J are the source functions known as scalar charge density and vector current density. The notation $\nabla\times$ and $\nabla\cdot$ is used for vector operator curl and divergence, respectively.

These four equations are the basics of electricity and magnetism in differential form. Equation (1.1) is the differential form of Faraday's law for induction and describes the creation of an induced electric field due to a time-varying magnetic flux. The creation of an induced magnetic field due to charge flow is described by Equation (1.2) known as Ampère's law. The divergence equations (1.3) and (1.4) are Gauss's electric law and Gauss's magnetic law, respectively. Equation (1.3) describes the relation between the electric field distribution and the charge distribution. Equation (1.4) is a statement of the absence of free magnetic monopoles. Faraday's and Ampère's law constitute a first-order hyperbolic system of equations. The two Gauss's laws can be derived from Faraday's and Ampère's laws provided that the initial conditions fulfill the Gauss's laws. Differentiating (1.3) with respect to time and using (1.2) gives

$$\frac{\partial \rho}{\partial t} + \nabla \cdot J = 0, \tag{1.5}$$

which expresses the conservation of the charge of the system. The Equation (1.5) is known as continuity equation. Equations (1.1) and (1.2) are also called curl equations, and equations (1.3) and (1.4) are divergence equations.

For time evolution only the curl equations are important and the divergence equations can be seen as constraints that have to be fulfilled for all times t. It is not hard to see that if the continuity equation (1.5) holds, then from the curl equations it follows that $\nabla \cdot D$ and $\nabla \cdot B$ are constant in time [11]. The charge density ρ and the current density J in equations (1.1)–(1.4) are the source of electromagnetic radiation. In many areas of optics, the propagation of electromagnetic radiation is analyzed in the regions far from the sources where ρ and J can be considered to be zero. This case is assumed in the all following derivation.

Although, the Maxwell's equations (1.1)–(1.4) fully describe the propagation of electromagnetic radiation in any medium, they are not sufficient to determine the electromagnetic field in matter and additional relations known as constitutive equations are needed to model the electromagnetic field interaction with matter.

1.1.1 Constitutive relations

The system of equations (1.1)–(1.4) constitutes the set of governing equations characterizing the behaviour of time and space varying electric and magnetic fields and their interaction with material structures. However, this system remains underdetermined, as long as the relations connecting D to E and B to H are undefined. To allow a unique determination of the field vectors from a given distribution of currents and charges, these equations must be supplemented by relations that describe the effect of the electromagnetic field on material. The relationships connecting these field vectors are called constitutive relations.

Although not being a part of the Maxwell's system, constitutive relations are of great importance to the uniqueness of the field quantities. For a general case, the constitutive relations can be written as,

$$D = D(E), \qquad B = B(H). \tag{1.6}$$

In free space, D differs from E and similarly B differs from H by two respective constants ε_0 and μ_0, which are called the permittivity and permeability of free space respectively. In this case, the constitutive relations (1.6) are written as

$$D = \varepsilon_0 E, \qquad B = \mu_0 H.$$

The values of ε_0 and μ_0 depend on the system of units used. In the standard SI or MKS units we have

$$\varepsilon_0 = 8.85 \times 10^{-12} F m^{-1}, \quad \text{and} \quad \mu_0 = 4\pi \times 10^{-7} H m^{-1}.$$

Furthermore the speed of light in a vacuum, denoted by c_0 is given by $c_0 = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$.

In the case of homogeneous isotropic medium where the physical properties of the medium in the neighbourhood is the same in all directions, the above relationship is given by

$$D = \varepsilon E, \qquad B = \mu H,$$

where $\varepsilon = \varepsilon_r \varepsilon_0$ and $\mu = \mu_r \mu_0$. ε_r and μ_r correspond to relative permittivity and relative permeability, respectively. The permittivity and permeability of the medium, ε and μ are positive, bounded and in the case of inhomogeneous materials, scalar functions of the position.

The electric or magnetic properties of the constituent materials may depend on the direction of the field. These phenomena known as anisotropy, are modeled using tensor permittivity $\varepsilon = [\varepsilon_{ij}]_{3\times 3}$ and tensor permeability $\mu = [\mu_{ij}]_{3\times 3}$ with directional dependence. In anisotropic media, the vector pairs, namely (B,H) and (D,E), are not always parallel and the material tensors, ε and μ are positive-definite matrices that may depend on the position. We will describe the basics of anisotropy later in Section 1.3.

With the assumptions of zero charge density and zero current density, along with the constitutive relations (Equation (1.6)), leads to the Maxwell's system

$$\mu \frac{\partial H}{\partial t} = -\nabla \times E, \tag{1.7}$$

$$\varepsilon \frac{\partial E}{\partial t} = \nabla \times H, \tag{1.8}$$

$$\nabla \cdot H = 0, \tag{1.9}$$

$$\nabla \cdot E = 0. \tag{1.10}$$

If the continuity equation, Equation (1.5) holds, the two divergence equations (1.9) and (1.10) are implicitly satisfied. The electromagnetic wave propagation in such a medium formulated as a set of first order coupled differential equations has the form

$$\varepsilon \frac{\partial E}{\partial t} = \nabla \times H, \tag{1.11}$$

$$\mu \frac{\partial H}{\partial t} = -\nabla \times E, \tag{1.12}$$

where $E = (E_x, E_y, E_z)$, $H = (H_x, H_y, H_z)$, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ are permittivity and permeability tensors. The equations (1.11) and (1.12) are completed with boundary conditions.

1.2 Interfaces and boundary conditions

A complete description of an electromagnetic problem should include complete information about both differential equations and boundary conditions. The typical boundary conditions imposed on the computational domain in an electromagnetic wave propagation simulation are of three types: reflecting, absorbing and periodic boundary conditions [82].

Perfect electric conductor (PEC) and perfect magnetic conductor (PMC) boundary conditions are typical reflecting boundary conditions that reflect all incident radiation and use to model cavities or to introduce symmetry planes into the system [12]. Many problems appearing in computational electromagnetics are posed in unbounded domains. To compute a numerical solution for such problems, it is necessary to truncate the space, by introducing artificial boundaries and regions that define a finite domain. The absorbing boundary conditions mimic open space by absorbing incident radiation in the truncated computational domain. An alternative to absorbing boundary conditions consists in using a perfectly matched layer [9, 10] which is constructed to absorb the electromagnetic waves entering the layer.

In the context of this thesis, we consider reflecting boundary conditions, PEC and PMC and first order Silver-Müller absorbing boundary condition (SM-ABC).

1.2.1 Tangential continuity condition

To solve the Maxwell's equations in the vicinity of boundaries, we shall need conditions relating the field components on either side of the boundary. It can be seen that [44], across the boundary of the domain the tangential components of E and H need to be continuous, that is

$$n \times (E_1 - E_2) = 0, \tag{1.13}$$

$$n \times (H_1 - H_2) = 0, \tag{1.14}$$

where n is the normal unit vector to the boundary and indexes 1 and 2 represent the field component inside and outside of the domain, respectively.

The tangential continuity conditions (1.13) and (1.14) ensure that the tangential component of the field vector is continuous on either side of the boundary regardless of the material. The tangential continuity condition is used to derive flux and interface conditions.

The continuity in normal direction of the fields B and D is achieved from the divergence equations

$$n \cdot (D_1 - D_2) = 0, \tag{1.15}$$

$$n \cdot (B_1 - B_2) = 0, \tag{1.16}$$

which is equivalent to the continuity of the normal component of E and H,

$$n \cdot (\varepsilon_1 E_1 - \varepsilon_2 E_2) = 0, \tag{1.17}$$

$$n \cdot (\mu_1 H_1 - \mu_2 H_2) = 0. \tag{1.18}$$

The equations (1.13)–(1.16) are called the interface conditions.

1.2.2 Perfect electric conductor boundary condition

Perfect electric conductor boundary condition is a reflective boundary condition which is typically used to model a metallic cavity. For a PEC surface the tangential component of the E field goes to zero and there is no field propagation into the PEC medium.

If we consider the case where the material is surrounded by a perfect conductor the interface conditions (1.13)–(1.16) yield the boundary conditions [11]

$$n \times E = 0, \tag{1.19}$$

$$n \cdot B = 0. \tag{1.20}$$

The conditions (1.19) and (1.20) imply that the tangential components of the electric field and the normal component of the magnetic field vanish at the boundary.

1.2.3 Perfect magnetic conductor boundary condition

Perfect magnetic conductor boundary condition is also a reflective boundary condition. For a PMC surface the tangential component of the H field goes to zero and there is no field propagation into the PMC medium. The interface conditions (1.13)–(1.16) yield the PMC boundary conditions as

$$n \times H = 0, \tag{1.21}$$

$$n \cdot D = 0. \tag{1.22}$$

The conditions (1.21) and (1.22) imply that the tangential components of the magnetic field and the normal component of the electric field vanish at the boundary.

1.2.4 Silver-Müller absorbing boundary condition

In unbounded electromagnetic simulations, the computational domain has to be truncated by an absorbing boundary condition to model the infinite space. In this case, when the electromagnetic wave hits the boundary it should not be reflected but absorbed. The effective modeling of waves on unbounded domains by numerical methods is dependent on the particular absorbing boundary condition used to truncate the computational domain [77].

One of the widely used absorbing boundary condition is the first-order accurate Silver-Müller absorbing boundary condition. Applying the Silver-Müller conditions at a finite distance from the scatterer results in an approximate absorbing boundary condition which is exact for outgoing spherical waves [45, 64].

In the case of isotropic material properties, the SM-ABC for Maxwell's equations can be imposed in two different ways as

$$n \times E = c\mu n \times (H \times n), \tag{1.23}$$
$$n \times H = -c\varepsilon n \times (E \times n), \tag{1.24}$$

where n is the unit outward normal vector to the boundary and $c = \frac{1}{\sqrt{\varepsilon\mu}}$ is the speed with which a wave travels along the direction of unit normal.

The SM-ABC depends on the material properties, ε and μ. In order to apply the SM-ABC in the case of anisotropic material properties, we consider, as in [49], the notion of effective permittivity

$$\varepsilon_{eff} = \frac{\det(\varepsilon)}{n^T \varepsilon n}. \tag{1.25}$$

In a similar way, we may define effective permeability as

$$\mu_{eff} = \frac{\det(\mu)}{n^T \mu n}. \tag{1.26}$$

In this case the SM-ABC (1.23) and (1.24) are changed to

$$n \times E = c\mu_{eff} n \times (H \times n), \tag{1.27}$$
$$n \times H = -c\varepsilon_{eff} n \times (E \times n), \tag{1.28}$$

where c is defined with the efficient permittivity and permeability as

$$c = \frac{1}{\sqrt{\varepsilon_{eff}\mu_{eff}}}. \tag{1.29}$$

The conditions (1.23)–(1.24) and (1.27)–(1.28) are based on considering that outside the computation domain, the fields propagate as plane waves normally to the interface. In these equations, the term $n \times (H \times n)$ is the tangential magnetic field and the term $n \times (E \times n)$ is the tangential electric field.

1.3 Wave propagation in anisotropic media

There are many materials whose optical properties depend on the direction of propagation. The history of research on wave propagation in anisotropic material is linked with the development of the history of the theory of elasticity in the early nineteenth century [59]. Wave propagation in anisotropic material is significantly more complex than in isotropic materials. Important anisotropic optical media are crystalline and their optical properties are closely related to various symmetry properties possessed by crystals [88]. Besides crystals and liquid crystals, two other important classes of optically anisotropic dielectric materials exist, optical anisotropy due to form birefringence [11], and optical anisotropy due to the photo-elastic effect or stress birefringence [72]. Optical anisotropy could play a role in biological waveguides, polarimetric fiber-optic sensors, and in mechanical stress sensors. Thus

it is important to have a complete understanding of light propagation in anisotropic media if these phenomena are to be used for practical applications.

In this section, basics of optical anisotropy is reviewed. Among the cases in which optical anisotropy happens, the form birefringence that plays a role in biological anisotropy is described in Subsection 1.3.2 in more detail.

1.3.1 Dielectric tensor of an anisotropic medium

At the macroscopic scale, a dielectric material is optically isotropic if, at any given spatial location in it, its optical properties are the same for any direction [11]. Then at a given spatial location in that medium, there is only one dielectric permittivity (for a given frequency of light) and, hence, only one refractive index of light. Gases, liquids, but not liquid crystals and amorphous solids are the examples of optically isotropic dielectric materials. Various general and specific aspects of the propagation and scattering of the electromagnetic field in optically isotropic materials are well studied and well documented [50, 84].

An optically anisotropic dielectric material is, by definition, one in which, for a given macroscopically small volume element, the optical properties depend on the chosen direction (also, for a given frequency of light), [11]. Then, the dielectric permittivity becomes a 3×3 symmetric tensor and there are more than one refractive indices of light. When the medium is anisotropic, the relation between the electric displacement and the electric field becomes tensorial. Therefore the vectors D and E are no longer parallel. We assumed that the medium is non-dispersive in the frequency range of interest. In this case, a tensorial relation also holds between D and E. On an orthonormal basis (e_1, e_2, e_3), this tensorial relation links the components of the vectors D and E and can be expressed in matrix form according to

$$\begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}. \tag{1.30}$$

These nine quantities $\varepsilon_{xx}, \varepsilon_{xy}, \cdots$ are characteristics of the medium and constitute the dielectric tensor.

In the case of non absorbing media that are magnetically isotropic (i.e. for which $B = \mu H$ where μ is a scalar quantity), the elements of the dielectric tensor are real quantities. Such materials will be considered in this thesis. So the permittivity tensor elements, ε_{ij} are real, dimensionless and functions of position.

An important property of the dielectric tensor that has implications on the propagation of electromagnetic waves in anisotropic media is its symmetry. This means that the permittivity tensor has in general six independent elements, since $\varepsilon_{ij} = \varepsilon_{ji}$ where $i, j = x, y, z$. The demonstration of the symmetry of the dielectric tensors raises some issues in many texts, *e.g.* [83].

There exists an orthonormal basis where the dielectric tensor is represented by a diagonal matrix

$$\bar{\varepsilon} = \begin{pmatrix} \varepsilon_x & 0 & 0 \\ 0 & \varepsilon_y & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}. \tag{1.31}$$

where ε_x, ε_y and ε_z are the eigenvalues of ε. Therefor the dielectric tensor ε becomes diagonal in the eigenvector coordinate system. The directions determined by x, y and z are known as the principal axes of the medium. The refractive indices can be defined along the three principal axes according to

$$r_i = \sqrt{\frac{\varepsilon_i}{\varepsilon_0}},$$

where $i = x, y, z$.

If $r_x = r_y = r_z$ the medium is isotropic. If two out of the three refractive indices are equal, for instance $r_x = r_y \neq r_z$ the anisotropic medium is said to be uniaxial. The $r_x = r_y$ is known as the ordinary index r_o and r_z is the extraordinary index r_e. If $r_o > r_e$ the medium is said to be negative, while it is described as positive when $r_e > r_o$. In more general case where all three eigenvalues are different $r_x \neq r_y \neq r_z$, the anisotropic medium is said to be biaxial. Quartz and calcite are the examples of positive and negative uniaxial crystals. An example of biaxial crystal is topaz.

The phenomenon known as form birefringence which refer the existence of two characteristic waves, behaves as a uniaxial optically anisotropic system. The form birefringence arise from the electrical properties of atoms and molecules which are typically isotropic. Among the cases in which optical anisotropy occurs, (crystals and liquid crystals and photo-elastic effects), form birefringence is described in the following.

1.3.2 Form birefringence

In all analysis and simulations through this thesis we consider the dielectric anisotropy, which motivates from the anisotropy in retinal tissue. The source of anisotropy in retina is its structure that cause a kind of anisotropy known as form birefringence. Form birefringence is directly related to an ordered arrangement of similar particles of optically isotropic material, whose size is large compared with the dimensions of molecules but small in comparison with the wavelength of light.

In 1912, Wiener [86] first showed that a stack of thin, non-absorbing, isotropic dielectric plates would exhibit effective anisotropic dielectric constants when the thickness of the plates and the dimensions of the overall structure were smaller and larger, respectively, than the wavelengths of light. Such birefringence is due to the boundary condition imposed by Maxwell's equations on the electric and displacement field vectors [11].

The form birefringence could arise from ordered structure of parallel fibrils and parallel discs. There is at least one direction along which an incident light has equal propagation velocities in a birefringent material. Such a direction is called an optic axis. An assembly of thin parallel discs behaves as a negative uniaxial crystal with its optic axis perpendicular to the plane of the discs. An assembly of parallel and similar thin cylindrical fibril behaves as a positive uniaxial crystal, with its optic axis parallel to the axes of the fibrils.

Observations on form birefringence are useful in biological waveguides. The sign of the difference between the refractive indexes indicates whether the shape of the particles is nearer to that of a discs or a fibril [11], and if the refractive indexes are known, it may be possible to estimate the fraction of the volume occupied by the particles.

The tissue birefringence includes two types of birefringence: form birefringence, due to ordered arrangement of cellular organelles and intrinsic birefringence, due to anisotropic molecular structure. A

quantitative estimation of the amount of birefringence can achieve with the so-called net birefringence, Δn which is a sum of differences of refractive indices. A non-vanishing value for the net birefringence ($\Delta n \neq 0$) implies the overall optical anisotropy. The net birefringence is the sum of the intrinsic, form and chromic birefringence:

$$\Delta n = \Delta n_I + \Delta n_F + \Delta n_C.$$

The value of form birefringence, Δn_F changes while the refractive index of the medium is varied, but the value of intrinsic birefringence, Δn_I is a fixed. The chromatic birefringence, Δn_C is associated with dichroism [37]. Dichroism is anisotropy in absorption and could be understood by recalling that the components ε_{ij} of the dielectric permittivity tensor could to be complex. The chromatic birefringence, Δn_C could be neglected.

The form birefringence happens in most of the ocular tissues including the cornea, the lens, Helen's fiber at the macula and retinal nerve fiber layer (RNFL) around the optic head. The retinal photoreceptors were modeled as arrays of approximately parallel cylindrical dielectric waveguides in [28]. Similar models will be assumed for the elements of which the retinal nerve fiber layer is composed [91]. The cylindrical shape of photoreceptor cells, their dimensions and the fact that their refractive index is higher than the surrounding medium are the reasons behind their waveguiding properties [28]. A theoretical model developed in [38] shows that an array of thick cylinders with low relative refractive index can produce form birefringence that varies with wavelength. This model is quite suitable for discussing form birefringence of RNFL.

To end up this discussion, measurement of net birefringence can be used as an early diagnosis tool. The net birefringence of RNFL is measured with different technics and groups, *e.g.* [14, 43]. The variation in net birefringence could imply structural changes in tissue. For instance, the RNFL birefringence measurements may provide an early indicator of structural changes caused by glaucoma [43].

1.4 Time-domain methods for solving Maxwell's equations

Several real world electromagnetic problems are not analytically calculable as they often involve large regions with complex geometry of inhomogeneous, anisotropic, lossy and even nonlinear materials. Computational numerical techniques can overcome the inability to derive closed form solutions of Maxwell's equations under various constitutive relations on media and different types of boundary conditions. In order to solve an electromagnetic field problem, one should take into account Maxwell's equations, boundary conditions, all interface and material conditions and all excitation conditions.

The propagation of electromagnetic waves and their interaction with matter is often investigated in a time domain setting. This is the case of the present thesis. In this section we make a brief comparative discussion about the most used numerical methods for time dependent wave problems, highlighting the advantages of the prominent discontinuous Galerkin (DG) as an interesting choice, which is the numerical method employed for space discretization in our derivations.

In the last decades there has been an increasing interest in solving Maxwell's equations because of their great importance and diversity of applications. The finite difference time domain (FDTD), which is also known as Yee's scheme, was introduced in 1966 in [89], and since then it has been applied to a wide range of electromagnetic problems, as for instance, in radar cross section computations and

electromagnetic compatibility investigations. The Yee's scheme continues to be the most widespread computational electromagnetic method in the time-domain [82], mostly because of its simplicity.

The Yee's scheme is an explicit finite difference scheme using central differences on a staggered Cartesian grid. It is second-order accurate in both time and space. The expression staggered grid here indicates that the different electromagnetic components are not computed at the same discretization points in the domain. Furthermore, the field components are not considered on the same time levels. The FDTD method is very efficient for simple geometries, as it gives naturally spurious free solution, is energy conservative and is second-order convergent. The main drawback of the scheme is the inability to represent curved boundaries and small geometrical details. Curved objects must be modeled by staircasing which reduces the accuracy of the scheme to first order [13, 82]. There are several FDTD methods (*e.g* [46, 65]) which effort to model curved objects and remain second order but most of them destroy the simplicity of Yee's scheme.

The finite volume time domain (FVTD) technique emerged as an alternative to FDTD aiming to overcome its geometrical discretization constraints, avoiding the staggered spatial discretization of the fields. The concept of the finite volume method originates in the context of computational fluid dynamics [79]. The most common formulation of FVTD is carried on tetrahedral elements for the Maxwell's curl equations [35, 61]. The scheme is formulated by defining a system of equations in which the time derivative of the electric field vector components integrated in volume equals to the sum of all surface integrals of the spatial derivative of the magnetic field vector components and vice-versa. Since the flux entering a given volume is identical to that leaving the adjacent volume, these methods are local conservative. The time discretization can be performed, in a similarly fashion to the FDTD method, using a second order leap-frog algorithm. Thus the FVTD method it is easily formulated to allow unstructured spatial grids. The main drawback of the FVTD is its order of convergence which is quite low. Moreover, the time step is limited by a condition that depends on the shape of the elements and that is usually more restrictive than for the FDTD method [68].

With the growing need to solve geometrically complex large scale problems, there has been an interest in the flexibility offered by the finite element schemes. Most applications of finite element methods to electromagnetic models were carried out in the frequency domain [45]. Finite element time domain solution of Maxwell's equations received more attention while focusing essentially on low order formulations [51]. These methods offer important advantages over the standard finite difference methods. The use of unstructured grids offers high facility in the modeling of complex geometries. Field and flux continuity conditions at material interfaces can be handled by the variational approach in a natural way [51]. Finite element based methods can handle irregular domains, achieve high order and allow adaptivity and error control. They also use a variational approach which inherits many properties of the continuous problems. In spite of their advantages, finite element methods suffer also from a number of drawbacks. The spatial discretization must be conforming and the mass matrix is global at every time step, which becomes an issue of significant importance in certain problems. While the mass matrix is sparse and typically well conditioned, if we seek for instance a steady solution, the global mass matrix must be inverted. The work associated with this inversion which increases for higher order and large scale problems, is a bottleneck for parallel computations.

The attention to the development of high-order accurate methods for solving time-domain Maxwell's equations in complex geometries brings to the use of a variation of finite element method

called discontinuous Galerkin (DG) finite element methods. Discontinuous Galerkin schemes retain most of the benefits of the finite element methods, as well as local refinement mesh strategies, due to the ability of the method to deal with non-conforming meshes with hanging nodes, and spatial high-order convergence, allowing to deal with problems where the required precision varies over the entire domain, or when the solution lacks smoothness. These advantages come together with other benefits as local conservation and flexibility on the choice of the numerical flux. Moreover, the mass matrix is local rather than global and the method is highly parallelizable. DG methods can be seen as a generalization of finite volume method. Thus the obvious advantage of DG methods over finite volume methods is that higher-order approximation in space can be achieved.

The one-step explicit time integration methods like Runge-Kutta (RK) and leap-frog schemes are computationally efficient per update cycle and easy to implement. The RK approach combined with a DG method was originally published by Cockburn and Shu [19]. Since then, several extensions, such as explicit low-storage Runge-Kutta and the fourth-order explicit, singly diagonally implicit Runge-Kutta [47], have been considered.

Resulting from the coupling of the DG method, for the integration in space, with a time integrator, the DG time domain approach gathers most of the advantages of FDTD, FVTD and finite element time domain methods.

Regarding the interaction of electromagnetic field with biological fields, several approaches have been proposed over the past decades, mostly based on single-scattering theory. The numerical methods for solving Maxwell's equations are providing methods to strengthen the knowledge of cellular-level as well as to accelerate the development of corresponding novel clinical technologies. FDTD is one of the most commonly used numerical methods for solving Maxwell's equations that has been applied to light scattering from cells [26]. The first application of the FDTD to cellular-level biophotonics was reported in [67], wherein visible light interactions with a retinal photoreceptors were modeled for the two-dimensional transverse magnetic (TM) and transverse electric (TE) polarization cases. The DG method, in particular the nodal formulation described in [39], has gained notorious popularity in recent years and it has been extensively used in electromagnetic problems since the first application of the method to Maxwell's equations in 2002 (see [41]). This is the method that we have chosen to use in the present work, motivated by our application of interest. Along the thesis we consider a nodal DG method for the space discretization combined with a leap-frog method for the time integration.

Chapter 2

Leap-frog DG Method

The DG finite element method appears to have been introduced in the framework of neutron transport in 1973 [69], and its first analysis was presented in 1974 in [54]. Important progresses took place in the next two decades, like their extension to conservation laws and development adaptive solution techniques. Since the years 2000 DG methods have become very widely used for solving a large range of problems. Being local methods capable of producing highly accurate numerical solutions, DG methods gather many desirable features over more standard continuous methods. The advantages include their flexibility on the choice of meshes and thus their capacity handle complicated geometries, their potential for error control and mesh adaptation, their possible definition on unstructured meshes, the fact that they are suitable parallelization attenuating their major drawbacks which are high memory requirements and computational cost. A main ingredient in the definition of any DG scheme is the so-called numerical flux, which serves as a connection between the single elements in order to make possible to construct the global numerical solution from all local approximations. The notion of the numerical flux has been taken from finite volume methods, where the numerical flux meets the same purpose, *i.e.* to transport the information from one local cell to another.

This chapter is dedicated to the leap-frog DG scheme for solving Maxwell's equations in anisotropic materials. We start by a 3D system of Maxwell's equations. After rewriting the system of equations in the conservation form we introduce transverse modes of Maxwell's system in Section 2.1. In Section 2.2 the nodal DG approach for solving Maxwell's equations is presented leading to a semi-discrete model where the spatial derivatives are discretized while the temporal derivates remain untouched. Then the temporal derivatives of the semi-discrete scheme are discretized, in a procedure that is known as the method of lines, using a leap-frog scheme which employs a centered approximation for time derivatives and the central flux terms and a backward approximation for the upwind flux terms. We arrive at the fully explicit leap-frog DG scheme in Section 2.4. The scheme is presented for a 2D model and further extended to 3D in the last part of this chapter.

2.1 Maxwell's equations in anisotropic materials

In this section we consider the homogeneous system of Maxwell's equations (1.11) and (1.12) which is completed by considering perfect electric conductor or perfect magnetic conductor boundary conditions or Silver-Müller boundary conditions. The system of equations is formulated as a hyperbolic

system of conservation laws which is later used to develop the discretization approach. The reduction to 2D is discussed in Subsection 2.1.2 and transverse electric mode is chosen as our 2D model.

2.1.1 Three dimensional model

Let us recall the Maxwell's equations (1.11) and (1.12) in three-dimensions

$$\varepsilon \frac{\partial E}{\partial t} = \nabla \times H \quad \text{in } \Omega \times (0, T_f], \tag{2.1}$$

$$\mu \frac{\partial H}{\partial t} = -\nabla \times E \quad \text{in } \Omega \times (0, T_f], \tag{2.2}$$

where $E = (E_x, E_y, E_z)$, $H = (H_x, H_y, H_z)$ and $\Omega \subset \mathbb{R}^3$ is a bounded polyhedral domain. The permittivity and permeability tensors ε and μ are space-dependent.

We assume that the electric permittivity and the magnetic permeability tensors ε and μ are symmetric and uniformly positive definite for almost every $(x, y, z) \in \Omega$, and are uniformly bounded with a strictly positive lower bound, *i.e.*, there are constants $\underline{\varepsilon} > 0$, $\overline{\varepsilon} > 0$ and $\underline{\mu} > 0$, $\overline{\mu} > 0$ such that, for almost every $(x, y, z) \in \Omega$,

$$\underline{\varepsilon}|\xi|^2 \leq \xi^T \varepsilon(x, y, z)\xi \leq \overline{\varepsilon}|\xi|^2, \qquad \underline{\mu}|\xi|^2 \leq \xi^T \mu(x, y, z)\xi \leq \overline{\mu}|\xi|^2, \qquad \forall \xi \in \mathbb{R}^3.$$

The model equations (2.1)–(2.2) must be complemented by proper boundary conditions. Here we consider the most common, either the perfect electric conductor boundary condition , the perfect magnetic conductor boundary condition or the first order Silver-Müller absorbing boundary condition as:

$$\begin{cases} n \times E = 0 & \text{on } \partial\Omega \times (0, T_f] \quad \text{for} \quad \text{PEC}, \\ n \times H = 0 & \text{on } \partial\Omega \times (0, T_f] \quad \text{for} \quad \text{PMC}, \\ n \times E = c\mu_{eff} n \times (H \times n) & \text{on } \partial\Omega \times (0, T_f] \quad \text{for} \quad \text{SM-ABC}, \end{cases} \tag{2.3}$$

where c and μ_{eff} were defined by (1.29) and (1.26) respectively.

The initial conditions

$$E(x, y, z, 0) = E_0(x, y, z) \quad \text{and} \quad H(x, y, z, 0) = H_0(x, y, z) \quad \text{in } \Omega, \tag{2.4}$$

must also be provided.

Maxwell's curl-equations can be reformulated in conservation form. The Maxwell's equations (2.1)–(2.2) in dimensionless units read

$$Q \frac{\partial q}{\partial t} + \nabla \cdot F(q) = 0 \quad \text{in } \Omega \times (0, T_f], \tag{2.5}$$

where the material matrix Q and the state vector q are defined by

$$Q = \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix} \quad \text{and} \quad q = \begin{pmatrix} E \\ H \end{pmatrix},$$

and the flux $F(q) = [F_x(q), F_y(q), F_z(q)]^T$ is given by

$$F_i(q) = \begin{pmatrix} -e_i \times H \\ e_i \times E \end{pmatrix},$$

where e_i signifies Cartesian unit vectors in i-direction where $i = x, y, z$. The 6×6 material matrix Q depends on ε and μ the material properties of the medium. The state vector q is the super-vector of the electric field $E = (E_x, E_y, E_z)$, and magnetic field $H = (H_x, H_y, H_z)$ and these are functions of (x, y, z, t).

In mathematical terms, the system (2.5) is said to be hyperbolic [55]. The physical meaning of this mathematical definition is that, the system has wavelike solution. This means that the material matrix Q is diagonalizable with real eigenvalues.

2.1.2 Reduction to two dimensions

If we suppose that the system of Maxwell's equations has some symmetries, it is possible to reduce the dimensions of the system. The system is often homogeneous in one direction. The structure being modeled extends to infinity in the z-direction with no change in the shape or position of its transverse cross section. If the incident wave is also uniform in the z-direction, then all z-derivatives will vanish [82]. Under these conditions the full set of Maxwell's curl equations given by (2.1) and (2.2) reduce to two decoupled sets of three equations.

The set of first three equations contains E_x, E_y and H_z, and it is called transverse electric mode. TE mode describes the propagation where the electric field lies in the plane of propagation. The set of the other three equations contains H_x, H_y and E_z components and it is called transverse magnetic (TM) mode. TM mode describes the propagation where the electric field is perpendicular to the plane of propagation. The TE and TM modes are decoupled since they do not contain any common field vector components. These two modes are completely independent for structures that composed of isotropic or anisotropic materials [82]. The TM and TE modes constitute the two possible ways that two-dimensional electromagnetic wave interaction problems can be set up for case of zero partial derivatives in the z-direction.

If the curl equations (2.1) and (2.2) are written component wise the reduction to TE mode is

$$\varepsilon_{xx}\frac{\partial E_x}{\partial t} + \varepsilon_{xy}\frac{\partial E_y}{\partial t} = \frac{\partial H_z}{\partial y} \tag{2.6}$$

$$\varepsilon_{yx}\frac{\partial E_x}{\partial t} + \varepsilon_{yy}\frac{\partial E_y}{\partial t} = -\frac{\partial H_z}{\partial x} \tag{2.7}$$

$$\mu\frac{\partial H_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{2.8}$$

Since we are dealing with anisotropic permittivity tensor, TE mode (2.6)–(2.8) is considered for our 2D model. This assumption is appropriate when studying *e.g.* truly 2D photonic crystals [24] or the electrodynamic properties of 2D materials like graphene [62].

In the major part of this thesis we shall analyze 2D time-domain Maxwell's equations in TE mode. For this case, and assuming no conductivity effects, the equations (2.6)–(2.8) in the non-dimensional

form are

$$\varepsilon \frac{\partial E}{\partial t} = \nabla \times H \quad \text{in } \Omega \times (0, T_f], \tag{2.9}$$

$$\mu \frac{\partial H}{\partial t} = -\text{curl } E \quad \text{in } \Omega \times (0, T_f], \tag{2.10}$$

where $E = (E_x, E_y)$ and $H = (H_z)$. These equations are set and solved on the bounded polygonal domain $\Omega \subset \mathbb{R}^2$. Note that we use the following notation for the vector and scalar curl operators

$$\nabla \times H = \left(\frac{\partial H_z}{\partial y}, -\frac{\partial H_z}{\partial x} \right)^T, \quad \text{curl } E = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}.$$

The electric permittivity of the medium, ε and the magnetic permeability of the medium μ are varying in space, being ε an anisotropic tensor

$$\varepsilon = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} \\ \varepsilon_{yx} & \varepsilon_{yy} \end{pmatrix}, \tag{2.11}$$

while we consider isotropic permeability μ.

We assume that the electric permittivity tensor ε is symmetric and uniformly positive definite for almost every $(x, y) \in \Omega$, and it is uniformly bounded with a strictly positive lower bound, *i.e.*, there are constants $\underline{\varepsilon} > 0$ and $\overline{\varepsilon} > 0$ such that, for almost every $(x, y) \in \Omega$,

$$\underline{\varepsilon} |\xi|^2 \leq \xi^T \varepsilon(x, y) \xi \leq \overline{\varepsilon} |\xi|^2, \qquad \forall \xi \in \mathbb{R}^2.$$

We also assume that there are constants $\underline{\mu} > 0$ and $\overline{\mu} > 0$ such that, for almost every $(x, y) \in \Omega$,

$$\underline{\mu} \leq \mu(x, y) \leq \overline{\mu}.$$

The model equations (2.9) and (2.10) is completed by boundary conditions, PEC, PMC or SM-ABC as:

$$\begin{cases} n \times E = 0 & \text{on } \partial\Omega \times (0, T_f] \quad \text{for PEC}, \\ n \times H = 0 & \text{on } \partial\Omega \times (0, T_f] \quad \text{for PMC}, \\ n \times E = c\mu n \times (H \times n) & \text{on } \partial\Omega \times (0, T_f] \quad \text{for SM-ABC}, \end{cases} \tag{2.12}$$

where

$$c = \sqrt{\frac{n^T \varepsilon n}{\mu \det(\varepsilon)}},$$

is the speed with which a wave travels along the direction of the unit normal n. The role of effective permittivity (1.25) is observed in defining c.

The initial conditions

$$E(x, y, 0) = E_0(x, y) \quad \text{and} \quad H(x, y, 0) = H_0(x, y) \quad \text{in } \Omega, \tag{2.13}$$

must also be provided.

2.2 Space discretization with discontinuous Galerkin method

In this section, the spatial discretization of Maxwell's equations in conservation form with nodal DG method is presented. The idea of DG method, like finite element methods, is to construct finite dimensional function spaces in which we search for an approximate solution. The first step in constructing a discrete function space is to discretize the domain. The concepts of the mesh and the broken polynomial space, which are known as discontinuous Galerkin function space are explained in Subsection 2.2.1.

The electromagnetic fields are expanded in terms of a set of basis functions which are defined locally. Let us mention that we follow the general DG method structure as in [39]. We consider the common choice of the basis functions in DG method which are Lagrange polynomials. In Subsection 2.2.3 we start by looking at the problem locally and working on one element. In order to arrive at global numerical approach, the elements are connected via numerical flux. The concepts of the numerical flux are explained in the of Subsection 2.2.3. In the last part of this section, we back to boundary conditions and the discretization of boundary conditions within DG approach is explained.

2.2.1 Definition of the mesh

Let Ω be a bounded polygonal region of $\mathbb{R}^2$ for which the boundary is $\partial\Omega$, where the numerical solutions of equations (2.9) and (2.10) are intended to be computed. Assume that the computational domain Ω is partitioned into K triangular elements T_k such that

$$\overline{\Omega} = \cup_k T_k,$$

where T_k is a straight-sided triangle. For simplicity, we consider that the resulting mesh $\mathscr{T}_h$ is conforming, that is the intersection of two elements is either empty, an edge or vertices.

Let h_k be the diameter of the triangle $T_k \in \mathscr{T}_h$, and h be the maximum element diameter,

$$h_k = \sup_{P_1,P_2 \in T_K} \|P_1 - P_2\|, \quad h = \max_{T_k \in \mathscr{T}_h} \{h_k\}.$$

We assume that the mesh is regular in the sense that there is a constant $\tau > 0$ such that

$$\forall T_k \in \mathscr{T}_h, \quad \frac{h_k}{\tau_k} \leq \tau, \tag{2.14}$$

where τ_k denotes the maximum diameter of a ball inscribed in T_k.

Let ν_k be the set of indices of the neighboring elements of T_k. For each $i \in \nu_k$, we consider the internal edge $f_{ik} = T_i \cap T_k$, and we denote by n_{ik} the unit normal oriented from T_i towards T_k. The boundary edge is $f_k = T_k \cap \partial\Omega$. The unitary outer normal vector to f_k is denoted by n_k. The set of internal edges denote by F^{int} and F^{ext} is the set of edges that belong to the boundary $\partial\Omega$.

By now we have constructed a mesh on the domain Ω and so we can turn to the second step, namely to the construction of the discrete finite element space. The discontinuous Galerkin finite

element space is chosen as a space consisting of piecewise polynomial functions. The finite element space is then taken to be

$$V_N = \{v \in L^2(\Omega)^3 : v|_{T_k} \in P_N(T_k)^3\}, \tag{2.15}$$

where $P_N(T_k)$ denotes the space of polynomials of degree less than or equal to N on the element T_k.

Let us define the local inner product and norm at each element T_k as

$$(u,v)_{T_k} = \int_{T_k} u(x,y).v(x,y)dxdy \quad \text{and} \quad (u,u)_{T_k} = \|u\|^2_{T_k},$$

and also the local inner product over the boundary ∂T_k

$$(u,v)_{\partial T_k} = \int_{\partial T_k} u(s).v(s)ds.$$

These local inner product and norm form the global broken measure as

$$(u,v)_\Omega = \sum_k (u,v)_{T_k}, \quad \|u\|^2_\Omega = \sum_k \|u\|^2_{T_k} \quad \text{and} \quad (u,v)_{\partial\Omega} = \sum_k (u,v)_{\partial T_k}.$$

2.2.2 Local approximate solution

On each element T_k, the solution fields are approximated by polynomials of degree less than or equal to N. The global solution $q(x,y,t)$ is then assumed to be approximated by the piecewise N order polynomials $\tilde{q}(x,y,t)$ defined as the direct sum of the K local polynomial solutions

$$q(x,y,t) \simeq \tilde{q}(x,y,t) = \bigoplus_{k=1}^{K} \tilde{q}_k(x,y,t),$$

where $\tilde{q}_k(x,t) = (\tilde{E}_{x_k}, \tilde{E}_{y_k}, \tilde{H}_{z_k})$. We use the following notation

$$\tilde{E}_x(x,y,t) = \bigoplus_{k=1}^{K} \tilde{E}_{x_k}(x,y,t), \quad \tilde{E}_y(x,y,t) = \bigoplus_{k=1}^{K} \tilde{E}_{y_k}(x,y,t), \quad \tilde{H}_z(x,y,t) = \bigoplus_{k=1}^{K} \tilde{H}_{z_k}(x,y,t).$$

The fields are expanded on each element T_k in terms of interpolating Lagrange polynomials $L_i(x,y)$ as [41],

$$\tilde{q}_k(x,y,t) = \sum_{i=1}^{N_p} \tilde{q}_k(x_i,y_i,t)L_i(x,y) = \sum_{i=1}^{N_p} \tilde{q}_{k_i}(t)L_i(x,y). \tag{2.16}$$

Here N_p denotes the number of coefficients that are utilized and $\tilde{q}_{k_i}(t) = \tilde{q}_k(x_i,y_i,t)$ serves as a short-hand notation for expansion coefficients. Since the expansion coefficients correspond to the field values at the nodes the representation of the fields as Equation (2.16) gives the so-called nodal representation.

The number of coefficients N_p is related with the polynomial order N via

$$N_p = \frac{(N+1)(N+2)}{2}.$$

Note that N_p also represents the number of interpolation node points per element and the number of basis functions.

The polynomial interpolation nodes in a triangle are discussed in [15, 40]. The equidistant points are the first choice, but they have large Lebesque constants and in the context of DG methods they lead to very ill conditioned linear systems [39]. The recent warp-blend method [85] performs sufficiently well for generating two dimensional point sets in triangles. The warp-blend points can be viewed as a two dimensional generalization of Legendre-Gauss-Lobatto points. These points coincide with the Legendre-Gauss-Lobatto points at the edges of the triangle. In one dimensional DG formulation, the Legendre-Gauss-Lobatto quadrature points are chosen as the set of optimal nodal points in [39]. For higher dimensions the warp-blend points are chosen as the set of optimal nodal points in nodal DG structure.

2.2.3 The DG semi-discrete form

Let's start by writing the 2D model (2.9) and (2.10) in a conservation form (2.5)

$$Q\frac{\partial q}{\partial t} + \nabla \cdot F(q) = 0 \quad \text{in } \Omega \times (0, T_f], \tag{2.17}$$

where the material matrix Q and the state vector q are defined by

$$Q = \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix}, \quad q = \begin{pmatrix} E_x \\ E_y \\ H_z \end{pmatrix},$$

respectively and the flux is given by

$$F(q) = \begin{pmatrix} 0 & H_z & E_y \\ -H_z & 0 & -E_x \end{pmatrix}^T.$$

In order to arrive at the semi-discrete scheme, let's start by noticing that the local residual, *i.e.* the error when $\tilde{q}_k$ is substituted in Equation (2.17), is required to vanish in the following way

$$\int_{T_k} \left(Q\frac{\partial \tilde{q}_k}{\partial t} + \nabla \cdot F(\tilde{q}_k) \right) \cdot v \, dxdy = 0, \tag{2.18}$$

for all test function $v \in V_N$. This can be recognized as a Galerkin approach, but on the local element only. Integration by parts once yields

$$\int_{T_k} \left(Q\frac{\partial \tilde{q}_k}{\partial t} \cdot v - F(\tilde{q}_k) \cdot \nabla v \right) dxdy = -\int_{\partial T_k} n \cdot F(\tilde{q}_k) \cdot v \, ds,$$

where n is the outward pointing unit normal vector of the contour.

The solution at the interfaces between elements is multiply defined and we will need to choose which solution, or combination of solutions, is correct. There are several possible choices for this issue. Delaying the detail of this choice, we substitute in the resulting contour integral the flux F by a numerical flux F^* that enforces a physically correct solution. With the numerical flux introduced on the right hand side to connect the elements, doing integration by parts once more yields the final form

$$\int_{T_k} \left(Q\frac{\partial \tilde{q}}{\partial t} + \nabla \cdot F(\tilde{q}_k) \right) \cdot v \, dxdy = \int_{\partial T_k} n \cdot (F(\tilde{q}_k) - F^*(\tilde{q}_k)) \cdot v \, ds. \tag{2.19}$$

Equation (2.19) is the strong variational formulation of Maxwell's curl equations. The left-hand side of Equation (2.19) remains a local expression, while in the right-hand side the numerical flux connects the neighbouring elements through their common edge.

In order to connect the solution between elements sharing a common edge, the continuous numerical flux of the tangential field components are defined at each interface. The concept of numerical flux in computational electromagnetic was inspired by finite volume time domain methods [63, 79]. Upwind flux evaluation is the usual way to exchange information between elements in finite volume methods. A discontinuous Galerkin method with polynomial order zero is nothing else than a finite volume method.

The proper choice of the numerical flux is essential for the accuracy of the DG scheme. The derivation of the numerical flux involves the solution of a Riemann problem and is discussed in detail for isotropic materials in [39]. This flux is given by the expression

$$n \cdot (F(\tilde{q}) - F^*(\tilde{q})) = \begin{pmatrix} \frac{-1}{Z^+ + Z^-} n \times \left(Z^+ [\tilde{H}] - \alpha n \times [\tilde{E}] \right) \\ \frac{1}{Y^+ + Y^-} n \times \left(Y^+ [\tilde{E}] + \alpha n \times [\tilde{H}] \right) \end{pmatrix}. \tag{2.20}$$

Here we need to introduce the notation for the jumps of the field values across the interfaces of the elements,

$$[\tilde{E}] = \tilde{E}^- - \tilde{E}^+, \quad \text{and} \quad [\tilde{H}] = \tilde{H}^- - \tilde{H}^+, \tag{2.21}$$

where the superscript "$+$" denotes the neighboring element and the superscript "$-$" refers to the local cell. Furthermore the cell-impedances and cell-conductances respectively are

$$Z^{\pm} = \sqrt{\frac{\mu^{\pm}}{\varepsilon^{\pm}}}, \quad \text{and} \quad Y^{\pm} = \left(Z^{\pm} \right)^{-1}.$$

The parameter $\alpha \in [0,1]$ in Equation (2.20) is called upwind parameter and can be used to control the dissipation. Taking $\alpha = 0$ yields a non dissipative central flux while $\alpha = 1$ corresponds to the classic upwind flux.

The treatment of anisotropic materials within a DG framework was discussed in [16, 32] where just the central flux is considered to interconnect the neighbouring elements. In [49] the flux formulation for isotropic materials is extended for two dimensional formulation of anisotropic materials. The numerical flux (2.20) involves material properties ε and μ via impedance Z and conductance Y. The numerical flux in isotropic materials is extended to an anisotropic materials with redefining the material properties, Z and Y with effective permittivity (1.25). The DG flux for the 2D model in

anisotropic material is

$$n\cdot(F(\tilde{q})-F^{*}(\tilde{q}))=\begin{pmatrix}\frac{-n_y}{Z^{+}+Z^{-}}\left(Z^{+}[\tilde{H}_z]-\alpha\left(n_x[\tilde{E}_y]-n_y[\tilde{E}_x]\right)\right)\\ \frac{n_x}{Z^{+}+Z^{-}}\left(Z^{+}[\tilde{H}_z]-\alpha\left(n_x[\tilde{E}_y]-n_y[\tilde{E}_x]\right)\right)\\ \frac{1}{Y^{+}+Y^{-}}\left(Y^{+}\left(n_x[\tilde{E}_y]-n_y[\tilde{E}_x]\right)-\alpha[\tilde{H}_z]\right)\end{pmatrix}, \tag{2.22}$$

where $Z^{\pm}=\mu^{\pm}c^{\pm}$ and $Y^{\pm}=(Z^{\pm})^{-1}$, while

$$c^{\pm}=\sqrt{\frac{n^T\varepsilon^{\pm}n}{\mu^{\pm}\det(\varepsilon^{\pm})}}.$$

2.2.4 Boundary conditions

Since the information between elements is exchanged by the numerical flux in DG methods, the boundary conditions are enforced by the numerical flux. The field differences (2.21) mediate in between neighbouring elements in numerical flux. Boundary elements are missing a neighbour to evaluate the field differences. To overcome this problem and according to different boundary conditions we consider that the jumps in Equation (2.22) are modified as following for different boundary conditions.

The PEC boundary condition can be implemented by applying the mirror principle as $n\times\tilde{E}^{+}=-n\times\tilde{E}^{-}$ and $n\times\tilde{H}^{+}=n\times\tilde{H}^{-}$. Thus the jumps at the outer boundary in this case are set as

$$[\tilde{E}_x]=2\tilde{E}_x^{-},\quad [\tilde{E}_y]=2\tilde{E}_y^{-},\quad [\tilde{H}_z]=0. \tag{2.23}$$

The jumps at the outer boundary for PMC are set as

$$[\tilde{E}_x]=0,\quad [\tilde{E}_y]=0,\quad [\tilde{H}_z]=2\tilde{H}_z^{-}. \tag{2.24}$$

In DGTD methods Silver-Müller absorbing boundary conditions can be applied by setting the incoming flux to zero [3]. For the upwind flux, this is directly implemented since it is equivalent to assuming that there is no contribution to the flux from outside the region of solution, so we have

$$Z^{-}\tilde{H}_z^{+}=n_x\tilde{E}_y^{+}-n_y\tilde{E}_x^{+},\quad \text{or equivalently}\quad \tilde{H}_z^{+}=Y^{-}(n_x\tilde{E}_y^{+}-n_y\tilde{E}_x^{+}). \tag{2.25}$$

For the central flux SM-ABC can also be employed [32]. Using the same kind of approach as in [4], for central flux we have

$$Z^{-}\tilde{H}_z^{+}=(n_x\tilde{E}_y^{-}-n_y\tilde{E}_x^{-}),\quad \text{and}\quad Y^{-}(n_x\tilde{E}_y^{+}-n_y\tilde{E}_x^{+})=\tilde{H}_z^{-}. \tag{2.26}$$

In order to incorporate the conditions (2.25) and (2.26) into the numerical flux (2.22), for both central and upwind fluxes we consider $\alpha=1$ for numerical flux at the outer boundary and

$$[\tilde{E}_x]=\tilde{E}_x^{-},\quad [\tilde{E}_y]=\tilde{E}_y^{-},\quad [\tilde{H}_z]=\tilde{H}_z^{-}. \tag{2.27}$$

We also consider (2.27) for the case $\alpha\in(0,1)$.

To complete the evaluation of fluxes at boundary edges, we mention that the material properties at the boundary set as

$$Z^+ = Z^- \quad \text{and} \quad Y^+ = Y^-.$$

2.3 Time discretization

The discretization by the DG method results in the semi-discrete form of Maxwell's equations which discretize the electromagnetic fields in space. The missing part towards a time-domain solver is the integration in time. The Equation (2.19) is a set of coupled first order ordinary differential equations in time, which is in more general form described by

$$\frac{d\tilde{q}_{k,h}}{dt} = \mathscr{F}_h(\tilde{q}_{k,h)}, \tag{2.28}$$

where $\tilde{q}_{k,h}$ denotes the vector of the coefficients of $\tilde{q}_k$ and $\mathscr{F}_h : V_N \to \mathbb{R}^{3N_p}$.

In order to achieve the fully discrete scheme, it is enough to employ a time integration method. The temporal integration methods can be divided to two major families: implicit and explicit schemes. Implicit schemes require the solution of large matrix system resulting in a high computational effort per time step and rely on the efficiency of the used linear system solver. The advantage of implicit schemes are their flexibility regarding the choice of time step since usually, these time integrations are unconditionally stable. Explicit schemes in contrast are easy to implement, produce greater accuracy with less computational effort than implicit methods, but are restricted by a stability criterion enforcing a relation between the time step and the spatial discretization parameter. This restriction may result in a large number of iterations per simulation, each iteration with low computational effort.

Explicit time integrators can exploit the block diagonal structure of the mass matrix of semi-discrete DG schemes and thus leads to fully explicit scheme. In this study, we employ the 2nd-order leap-frog scheme for the time integration which is described in the FDTD literature in [89]. The staggered leapfrog time-stepping algorithm is a popular choice for time domain Maxwell's equations (*e.g.* [4, 32, 82]) due to its simplicity, as it does not require to save in memory previous states, accuracy and robustness. It samples the unknown fields in a staggered way: the electric field at $t^m = m\Delta t$, and the magnetic field at $t^{m+1/2} = (m+\frac{1}{2})\Delta t$. The staggered sampling yields an explicit marching-on-in-time algorithm.

2.3.1 Leap-frog time integrator

In this work, we use a second order leap-frog scheme, which employs a centered approximation for the time derivatives. Consider the semi-discrete form Equation (2.19). To define a fully discrete scheme, we divide the time interval $[0,T]$ into M subintervals by points

$$0 = t^0 < t^1 < \cdots < t^M = T,$$

where $t^m = m\Delta t$ while Δt is the time step size and $T + \Delta t/2 \leq T_f$.

The unknowns related to the electric field are approximated at integer time-stations t^m and are denoted by $\tilde{E}_k^m = \tilde{E}_k(.,t^m)$. The unknowns related to the magnetic field are approximated at half-

integer time-stations $t^{m+1/2} = (m+\frac{1}{2})\Delta t$ and are denoted by $\tilde{H}_k^{m+1/2} = \tilde{H}_k(.,t^{m+1/2})$. This implies that we do not have a fully defined state vector in the sense of Equation (2.28) for a given time t. The time derivatives in Equation (2.19) are replaced by 2nd-order accurate central differences

$$\frac{\partial \tilde{E}^{m+1/2}}{\partial t} = \frac{\tilde{E}^{m+1} - \tilde{E}^{m}}{\Delta t} + \mathscr{O}(\Delta t^2) \quad \text{and} \quad \frac{\partial \tilde{H}^{m+1}}{\partial t} = \frac{\tilde{H}^{m+3/2} - \tilde{H}^{m+1/2}}{\Delta t} + \mathscr{O}(\Delta t^2). \tag{2.29}$$

To obtain the future values from a present state the following algorithm is applied

$$\tilde{E}_{k,h}^{m+1} = \tilde{E}_{k,h}^{m} + \Delta t \mathscr{F}_{h_1}(\tilde{E}_{k,h}^{m}, \tilde{H}_{k,h}^{m+1/2}), \tag{2.30}$$

$$\tilde{H}_{k,h}^{m+3/2} = \tilde{H}_{k,h}^{m+1/2} + \Delta t \mathscr{F}_{h_2}(\tilde{E}_{k,h}^{m+1}, \tilde{H}_{k,h}^{m+1/2}), \tag{2.31}$$

where $\tilde{E}_{k,h}^{m}$ denotes the vector of the coefficients of $\tilde{E}_k^m$ which is an approximation for $E_k(t^m)$ and $\tilde{H}_{k,h}^{m+1/2}$ is defined analogously. The function $\mathscr{F}_h = (\mathscr{F}_{h_1}, \mathscr{F}_{h_2})^T$ is a function representing the result of applying the spatial semi-discretization.

Usually evaluating the upwind flux terms imply the need of averaging between the next and previous time steps and thus resulting on a globally implicit scheme due to the coupling terms from the adjacent elements [56]. To avoid this, we used in (2.30)–(2.31) the backward approximation [25] for the two extra dissipative terms arising from the upwind flux formulation. As discussed in [2, 66] this backward approximation for the flux terms is enough to attenuate spurious modes in space more strongly than physical modes, which is the only aim of these terms. Using this backward approximation introduces a slight penalization in stability condition for upwind flux. This will be observed later in stability analysis and in our numerical experiments.

2.4 Full-discrete scheme

The leap-frog time integrator is applied on semi discrete DG formulation where the first order time derivates are replaced by central differences and backward approximations for the dissipative terms of flux. We arrive at a fully explicit leapfrog DG scheme in 2D. In Subsection 2.4.2, we consider the system of 3D Maxwell's equations and extend the leap-frog DG formulation to 3D.

2.4.1 Leap-frog DG scheme in 2D

In order to get the fully explicit leap-frog DG scheme, the leap-frog time integrator (2.29) is applied on Equation (2.19). We can now formulate the leap-frog DG method: given an initial approximation $(\tilde{E}_{x_k}^0, \tilde{E}_{y_k}^0, \tilde{H}_{z_k}^{1/2}) \in V_N$, for each $m = 0, 1, \ldots, M-1$, find $(\tilde{E}_{x_k}^{m+1}, \tilde{E}_{y_k}^{m+1}, \tilde{H}_{z_k}^{m+1/2}) \in V_N$ such that

$\forall (u_k, v_k, w_k) \in V_N$,

$$\left(\varepsilon_{xx}\frac{\tilde{E}_{x_k}^{m+1}-\tilde{E}_{x_k}^{m}}{\Delta t}+\varepsilon_{xy}\frac{\tilde{E}_{y_k}^{m+1}-\tilde{E}_{y_k}^{m}}{\Delta t},u_k\right)_{T_k}=\left(\partial_y\tilde{H}_{z_k}^{m+1/2},u_k\right)_{T_k}$$
$$+\left(\frac{-n_y}{Z^+ + Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),u_k\right)_{\partial T_k}, \tag{2.32}$$

$$\left(\varepsilon_{yx}\frac{\tilde{E}_{x_k}^{m+1}-\tilde{E}_{x_k}^{m}}{\Delta t}+\varepsilon_{yy}\frac{\tilde{E}_{y_k}^{m+1}-\tilde{E}_{y_k}^{m}}{\Delta t},v_k\right)_{T_k}=-\left(\partial_x\tilde{H}_{z_k}^{m+1/2},v_k\right)_{T_k}$$
$$+\left(\frac{n_x}{Z^+ + Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),v_k\right)_{\partial T_k}, \tag{2.33}$$

$$\left(\mu\frac{\tilde{H}_{z_k}^{m+3/2}-\tilde{H}_{z_k}^{m+1/2}}{\Delta t},w_k\right)_{T_k}=\left(\partial_y\tilde{E}_{x_k}^{m+1}-\partial_x\tilde{E}_{y_k}^{m+1},w_k\right)_{T_k}$$
$$+\left(\frac{1}{Y^+ + Y^-}\left(Y^+(n_x[\tilde{E}_y^{m+1}]-n_y[\tilde{E}_x^{m+1}])-\alpha[\tilde{H}_z^{m+1/2}]\right),w_k\right)_{\partial T_k}, \tag{2.34}$$

where $(\cdot,\cdot)_{T_k}$ and $(\cdot,\cdot)_{\partial T_k}$ denote the classical $L^2(T_k)$ and $L^2(\partial T_k)$ inner-products we defined before on each element. The boundary conditions are considered as what described in Subsection 2.2.4.

We want to emphasize that the scheme (2.32)–(2.34) is fully explicit in time, in opposition to [56], where the scheme is defined with the upwind fluxes involving the unknowns E_k^{m+1} and $H_k^{m+3/2}$ and to [32], where the scheme that is defined with the central fluxes leads to a locally implicit time method in the case of Silver-Müller absorbing boundary conditions.

2.4.2 Leap-frog DG scheme in 3D

Recall the 3D Maxwell's equations (2.1)–(2.2). We assume that polyhedral domain Ω is partitioned into K disjoint tetrahedral elements T_k. We consider that the resulting mesh $\mathscr{T}_h$ is conforming and regular in the same sense as (2.14), where h_k and τ_k denotes the diameter of the element T_k and maximum diameter of a sphere inscribed in T_k respectively.

Keeping the same notation from 2D, ν_k denotes the set of indices of the neighbouring elements of T_k. For each $i \in \nu_k$, we consider the internal face $f_{ik} = T_i \cap T_k$, and we denote by n_{ik} the unit normal oriented from T_i towards T_k. The boundary face is $f_k = T_k \cap \partial\Omega$. The unitary outer normal vector to f_k is denoted by n_k. The set of internal faces denote by F^{int} and F^{ext} is the set of faces that belong to the boundary $\partial\Omega$.

Writing the equations as a conservation form (2.5) the semi-discrete DG formulation is an extension of the same approach as for 2D. The tensorial permittivity and permeability can be reduced to effective scalar values (1.25) and (1.26) respectively. The numerical flux is introduced as

$$n\cdot(F(\tilde{q})-F^*(\tilde{q}))=\begin{pmatrix}\frac{-1}{Z^+ + Z^-}n\times\left(Z^+[\tilde{H}]-\alpha n\times[\tilde{E}]\right)\\ \frac{1}{Y^+ + Y^-}n\times\left(Y^+[\tilde{E}]+\alpha n\times[\tilde{H}]\right)\end{pmatrix},$$

where $Z^{\pm} = \mu_{eff}^{\pm} c^{\pm}$ and $Y^{\pm} = (Z^{\pm})^{-1}$ while

$$\mu_{eff}^{\pm} = \frac{\det(\mu)}{n^T \mu^{\pm} n}, \quad \text{and} \quad c^{\pm} = \sqrt{\frac{n^T \mu^{\pm} \varepsilon^{\pm} n}{\det(\mu)\det(\varepsilon)}}.$$

With this flux formulation, the leap-frog DG scheme (2.32)–(2.34) is extended to 3D. Given an initial approximation $(\tilde{E}_k^0, \tilde{H}_k^{1/2}) \in V_N$, for each $m = 0, 1, \ldots, M-1$, find $(\tilde{E}_k^{m+1}, \tilde{H}_k^{m+1/2}) \in V_N$ such that $\forall (u_k, v_k) \in V_N$,

$$\begin{aligned}
\left(\varepsilon \frac{\tilde{E}_k^{m+1} - \tilde{E}_k^m}{2}, u_k \right)_{T_k} &= \left(\nabla \times \tilde{H}_k^{m+1/2}, u_k \right)_{T_k} \\
&+ \left(\frac{1}{Z^+ + Z^-} n \times \left(Z^+ [\tilde{H}_k^{m+1/2}] - \alpha n \times [\tilde{E}_k^m] \right), u_k \right)_{\partial T_k}, \qquad (2.35) \\
\left(\mu \frac{\tilde{H}_k^{m+3/2} - \tilde{H}_k^{m+1/2}}{2}, v_k \right)_{T_k} &= - \left(\nabla \times \tilde{E}_k^{m+1}, v_k \right)_{T_k} \\
&+ \left(\frac{1}{Y^+ + Y^-} n \times \left(Y^+ [\tilde{E}_k^{m+1}] + \alpha n \times [\tilde{H}_k^{m+1/2}] \right), v_k \right)_{\partial T_k}, \qquad (2.36)
\end{aligned}$$

where V_N is the space of finite element space defined by (2.15) on tetrahedral elements. The scheme (2.35)–(2.36) is a fully explicit leap-frog DG scheme for 3D Maxwell's equations with anisotropic permittivity and the permeability tensors. The stability of this scheme will be analyzed in next chapter.

Chapter 3

Stability Analysis

The explicit leap-frog time integration scheme (2.30)–(2.31), which employed a centered approximation for time derivatives and used a backward approximation in upwind flux terms yielded a fully explicit leap-frog DG scheme which is conditionally stable. The conditional stability imposes a condition on the time step Δt, in the same fashion as the Courant-Friedrichs-Lewy (CFL) criterion [22], defined with respect to the time integration scheme in use.

In this chapter we first present a rigorous proof of the conditional stability of leap-frog DG scheme in 2D. The analysis shows the influence of the mesh size, the choice of the numerical flux and the choice of the degree of the polynomials used in the construction of the finite element space and the boundary conditions, which can be either perfect electric, perfect magnetic or first order Silver-Müller. The stability analysis is extended to the 3D model in Section 3.2 and the correspondent stability region is presented.

3.1 Stability analysis of the 2D model

The aim of this section is to provide a sufficient condition for the L^2-stability of the leap-frog DG method (2.32)–(2.34). We start by choosing a proper test function and try to find the estimations for the terms related to inner boundary edges and outer boundary terms in Lemma 3.1.1 and Lemma 3.1.2 respectively. The main result of this section is stated in Theorem 3.1.3 where the stability condition is derived for different boundary conditions and different fluxes.

Recall Equations (2.32)–(2.34), choosing

$$u_k = \Delta t \tilde{E}_{x_k}^{[m+1/2]}, \quad v_k = \Delta t \tilde{E}_{y_k}^{[m+1/2]}, \quad w_k = \Delta t \tilde{H}_{z_k}^{[m+1]},$$

where

$$\tilde{E}^{[m+1/2]} = \frac{\tilde{E}^m + \tilde{E}^{m+1}}{2} \quad \text{and} \quad \tilde{H}^{[m+1]} = \frac{\tilde{H}^{m+1/2} + \tilde{H}^{m+3/2}}{2},$$

we have

$$\begin{aligned}\left(\varepsilon\tilde{E}_k^{m+1},\tilde{E}_k^{m+1}\right)_{T_k}-\left(\varepsilon\tilde{E}_k^{m},\tilde{E}_k^{m}\right)_{T_k}&=2\Delta t\left(\nabla\times\tilde{H}_{z_k}^{m+1/2},\tilde{E}_k^{[m+1/2]}\right)_{T_k}\\&+2\Delta t\left(\frac{-n_y}{Z^++Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),\tilde{E}_{x_k}^{[m+1/2]}\right)_{\partial T_k}\\&+2\Delta t\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),\tilde{E}_{y_k}^{[m+1/2]}\right)_{\partial T_k},\end{aligned}\tag{3.1}$$

$$\begin{aligned}\left(\mu\tilde{H}_{z_k}^{m+3/2},\tilde{H}_{z_k}^{m+3/2}\right)_{T_k}-\left(\mu\tilde{H}_{z_k}^{m+1/2},\tilde{H}_{z_k}^{m+1/2}\right)_{T_k}&=-2\Delta t\left(\operatorname{curl}\tilde{E}_k^{m+1},\tilde{H}_{z_k}^{[m+1]}\right)_{T_k}\\&+2\Delta t\left(\frac{1}{Y^++Y^-}\left(Y^+\left(n_x[\tilde{E}_y^{m+1}]-n_y[\tilde{E}_x^{m+1}]\right)-\alpha[\tilde{H}_z^{m+1/2}]\right),\tilde{H}_{z_k}^{[m+1]}\right)_{\partial T_k}.\end{aligned}\tag{3.2}$$

Using the identity,

$$\left(\operatorname{curl}\tilde{E}_k^{m+1},\tilde{H}_{z_k}^{[m+1]}\right)_{T_k}=\left(\nabla\times\tilde{H}_{z_k}^{[m+1]},\tilde{E}_k^{m+1}\right)_{T_k}+\left(n_x\tilde{E}_{y_k}^{m+1}-n_y\tilde{E}_{x_k}^{m+1},\tilde{H}_{z_k}^{[m+1]}\right)_{\partial T_k},$$

summing (3.1) and (3.2) from $m=0$ to $m=M-1$, and integrating by parts, we get

$$\begin{aligned}\left(\varepsilon\tilde{E}_k^{M},\tilde{E}_k^{M}\right)_{T_k}+\left(\mu\tilde{H}_{z_k}^{M+1/2},\tilde{H}_{z_k}^{M+1/2}\right)_{T_k}&=\left(\varepsilon\tilde{E}_k^{0},\tilde{E}_k^{0}\right)_{T_k}+\left(\mu\tilde{H}_{z_k}^{1/2},\tilde{H}_{z_k}^{1/2}\right)_{T_k}\\&+\Delta t\left(\nabla\times\tilde{H}_{z_k}^{1/2},\tilde{E}_k^{0}\right)_{T_k}-\Delta t\left(\nabla\times\tilde{H}_{z_k}^{M+1/2},\tilde{E}_k^{M}\right)_{T_k}+2\Delta t\sum_{m=0}^{M-1}A_k^m,\end{aligned}\tag{3.3}$$

where

$$\begin{aligned}A_k^m&=\left(\frac{-n_y}{Z^++Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),\tilde{E}_{x_k}^{[m+1/2]}\right)_{\partial T_k}\\&+\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}]-\alpha\left(n_x[\tilde{E}_y^m]-n_y[\tilde{E}_x^m]\right)\right),\tilde{E}_{y_k}^{[m+1/2]}\right)_{\partial T_k}\\&+\left(\frac{1}{Y^++Y^-}\left(Y^+\left(n_x[\tilde{E}_y^{m+1}]-n_y[\tilde{E}_x^{m+1}]\right)-\alpha[\tilde{H}_z^{m+1/2}]\right),\tilde{H}_{z_k}^{[m+1]}\right)_{\partial T_k}\\&-\left(n_x\tilde{E}_{y_k}^{m+1}-n_y\tilde{E}_{x_k}^{m+1},\tilde{H}_{z_k}^{[m+1]}\right)_{\partial T_k}.\end{aligned}$$

Summing over all elements $T_k\in\mathscr{T}_h$ we obtain

$$\sum_{T_k\in\mathscr{T}_h}A_k^m=B_1^m+B_2^m,\tag{3.4}$$

where B_1^m contains the terms on internal edges and B_2^m contains the boundary terms. We can write B_1^m as

$$B_1^m=B_{11}^m+B_{12}^m+B_{13}^m,$$

where

$$
\begin{aligned}
B_{11}^m \;=\; & \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(\frac{-(n_y)_{ki}}{Z_i+Z_k} \Big(Z_i[\tilde{H}_{z_k}^{m+1/2}] - \alpha \big((n_x)_{ki}[\tilde{E}_{y_k}^m] - (n_y)_{ki}[\tilde{E}_{x_k}^m] \big) \Big) \tilde{E}_{x_k}^{[m+1/2]} \\
& + \frac{-(n_y)_{ik}}{Z_i+Z_k} \Big(Z_k[\tilde{H}_{zi}^{m+1/2}] - \alpha \big((n_x)_{ik}[\tilde{E}_{yi}^m] - (n_y)_{ik}[\tilde{E}_{xi}^m] \big) \Big) \tilde{E}_{xi}^{[m+1/2]} \\
& - \frac{Y_i (n_y)_{ki}}{Y_i+Y_k} [\tilde{E}_{x_k}^{m+1}] \tilde{H}_{z_k}^{[m+1]} - \frac{Y_k (n_y)_{ik}}{Y_i+Y_k} [\tilde{E}_{xi}^{m+1}] \tilde{H}_{zi}^{[m+1]} \\
& + (n_y)_{ki} \tilde{E}_{x_k}^{m+1} \tilde{H}_{z_k}^{[m+1]} + (n_y)_{ik} \tilde{E}_{xi}^{m+1} \tilde{H}_{zi}^{[m+1]} \Bigg) \, ds,
\end{aligned} \tag{3.5}
$$

$$
\begin{aligned}
B_{12}^m \;=\; & \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(\frac{(n_x)_{ki}}{Z_i+Z_k} \Big(Z_i[\tilde{H}_{z_k}^{m+1/2}] - \alpha \big((n_x)_{ki}[\tilde{E}_{y_k}^m] - (n_y)_{ki}[\tilde{E}_{x_k}^m] \big) \Big) \tilde{E}_{y_k}^{[m+1/2]} \\
& + \frac{(n_x)_{ik}}{Z_i+Z_k} \Big(Z_k[\tilde{H}_{zi}^{m+1/2}] - \alpha \big((n_x)_{ik}[\tilde{E}_{yi}^m] - (n_y)_{ik}[\tilde{E}_{xi}^m] \big) \Big) \tilde{E}_{yi}^{[m+1/2]} \\
& + \frac{Y_i (n_x)_{ki}}{Y_i+Y_k} [\tilde{E}_{y_k}^{m+1}] \tilde{H}_{z_k}^{[m+1]} + \frac{Y_k (n_x)_{ik}}{Y_i+Y_k} [\tilde{E}_{yi}^{m+1}] \tilde{H}_{zi}^{[m+1]} \\
& - (n_x)_{ki} \tilde{E}_{y_k}^{m+1} \tilde{H}_{z_k}^{[m+1]} - (n_x)_{ik} \tilde{E}_{yi}^{m+1} \tilde{H}_{zi}^{[m+1]} \Bigg) \, ds,
\end{aligned} \tag{3.6}
$$

$$
B_{13}^m = - \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \left(\frac{\alpha}{Y_i+Y_k} [\tilde{H}_{z_k}^{m+1/2}] \tilde{H}_{z_k}^{[m+1]} + \frac{\alpha}{Y_i+Y_k} [\tilde{H}_{zi}^{m+1/2}] \tilde{H}_{zi}^{[m+1]} \right) ds. \tag{3.7}
$$

The terms related with the outer boundary are

$$
\begin{aligned}
B_2^m \;=\; & \sum_{f_k\in F^{ext}} \int_{f_k} \Bigg(\frac{-(n_y)_k}{2Z_k} \Big(Z_k[\tilde{H}_{z_k}^{m+1/2}] - \alpha \big((n_x)_k[\tilde{E}_{y_k}^m] - (n_y)_k[\tilde{E}_{x_k}^m] \big) \Big) \tilde{E}_{x_k}^{[m+1/2]} \\
& + \frac{(n_x)_k}{2Z_k} \Big(Z_k[\tilde{H}_{z_k}^{m+1/2}] - \alpha \big((n_x)_k[\tilde{E}_{y_k}^m] - (n_y)_k[\tilde{E}_{x_k}^m] \big) \Big) \tilde{E}_{y_k}^{[m+1/2]} \\
& + \frac{1}{2Y_k} \Big(Y_k \big((n_x)_k[\tilde{E}_{y_k}^{m+1}] - (n_y)_k[\tilde{E}_{x_k}^{m+1}] \big) - \alpha[\tilde{H}_{z_k}^{m+1/2}] \Big) \tilde{H}_{z_k}^{[m+1]} \\
& - \big((n_x)_k \tilde{E}_{y_k}^{m+1} - (n_y)_k \tilde{E}_{x_k}^{m+1} \big) \tilde{H}_{z_k}^{[m+1]} \Bigg) \, ds.
\end{aligned} \tag{3.8}
$$

Now we find estimates for the terms of internal edges in Lemma 3.1.1 by finding an estimation for B_1^m and then an estimate for the terms on edges that belong to the boundary in Lemma 3.1.2 by finding an estimate for B_2^m.

Lemma 3.1.1. *Let B_{11}^m, B_{12}^m and B_{13}^m be defined by (3.5), (3.6) and (3.7), respectively and $B_1^m = B_{11}^m + B_{12}^m + B_{13}^m$. Then*

$$\begin{aligned}\sum_{m=0}^{M-1} B_1^m \leq \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{1}{4(Z_i+Z_k)} \Bigg(& -\alpha \left((n_y)_{ki}[\tilde{E}_{x_k}^0] - (n_x)_{ki}[\tilde{E}_{y_k}^0]\right)^2 \\ & +2\left((n_x)_{ki}\left(Z_i\tilde{E}_{y_k}^0 + Z_k\tilde{E}_{yi}^0\right) - (n_y)_{ki}\left(Z_i\tilde{E}_{x_k}^0 + Z_k\tilde{E}_{xi}^0\right)\right)[\tilde{H}_{z_k}^{1/2}] \\ & +\alpha\left(\left((n_y)_{ki}[\tilde{E}_{x_k}^M] - (n_x)_{ki}[\tilde{E}_{y_k}^M]\right)^2 - \left([\tilde{H}_{z_k}^{1/2}]^2 - [\tilde{H}_{z_k}^{M+1/2}]^2\right)\right) \\ & +2\left((n_y)_{ki}\left(Z_i\tilde{E}_{x_k}^M + Z_k\tilde{E}_{xi}^M\right) - (n_x)_{ki}\left(Z_i\tilde{E}_{y_k}^M + Z_k\tilde{E}_{yi}^M\right)\right)[\tilde{H}_{z_k}^{M+1/2}]\Bigg) ds.\end{aligned}$$

Proof. Since

$$\frac{Z_i}{Z_i+Z_k} + \frac{Y_i}{Y_i+Y_k} = \frac{Z_k}{Z_i+Z_k} + \frac{Y_k}{Y_i+Y_k} = 1 \tag{3.9}$$

and

$$\frac{Z_i}{Z_i+Z_k} = \frac{Y_k}{Y_i+Y_k}, \qquad \frac{Z_k}{Z_i+Z_k} = \frac{Y_i}{Y_i+Y_k}, \tag{3.10}$$

for B_{11}^m we have

$$\begin{aligned}B_{11}^m = \frac{1}{2}\sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& \frac{-(n_y)_{ki}}{Z_i+Z_k}\left(Z_i[\tilde{H}_{z_k}^{m+1/2}] - \alpha\left((n_x)_{ki}[\tilde{E}_{y_k}^m] - (n_y)_{ki}[\tilde{E}_{x_k}^m]\right)\right)\tilde{E}_{x_k}^m \\ & +\frac{-(n_y)_{ki}}{Z_i+Z_k}\left(-\alpha\left((n_x)_{ki}[\tilde{E}_{y_k}^m] - (n_y)_{ki}[\tilde{E}_{x_k}^m]\right)\right)\tilde{E}_{x_k}^{m+1} \\ & +\frac{-(n_y)_{ik}}{Z_i+Z_k}\left(Z_k[\tilde{H}_{zi}^{m+1/2}] - \alpha\left((n_x)_{ik}[\tilde{E}_{yi}^m] - (n_y)_{ik}[\tilde{E}_{xi}^m]\right)\right)\tilde{E}_{xi}^m \\ & +\frac{-(n_y)_{ik}}{Z_i+Z_k}\left(-\alpha\left((n_x)_{ik}[\tilde{E}_{yi}^m] - (n_y)_{ik}[\tilde{E}_{xi}^m]\right)\right)\tilde{E}_{xi}^{m+1} \\ & -\frac{Y_i(n_y)_{ki}}{Y_i+Y_k}[\tilde{E}_{x_k}^{m+1}]\tilde{H}_{z_k}^{m+3/2} - \frac{Y_k(n_y)_{ik}}{Y_i+Y_k}[\tilde{E}_{xi}^{m+1}]\tilde{H}_{zi}^{m+3/2} \\ & +(n_y)_{ki}\tilde{E}_{x_k}^{m+1}\tilde{H}_{z_k}^{m+3/2} + (n_y)_{ik}\tilde{E}_{xi}^{m+1}\tilde{H}_{zi}^{m+3/2}\Bigg) ds.\end{aligned}$$

Summing from $m=0$ to $m=M-1$ we conclude that

$$\begin{aligned}\sum_{m=0}^{M-1} B_{11}^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{(n_y)_{ki}}{2(Z_i+Z_k)} \Bigg(& -\left(Z_i\tilde{E}_{x_k}^0 + Z_k\tilde{E}_{xi}^0\right)[\tilde{H}_{z_k}^{1/2}] + \alpha\left((n_x)_{ki}[\tilde{E}_{y_k}^0] - (n_y)_{ki}[\tilde{E}_{x_k}^0]\right)[\tilde{E}_{x_k}^0] \\ & +\alpha\sum_{m=0}^{M-1}\left((n_x)_{ki}[\tilde{E}_{y_k}^{m+1}] - (n_y)_{ki}[\tilde{E}_{x_k}^{m+1}] + (n_x)_{ki}[\tilde{E}_{y_k}^m] - (n_y)_{ki}[\tilde{E}_{x_k}^m]\right)[\tilde{E}_{x_k}^{m+1}] \\ & +\left(Z_i\tilde{E}_{x_k}^M + Z_k\tilde{E}_{xi}^M\right)[\tilde{H}_{z_k}^{M+1/2}] - \alpha\left((n_x)_{ki}[\tilde{E}_{y_k}^M] - (n_y)_{ki}[\tilde{E}_{x_k}^M]\right)[\tilde{E}_{x_k}^M]\Bigg) ds.\end{aligned}$$

In the same way, for B_{12}^m we have

$$\begin{aligned}\sum_{m=0}^{M-1} B_{12}^m &= \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{(n_x)_{ki}}{2(Z_i+Z_k)} \Bigg(\left(Z_i\tilde{E}_{y_k}^0 + Z_k\tilde{E}_{yi}^0\right)[\tilde{H}_{z_k}^{1/2}] - \left((n_x)_{ki}[\tilde{E}_{y_k}^0] - (n_y)_{ki}[\tilde{E}_{x_k}^0]\right)[\tilde{E}_{y_k}^0] \\ &\quad -\alpha \sum_{m=0}^{M-1} \left((n_x)_{ki}[\tilde{E}_{y_k}^{m+1}] - (n_y)_{ki}[\tilde{E}_{x_k}^{m+1}] + (n_x)_{ki}[\tilde{E}_{y_k}^{m}] - (n_y)_{ki}[\tilde{E}_{x_k}^{m}]\right)[\tilde{E}_{y_k}^{m+1}] \\ &\quad - \left(Z_i\tilde{E}_{y_k}^M + Z_k\tilde{E}_{yi}^M\right)[\tilde{H}_{z_k}^{M+1/2}] + \alpha\left((n_x)_{ki}[\tilde{E}_{y_k}^M] - (n_y)_{ki}[\tilde{E}_{x_k}^M]\right)[\tilde{E}_{y_k}^M] \Bigg) ds,\end{aligned}$$

and for B_{13}^m

$$\sum_{m=0}^{M-1} B_{13}^m = -\sum_{m=0}^{M-1} \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{\alpha}{2(Y_i+Y_k)} [\tilde{H}_{z_k}^{m+1/2}] \left([\tilde{H}_{z_k}^{m+1/2}] + [\tilde{H}_{z_k}^{m+3/2}] \right) ds.$$

Observing that, for general sequences $\{a^m\}$ and $\{b^m\}$ hold

$$\sum_{m=0}^{M-1} \left(a^{m+1}+a^m\right) a^{m+1} = \frac{1}{2}\left(-(a^0)^2 + (a^M)^2 + \sum_{m=0}^{M-1} \left(a^m + a^{m+1}\right)^2 \right),$$

$$\sum_{m=0}^{M-1} \left(a^{m+1}+a^m\right) b^{m+1} = \frac{1}{2}\left(-a^0b^0 + a^Mb^M + \sum_{m=0}^{M-1} \left(a^mb^m + 2a^mb^{m+1} + a^{m+1}b^{m+1}\right) \right),$$

we get

$$\begin{aligned}\sum_{m=0}^{M-1} (B_{11}^m + B_{12}^m) &\le \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{1}{4(Z_i+Z_k)} \Bigg(-\alpha(n_y)_{ki}^2 \left(-[\tilde{E}_{x_k}^0]^2 + [\tilde{E}_{x_k}^M]^2\right) \\ &\quad +\alpha(n_x)_{ki}(n_y)_{ki}\left(-[\tilde{E}_{x_k}^0][\tilde{E}_{y_k}^0] + [\tilde{E}_{x_k}^M][\tilde{E}_{y_k}^M]\right) - 2(n_y)_{ki}\left(Z_i\tilde{E}_{x_k}^0 + Z_k\tilde{E}_{xi}^0\right)[\tilde{H}_{z_k}^{1/2}] \\ &\quad +2\alpha(n_y)_{ki}\left((n_x)_{ki}[\tilde{E}_{y_k}^0] - (n_y)_{ki}[\tilde{E}_{x_k}^0]\right)[\tilde{E}_{x_k}^0] + 2(n_y)_{ki}\left(Z_i\tilde{E}_{x_k}^M + Z_k\tilde{E}_{xi}^M\right)[\tilde{H}_{z_k}^{M+1/2}] \\ &\quad -2\alpha(n_y)_{ki}\left((n_x)_{ki}[\tilde{E}_{y_k}^M] - (n_y)_{ki}[\tilde{E}_{x_k}^M]\right)[\tilde{E}_{x_k}^M] - \alpha(n_x)_{ki}^2\left(-[\tilde{E}_{y_k}^0]^2 + [\tilde{E}_{y_k}^M]^2\right) \\ &\quad +\alpha(n_x)_{ki}(n_y)_{ki}\left(-[\tilde{E}_{x_k}^0][\tilde{E}_{y_k}^0] + [\tilde{E}_{x_k}^M][\tilde{E}_{y_k}^M]\right) + 2(n_x)_{ki}\left(Z_i\tilde{E}_{y_k}^0 + Z_k\tilde{E}_{yi}^0\right)[\tilde{H}_{z_k}^{1/2}] \\ &\quad -2\alpha(n_x)_{ki}\left((n_x)_{ki}[\tilde{E}_{y_k}^0] - (n_y)_{ki}[\tilde{E}_{x_k}^0]\right)[\tilde{E}_{y_k}^0] - 2(n_x)_{ki}\left(Z_i\tilde{E}_{y_k}^M + Z_k\tilde{E}_{yi}^M\right)[\tilde{H}_{z_k}^{M+1/2}] \\ &\quad +2\alpha(n_x)_{ki}\left((n_x)_{ki}[\tilde{E}_{y_k}^M] - (n_y)_{ki}[\tilde{E}_{x_k}^M]\right)[\tilde{E}_{y_k}^M] \Bigg) ds.\end{aligned}$$

We also have

$$\begin{aligned}\sum_{m=0}^{M-1} B_{13}^m &= -\sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{\alpha}{4(Y_i+Y_k)} \left([\tilde{H}_{z_k}^{1/2}]^2 - [\tilde{H}_{z_k}^{M+1/2}]^2 + \sum_{m=0}^{M-1} \left([\tilde{H}_{z_k}^{m+1/2}] + [\tilde{H}_{z_k}^{m+3/2}] \right)^2 \right) ds \\ &\le -\sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{\alpha}{4(Y_i+Y_k)} \left([\tilde{H}_{z_k}^{1/2}]^2 - [\tilde{H}_{z_k}^{M+1/2}]^2 \right) ds,\end{aligned}$$

which concludes the proof. □

Let us now analyze the term B_2^m for different kinds of boundary conditions. The parameter β_1, β_2 and β_3 which have appeared in the following estimation depend on the type of fluxes and boundary conditions that we consider.

Lemma 3.1.2. *Let B_2^m be defined by 3.8. Then*

$$\begin{aligned}\sum_{m=0}^{M-1} B_2^m \;\leq\; & \sum_{f_k\in F^{ext}} \int_{f_k} \frac{\beta_1}{4Z_k}\left(-\left((n_y)_k\tilde{E}^0_{x_k} - (n_x)_k\tilde{E}^0_{y_k}\right)^2 + \left((n_y)_k\tilde{E}^M_{x_k} - (n_x)_k\tilde{E}^M_{y_k}\right)^2 \right) \\ & + \frac{\beta_2}{2}\left(\tilde{H}^{1/2}_{z_k}\left((n_x)_k\tilde{E}^0_{y_k} - (n_y)_k\tilde{E}^0_{x_k} - \frac{\beta_3}{2Y_k}\tilde{H}^{1/2}_{z_k}\right)\right. \\ & \left. -\tilde{H}^{M+1/2}_{z_k}\left((n_x)_k\tilde{E}^M_{y_k} - (n_y)_k\tilde{E}^M_{x_k} - \frac{\beta_3}{2Y_k}\tilde{H}^{M+1/2}_{z_k}\right)\right) ds,\end{aligned}$$

where $\beta_1 = \alpha, \beta_2 = 0$ for PEC, $\beta_1 = 0, \beta_2 = 1$, $\beta_3 = \alpha$ for PMC, and $\beta_1 = \beta_2 = \frac{1}{2}$, $\beta_3 = 1$ for Silver-Müller boundary conditions.

Proof. First we consider PEC boundary conditions. We have

$$B_2^m = \sum_{f_k\in F^{ext}} \int_{f_k} \frac{\alpha}{Z_k}\left((n_y)_k\left((n_x)_k\tilde{E}^m_{y_k} - (n_y)_k\tilde{E}^m_{x_k}\right)\tilde{E}^{[m+1/2]}_{x_k} - (n_x)_k\left((n_x)_k\tilde{E}^m_{y_k} - (n_y)_k\tilde{E}^m_{x_k}\right)\tilde{E}^{[m+1/2]}_{y_k}\right) ds.$$

Summing from $m = 0$ to $m = M-1$ we obtain

$$\begin{aligned}\sum_{m=0}^{M-1} B_2^m \;=\; & \sum_{f_k\in F^{ext}} \int_{f_k} \frac{\alpha}{4Z_k}\left(-\left((n_x)_k\tilde{E}^0_{y_k} - (n_y)_k\tilde{E}^0_{x_k}\right)^2 + \left((n_x)_k\tilde{E}^M_{y_k} - (n_y)_k\tilde{E}^M_{x_k}\right)^2 \right. \\ & \left. -4\sum_{m=0}^{M-1}\left((n_x)_k\tilde{E}^{[m+1/2]}_{y_k} - (n_y)_k\tilde{E}^{[m+1/2]}_{x_k}\right)^2\right) ds,\end{aligned}$$

and then

$$\sum_{m=0}^{M-1} B_2^m \;\leq\; \sum_{f_k\in F^{ext}} \int_{f_k} \frac{\alpha}{4Z_k}\left(-\left((n_x)_k\tilde{E}^0_{y_k} - (n_y)_k\tilde{E}^0_{x_k}\right)^2 + \left((n_x)_k\tilde{E}^M_{y_k} - (n_y)_k\tilde{E}^M_{x_k}\right)^2 \right) ds. \quad (3.11)$$

For PMC boundary conditions we have

$$\begin{aligned}B_2^m \;=\; & \sum_{f_k\in F^{ext}} \int_{f_k} \left(\tilde{H}^{m+1/2}_{z_k}\left((n_x)_k\tilde{E}^{[m+1/2]}_{y_k} - (n_y)_k\tilde{E}^{[m+1/2]}_{x_k}\right)\right. \\ & \left. -\left(\frac{\alpha}{Y_k}\tilde{H}^{m+1/2}_{z_k} + (n_x)_k\tilde{E}^{m+1}_{y_k} - (n_y)_k\tilde{E}^{m+1}_{x_k}\right)\tilde{H}^{[m+1]}_{z_k}\right) ds.\end{aligned}$$

Summing from $m = 0$ to $m = M-1$ results

$$\begin{aligned}\sum_{m=0}^{M-1} B_2^m \;=\; & \sum_{f_k\in F^{ext}} \int_{f_k} \left(\frac{\tilde{H}^{1/2}_{z_k}}{2}\left((n_x)_k\tilde{E}^0_{y_k} - (n_y)_k\tilde{E}^0_{x_k}\right) - \frac{\alpha}{4Y_k}\left(\tilde{H}^{1/2}_{z_k}\right)^2 - \sum_{m=0}^{M-1}\frac{\alpha}{Y_k}\left(\tilde{H}^{[m+1]}_{z_k}\right)^2\right. \\ & \left. -\frac{\tilde{H}^{M+1/2}_{z_k}}{2}\left((n_x)_k\tilde{E}^M_{y_k} - (n_y)_k\tilde{E}^M_{x_k}\right) - \frac{\alpha}{4Y_k}\left(\tilde{H}^{M+1/2}_{z_k}\right)^2\right) ds,\end{aligned}$$

so we have

$$\sum_{m=0}^{M-1} B_2^m \leq \frac{1}{2}\sum_{f_k\in F^{ext}}\int_{f_k}\left(\tilde{H}_{z_k}^{1/2}\left((n_x)_k\tilde{E}_{y_k}^0-(n_y)_k\tilde{E}_{x_k}^0-\frac{\alpha}{2Y_k}\tilde{H}_{z_k}^{1/2}\right)\right.$$
$$\left.-\tilde{H}_{z_k}^{M+1/2}\left((n_x)_k\tilde{E}_{y_k}^M-(n_y)_k\tilde{E}_{x_k}^M-\frac{\alpha}{2Y_k}\tilde{H}_{z_k}^{M+1/2}\right)\right)ds. \tag{3.12}$$

For Silver-Müller absorbing boundary conditions we have

$$\begin{aligned}B_2^m = \frac{1}{2}\sum_{f_k\in F^{ext}}\int_{f_k}\left(\left(-(n_y)_k\tilde{H}_{z_k}^{m+1/2}+\frac{(n_y)_k}{Z_k}\left((n_x)_k\tilde{E}_{y_k}^m-(n_y)_k\tilde{E}_{x_k}^m\right)\right)\tilde{E}_{x_k}^{[m+1/2]}\right.\\ +\left((n_x)_k\tilde{H}_{z_k}^{m+1/2}-\frac{(n_x)_k}{Z_k}\left((n_x)_k\tilde{E}_{y_k}^m-(n_y)_k\tilde{E}_{x_k}^m\right)\right)\tilde{E}_{y_k}^{[m+1/2]}\\ \left.-\left(\frac{1}{Y_k}\tilde{H}_{z_k}^{m+1/2}+(n_x)_k\tilde{E}_{y_k}^{m+1}-(n_y)_k\tilde{E}_{x_k}^{m+1}\right)\tilde{H}_{z_k}^{[m+1]}\right)ds.\end{aligned}$$

Summing from $m=0$ to $m=M-1$ and taking into account the previous cases, we deduce that

$$\sum_{m=0}^{M-1} B_2^m \leq \sum_{f_k\in F^{ext}}\int_{f_k}\frac{1}{8Z_k}\left(-\left((n_y)_k\tilde{E}_{x_k}^0-(n_x)_k\tilde{E}_{y_k}^0\right)^2+\left((n_y)_k\tilde{E}_{x_k}^M-(n_x)_k\tilde{E}_{y_k}^M\right)^2\right)$$
$$+\frac{1}{4}\left(\tilde{H}_{z_k}^{1/2}\left((n_x)_k\tilde{E}_{y_k}^0-(n_y)_k\tilde{E}_{x_k}^0-\frac{1}{2Y_k}\tilde{H}_{z_k}^{1/2}\right)\right.$$
$$\left.-\tilde{H}_{z_k}^{M+1/2}\left((n_x)_k\tilde{E}_{y_k}^M-(n_y)_k\tilde{E}_{x_k}^M-\frac{1}{2Y_k}\tilde{H}_{z_k}^{M+1/2}\right)\right)ds. \tag{3.13}$$

Considering the estimates we found for PEC, PMC and SM-ABC in (3.11)–(3.13) and employing the parameters β_1, β_2 and β_3 concludes the proof. □

By theses two lemmas, we found an upper bound for (3.4). Now we recall (3.3) and derive the stability condition of the scheme for different fluxes and different types of boundary conditions in the following theorem.

Theorem 3.1.3. *Let us consider the leap-frog DG method (2.32)–(2.34) complemented with the discrete boundary conditions defined in Subsection 2.2.4. If the time step Δt is such that*

$$\Delta t < \frac{\min\{\underline{\varepsilon},\underline{\mu}\}}{\max\{C_E,C_H\}}\min\{h_k\}, \tag{3.14}$$

where

$$C_E = \frac{1}{2}C_{inv}N^2+C_\tau^2(N+1)(N+2)\left(2+\beta_2+\frac{2\alpha+\beta_1}{2\min\{Z_k\}}\right),$$
$$C_H = \frac{1}{2}C_{inv}N^2+C_\tau^2(N+1)(N+2)\left(2+\beta_2+\frac{\alpha+\beta_2\beta_3}{\min\{Y_k\}}\right),$$

with C_τ defined by (A.4) of Lemma A.1.1 and C_{inv} defined by (A.6) of Lemma A.1.2, and $\beta_1=\alpha,\beta_2=0$ for PEC, $\beta_1=0,\beta_2=1$, $\beta_3=\alpha$ for PMC, and $\beta_1=\beta_2=\frac{1}{2}$, $\beta_3=1$ for Silver-Müller boundary conditions, then the method is stable.

Proof. From (3.3) and the previous lemmata (Lemma 3.1.1 and Lemma 3.1.2), considering the Cauchy-Schwarz's inequality and taking into account that $Z_i/(Z_i+Z_k)<1$, we obtain

$$\begin{aligned}
\sum_{T_k\in\mathscr{T}_h}\left(\left(\varepsilon\tilde{E}_k^M,\tilde{E}_k^M\right)_{T_k}+\left(\mu\tilde{H}_{z_k}^{M+1/2},\tilde{H}_{z_k}^{M+1/2}\right)_{T_k}\right) &\leq \sum_{T_k\in\mathscr{T}_h}\left(\left(\varepsilon\tilde{E}_k^0,\tilde{E}_k^0\right)_{T_k}+\left(\mu\tilde{H}_{z_k}^{1/2},\tilde{H}_{z_k}^{1/2}\right)_{T_k}\right)\\
&+\Delta t\sum_{T_k\in\mathscr{T}_h}\left(\|\nabla\times\tilde{H}_{z_k}^{1/2}\|_{L^2(T_k)}\|\tilde{E}_k^0\|_{L^2(T_k)}+\|\nabla\times\tilde{H}_{z_k}^{M+1/2}\|_{L^2(T_k)}\|\tilde{E}_k^M\|_{L^2(T_k)}\right)\\
&+2\Delta t\sum_{f_{ik}\in F^{int}}\left(\|\tilde{E}_k^M\|_{L^2(f_{ik})}\|[\tilde{H}_{z_k}^{M+1/2}]\|_{L^2(f_{ik})}+\|\tilde{E}_k^0\|_{L^2(f_{ik})}\|[\tilde{H}_{z_k}^{1/2}]\|_{L^2(f_{ik})}\right)\\
&+\frac{\alpha\Delta t}{4\min\{Z_k\}}\sum_{f_{ik}\in F^{int}}\|[\tilde{E}_k^M]\|_{L^2(f_{ik})}^2+\frac{\alpha\Delta t}{4\min\{Y_k\}}\sum_{f_{ik}\in F^{int}}\|[\tilde{H}_{z_k}^{M+1/2}]\|_{L^2(f_{ik})}^2\\
&+\frac{\beta_1\Delta t}{2\min\{Z_k\}}\sum_{f_k\in F^{ext}}\|\tilde{E}_k^M\|_{L^2(f_k)}^2+\frac{\beta_2\beta_3\Delta t}{\min\{Y_k\}}\sum_{f_k\in F^{ext}}\|\tilde{H}_{z_k}^{M+1/2}\|_{L^2(f_k)}^2\\
&+2\beta_2\Delta t\sum_{f_k\in F^{ext}}\left(\|\tilde{H}_{z_k}^{1/2}\|_{L^2(f_k)}\|\tilde{E}_k^0\|_{L^2(f_k)}+\|\tilde{H}_{z_k}^{M+1/2}\|_{L^2(f_k)}\|\tilde{E}_k^M\|_{L^2(f_k)}\right).
\end{aligned}$$

Using the inequality (A.4) of Lemma A.1.1 and the inequality (A.6) of Lemma A.1.2 (both in Appendix), we get

$$\begin{aligned}
\min\{\underline{\varepsilon},\underline{\mu}\}\left(\|\tilde{E}^M\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\right) &\leq \max\{\bar{\varepsilon},\bar{\mu}\}\left(\|\tilde{E}^0\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2\right)\\
&+\frac{\Delta t}{2}C_{inv}N^2\max\left\{h_k^{-1}\right\}\left(\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2+\|\tilde{E}^0\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2+\|\tilde{E}^M\|_{L^2(\Omega)}^2\right)\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+\beta_1}{2\min\{Z_k\}}\right)\|\tilde{E}^M\|_{L^2(\Omega)}^2\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+2\beta_2\beta_3}{2\min\{Y_k\}}\right)\|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}(2+\beta_2)\left(\|\tilde{E}^0\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2\right).
\end{aligned}$$

and so, taking $C_0=\frac{1}{2}C_{inv}N^2+C_\tau^2(N+1)(N+2)(2+\beta_2)$,

$$\begin{aligned}
\left(\min\{\underline{\varepsilon},\underline{\mu}\}-\Delta t\max\left\{h_k^{-1}\right\}\max\{C_E,C_H\}\right)\left(\|\tilde{E}^M\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\right) &\leq\\
\left(\max\{\bar{\varepsilon},\bar{\mu}\}+\Delta t\max\left\{h_k^{-1}\right\}C_0\right)\left(\|\tilde{E}^0\|_{L^2(\Omega)}^2+\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2\right),
\end{aligned}$$

which concludes the proof. □

The stability condition (3.14) shows that the method is conditionally stable, which is natural since we considered an explicit time discretization. Further, it discloses the influence of the values of α, h_{min} and N on the bounds of the stable region. This is of utmost importance to balance accuracy *versus* stability.

3.2 Stability analysis of the 3D model

In this section we extend the analysis in Section 3.1 of the TE form of Maxwell's equations in two-dimensions to the full three-dimensional time-dependent Maxwell's equations (2.1) and (2.2), with the equations set on a bounded polyhedral domain $\Omega\subset\mathbb{R}^3$.

We start by noticing that the following vector equalities hold

$$u \times v \cdot w = -u \times w \cdot v, \tag{3.15}$$

and

$$u \times (v \times w) = v(u \cdot w) - w(u \cdot v). \tag{3.16}$$

Recall the leap-frog DG scheme (2.35)–(2.36). Let $u_k = \tilde{E}_k^{[m+1/2]}$ and $v_k = \tilde{H}_k^{[m+1]}$ in (2.35) and (2.36) respectively, we have

$$\begin{aligned}\left(\varepsilon \tilde{E}_k^{m+1}, \tilde{E}_k^{m+1}\right)_{T_k} - \left(\varepsilon \tilde{E}_k^{m}, \tilde{E}_k^{m}\right)_{T_k} &= 2\Delta t \left(\nabla \times \tilde{H}_k^{m+1/2}, \tilde{E}_k^{[m+1/2]}\right)_{T_k} \\ -2\Delta t \left(\frac{1}{Z^+ + Z^-} n \times \left(Z^+ [\tilde{H}^{m+1/2}] - \alpha n \times [\tilde{E}^m]\right), \tilde{E}_k^{[m+1/2]}\right)_{\partial T_k}\end{aligned} \tag{3.17}$$

and

$$\begin{aligned}\left(\mu \tilde{H}_k^{m+3/2}, \tilde{H}_k^{m+3/2}\right)_{T_k} - \left(\mu \tilde{H}_k^{m+1/2}, \tilde{H}_k^{m+1/2}\right)_{T_k} &= -2\Delta t \left(\nabla \times \tilde{E}_k^{m+1}, \tilde{H}_k^{[m+1]}\right)_{T_k} \\ +2\Delta t \left(\frac{1}{Y^+ + Y^-} n \times \left(Y^+ [\tilde{E}^{m+1}] + \alpha n \times [\tilde{H}^{m+1/2}]\right), \tilde{H}_k^{[m+1]}\right)_{\partial T_k}.\end{aligned} \tag{3.18}$$

Using the identity

$$\left(\nabla \times \tilde{E}_k^{m+1}, \tilde{H}_k^{[m+1]}\right)_{T_k} = \left(\nabla \times \tilde{H}_k^{[m+1]}, \tilde{E}_k^{m+1}\right)_{T_k} + \left(n \times \tilde{E}_k^{m+1}, \tilde{H}_k^{[m+1]}\right)_{\partial T_k},$$

and summing (2.35) and (2.36) from $m = 0$ to $m = M-1$, and integrating by parts, we get

$$\begin{aligned}\left(\varepsilon \tilde{E}_k^{M}, \tilde{E}_k^{M}\right)_{T_k} + \left(\mu \tilde{H}_k^{M+1/2}, \tilde{H}_k^{M+1/2}\right)_{T_k} &= \left(\varepsilon \tilde{E}_k^{0}, \tilde{E}_k^{0}\right)_{T_k} + \left(\mu \tilde{H}_k^{1/2}, \tilde{H}_{z_k}^{1/2}\right)_{T_k} \\ +\Delta t \left(\nabla \times \tilde{H}_k^{1/2}, \tilde{E}_k^{0}\right)_{T_k} - \Delta t \left(\nabla \times \tilde{H}_k^{M+1/2}, \tilde{E}_k^{M}\right)_{T_k} &+ 2\Delta t \sum_{m=0}^{M-1} A_k^m,\end{aligned} \tag{3.19}$$

where

$$\begin{aligned}A_k^m = &- \left(\frac{1}{Z^+ + Z^-} n \times \left(Z^+ [\tilde{H}^{m+1/2}] - \alpha n \times [\tilde{E}^m]\right), \tilde{E}_k^{[m+1/2]}\right)_{\partial T_k} \\ &+ \left(\frac{1}{Y^+ + Y^-} n \times \left(Y^+ [\tilde{E}^{m+1}] + \alpha n \times [\tilde{H}^{m+1/2}]\right), \tilde{H}_k^{[m+1]}\right)_{\partial T_k} - \left(n \times \tilde{E}_k^{m+1}, \tilde{H}_k^{[m+1]}\right)_{\partial T_k}.\end{aligned}$$

Let us consider the following decomposition

$$\sum_{T_k \in \mathcal{T}_h} A_k^m = B_1^m + B_2^m,$$

where B_1 corresponds to the terms on the internal edges

$$\begin{aligned} B_1^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& \frac{-1}{Z_i+Z_k} n_{ki} \times \left(Z_i [\tilde{H}_k^{m+1/2}] - \alpha n_{ki} \times [\tilde{E}_k^m] \right) \cdot \tilde{E}_k^{[m+1/2]} \\ & + \frac{1}{Z_i+Z_k} n_{ki} \times \left(Z_k [\tilde{H}_i^{m+1/2}] + \alpha n_{ki} \times [\tilde{E}_i^m] \right) \cdot \tilde{E}_i^{[m+1/2]} \\ & + \frac{1}{Y_i+Y_k} n_{ki} \times \left(Y_i [\tilde{E}_k^{m+1}] + \alpha n_{ki} \times [\tilde{H}_k^{m+1/2}] \right) \cdot \tilde{H}_k^{[m+1]} \\ & - \frac{1}{Y_i+Y_k} n_{ki} \times \left(Y_k [\tilde{E}_i^{m+1}] - \alpha n_{ki} \times [\tilde{H}_i^{m+1/2}] \right) \cdot \tilde{H}_i^{[m+1]} \\ & - n_{ki} \times \tilde{E}_k^{m+1} \cdot \tilde{H}_k^{[m+1]} + n_{ki} \times \tilde{E}_i^{m+1} \cdot \tilde{H}_i^{[m+1]} \Bigg) \, ds, \end{aligned} \tag{3.20}$$

and B_2 corresponds to the terms on the boundary $\partial\Omega$

$$\begin{aligned} B_2^m = \sum_{f_k\in F^{ext}} \int_{f_k} \Bigg(& - \frac{1}{2Z_k} n_k \times \left(Z_k [\tilde{H}_k^{m+1/2}] - \alpha n_k \times [\tilde{E}_k^m] \right) \cdot \tilde{E}_k^{[m+1/2]} \\ & + \frac{1}{2Y_k} n_k \times \left(Y_k [\tilde{E}_k^{m+1}] + \alpha n_k \times [\tilde{H}_k^{m+1/2}] \right) \cdot \tilde{H}_k^{[m+1]} - n_k \times \tilde{E}_k^{m+1} \cdot \tilde{H}_k^{[m+1]} \Bigg) \, ds. \end{aligned} \tag{3.21}$$

Lemma 3.2.1. *Let B_1^m be defined by (3.20). Then*

$$\begin{aligned} \sum_{m=0}^{M-1} B_1^m \leq \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& \frac{\alpha}{4(Z_i+Z_k)} [\tilde{E}_k^M] \cdot [\tilde{E}_k^M] + \frac{\alpha}{4(Y_i+Y_k)} [\tilde{H}_k^{M+1/2}] \cdot [\tilde{H}_k^{M+1/2}] \\ & + \frac{Z_k}{2(Z_i+Z_k)} \left(n_{ki} \times [\tilde{H}_k^{1/2}] \cdot [E_k^0] - n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot [\tilde{E}_k^M] \right) \\ & + \frac{1}{2} \left(n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot \tilde{E}_k^M - n_{ki} \times [\tilde{H}_k^{1/2}] \cdot \tilde{E}_k^0 \right) \Bigg) \, ds. \end{aligned}$$

Proof. Summing from $m=0$ to $m=M-1$, using the relations (3.9), (3.10) and employing the vector equalities (3.15) and (3.16), we obtain

$$\begin{aligned}\sum_{m=0}^{M-1} B_1^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& \frac{\alpha}{4(Z_i+Z_k)} \Bigg(- [\tilde{E}_k^0]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{E}_k^0] + [\tilde{E}_k^M]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{E}_k^M] \\ & - \sum_{m=0}^{M-1} \left([\tilde{E}_k^m] + [\tilde{E}_k^{m+1}]\right)^T \left(I - n_{ki}n_{ik}^T\right)\left([\tilde{E}_k^m] + [\tilde{E}_k^{m+1}]\right) \Bigg) \\ & + \frac{\alpha}{4(Y_i+Y_k)} \Bigg(- [\tilde{H}_k^{1/2}]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{H}_k^{1/2}] + [\tilde{H}_k^{M+1/2}]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{H}_k^{M+1/2}] \\ & - \sum_{m=0}^{M-1} \left([\tilde{H}_k^{m+1/2}] + [\tilde{H}_k^{m+3/2}]\right)^T \left(I - n_{ki}n_{ik}^T\right)\left([\tilde{H}_k^{m+1/2}] + [\tilde{H}_k^{m+3/2}]\right) \Bigg) \\ & + \frac{Z_k}{2(Z_i+Z_k)} \left(n_{ki} \times [\tilde{H}_k^{1/2}] \cdot [E_k^0] - n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot [\tilde{E}_k^M]\right) \\ & + \frac{1}{2} \left(n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot \tilde{E}_k^M - n_{ki} \times [\tilde{H}_k^{1/2}] \cdot \tilde{E}_k^0\right) \Bigg) ds,\end{aligned}$$

where I is an identity matrix. Since $I - n_{ki}n_{ik}^T$ is a positive semidefinite matrix,

$$\begin{aligned}\sum_{m=0}^{M-1} B_1^m \leq \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& \frac{\alpha}{4(Z_i+Z_k)} \left([\tilde{E}_k^M]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{E}_k^M] \right) \\ & + \frac{\alpha}{4(Y_i+Y_k)} \left([\tilde{H}_k^{M+1/2}]^T \left(I - n_{ki}n_{ik}^T\right)[\tilde{H}_k^{M+1/2}] \right) \\ & + \frac{Z_k}{2(Z_i+Z_k)} \left(n_{ki} \times [\tilde{H}_k^{1/2}] \cdot [E_k^0] - n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot [\tilde{E}_k^M]\right) \\ & + \frac{1}{2} \left(n_{ki} \times [\tilde{H}_k^{M+1/2}] \cdot \tilde{E}_k^M - n_{ki} \times [\tilde{H}_k^{1/2}] \cdot \tilde{E}_k^0\right) \Bigg) ds.\end{aligned}$$

The proof follows from the fact that, since the matrix $I - n_{ki}n_{ik}^T$ is an orthogonal projector,

$$x^T (I - n_{ki}n_{ik}^T)x \leq x^T x \quad \text{for all vector} \quad x.$$

□

Lemma 3.2.2. *Let B_2^m be defined by (3.21).Then*

$$\begin{aligned}\sum_{m=0}^{M-1} B_2^m \leq \sum_{f_k\in F^{ext}} \int_{f_k} \Bigg(& \frac{\beta_1}{4Z_k} \Bigg(- \left(n_k \times \tilde{E}_k^0\right) \cdot \left(n_k \times \tilde{E}_k^0\right) + \left(n_k \times \tilde{E}_k^M\right) \cdot \left(n_k \times \tilde{E}_k^M\right) \Bigg) \\ & + \frac{\beta_3}{4Y_k} \Bigg(- \left(n_k \times \tilde{H}_k^{1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{1/2}\right) + \left(n_k \times \tilde{H}_k^{M+1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{M+1/2}\right) \Bigg) \\ & + \frac{\beta_2}{2} \left(n_k \times \tilde{E}_k^0 \cdot \tilde{H}_k^{1/2} - n_k \times \tilde{E}_k^M \cdot \tilde{H}_k^{M+1/2}\right) \Bigg) ds,\end{aligned}$$

where $\beta_1 = \alpha, \beta_2 = 0$ *for PEC,* $\beta_1 = 0, \beta_2 = 1$, $\beta_3 = \alpha$ *for PMC, and* $\beta_1 = \beta_2 = \beta_3 = \frac{1}{2}$ *for Silver-Müller boundary conditions.*

Proof. First we consider PEC boundary conditions. We have

$$B_2^m = \sum_{f_k \in F^{ext}} \int_{f_k} \frac{1}{Z_k} n_k \times \left(\alpha n_k \times \tilde{E}_k^m\right) \cdot \tilde{E}_k^{[m+1/2]} \, ds,$$

and then

$$\sum_{m=0}^{M-1} B_2^m \leq \sum_{f_k \in F^{ext}} \int_{f_k} \frac{\alpha}{4Z_k} \left(- \left(n_k \times \tilde{E}_k^0\right) \cdot \left(n_k \times \tilde{E}_k^0\right) + \left(n_k \times \tilde{E}_k^M\right) \cdot \left(n_k \times \tilde{E}_k^M\right) \right) ds.$$

If we consider PMC boundary conditions we have

$$B_2^m = \sum_{f_k \in F^{ext}} \int_{f_k} \left(- n_k \times \tilde{H}_k^{m+1/2} \cdot \tilde{E}_k^{[m+1/2]} + \frac{\alpha}{Y_k} n_k \times \left(n_k \times \tilde{H}_k^{m+1/2}\right) \cdot \tilde{H}_k^{[m+1]} - n_k \times \tilde{E}_k^{m+1} \cdot \tilde{H}_k^{[m+1]} \right) ds,$$

and then

$$\sum_{m=0}^{M-1} B_2^m \leq \sum_{f_k \in F^{ext}} \int_{f_k} \left(\frac{\alpha}{4Y_k} \left(- \left(n_k \times \tilde{H}_k^{1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{1/2}\right) + \left(n_k \times \tilde{H}_k^{M+1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{M+1/2}\right) \right) + \frac{1}{2} n_k \times \tilde{E}_k^0 \cdot \tilde{H}_k^{1/2} - \frac{1}{2} n_k \times \tilde{E}_k^M \cdot \tilde{H}_k^{M+1/2} \right) ds.$$

For Silver-Müler boundary conditions, we have

$$\begin{aligned} B_2^m &= \sum_{f_k \in F^{ext}} \int_{f_k} \left(- \frac{1}{2Z_k} n_k \times \left(Z_k \tilde{H}_k^{m+1/2} - n_k \times \tilde{E}_k^m\right) \cdot \tilde{E}_k^{[m+1/2]} \right. \\ &\qquad \left. + \frac{1}{2Y_k} n_k \times \left(Y_k \tilde{E}_k^{m+1} + n_k \times \tilde{H}_k^{m+1/2}\right) \cdot \tilde{H}_k^{[m+1]} - n_k \times \tilde{E}_k^{m+1} \cdot \tilde{H}_k^{[m+1]} \right) ds \\ &= \sum_{f_k \in F^{ext}} \int_{f_k} \left(- \frac{1}{2Z_k} n_k \times \left(Z_k \tilde{H}_k^{m+1/2} - n_k \times \tilde{E}_k^m\right) \cdot \tilde{E}_k^{[m+1/2]} \right. \\ &\qquad \left. + \frac{1}{2Y_k} n_k \times \left(n_k \times \tilde{H}_k^{m+1/2}\right) \cdot \tilde{H}_k^{[m+1]} . - \frac{1}{2} n_k \times \tilde{E}_k^{m+1} \cdot \tilde{H}_k^{[m+1]} \right) ds, \end{aligned}$$

and then

$$\sum_{m=0}^{M-1} B_2^m \leq \sum_{f_k \in F^{ext}} \int_{f_k} \left(\frac{1}{8Z_k} \left(-\left(n_k \times \tilde{E}_k^0\right) \cdot \left(n_k \times \tilde{E}_k^0\right) + \left(n_k \times \tilde{E}_k^M\right) \cdot \left(n_k \times \tilde{E}_k^M\right) \right) \right.$$
$$+ \frac{1}{8Y_k} \left(-\left(n_k \times \tilde{H}_k^{1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{1/2}\right) + \left(n_k \times \tilde{H}_k^{M+1/2}\right) \cdot \left(n_k \times \tilde{H}_k^{M+1/2}\right) \right)$$
$$\left. + \frac{1}{4} n_k \times \tilde{E}_k^0 \cdot \tilde{H}_k^{1/2} - \frac{1}{4} n_k \times \tilde{E}_k^M \cdot \tilde{H}_k^{M+1/2} \right) ds,$$

which concludes the proof. □

Employing the estimates we already achieved in lemmas 3.2.1 and 3.2.2 and using Lemma A.1.1 and Lemma A.1.2 (both in Appendix) we will show in the following theorem that the 3D model is conditionally stable.

Theorem 3.2.3. *Let us consider the leap-frog DG method (2.35)–(2.36) complemented with the discrete boundary conditions. If the time step Δt is such that*

$$\Delta t < \frac{\min\{\underline{\varepsilon}, \underline{\mu}\}}{\max\{C_E, C_H\}} \min\{h_k\}, \tag{3.22}$$

where

$$C_E = \frac{1}{2} C_{inv} N^2 + C_\tau^2 (N+1)(N+3) \left(3 + \frac{\beta_2}{2} + \frac{\alpha + \beta_1}{2 \min\{Z_k\}} \right),$$
$$C_H = \frac{1}{2} C_{inv} N^2 + C_\tau^2 (N+1)(N+3) \left(3 + \frac{\beta_2}{2} + \frac{\alpha + \beta_3}{2 \min\{Y_k\}} \right),$$

with C_τ defined by (A.5) of Lemma A.1.1 and C_{inv} defined by (A.6) of Lemma A.1.2, and $\beta_1 = \alpha, \beta_2 = 0$ for PEC, $\beta_1 = 0, \beta_2 = 1$, $\beta_3 = \alpha$ for PMC, and $\beta_1 = \beta_2 = \frac{1}{2}$, $\beta_3 = 1$ for Silver-Müller boundary conditions, then the method is stable.

Proof. As for the 2D case, from (3.19) and the previous lemmata, considering the Cauchy-Schwarz's and triangular inequality inequality, taking into account that $Z_i/(Z_i + Z_k) < 1$, and using the inequality (A.5) of Lemma A.1.1 and the inequality (A.6) of Lemma A.1.2 (both in Appendix), we get

$$\min\{\underline{\varepsilon}, \underline{\mu}\} \left(\|\tilde{E}^M\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2 \right) \leq \max\{\bar{\varepsilon}, \bar{\mu}\} \left(\|\tilde{E}^0\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2 \right)$$
$$+ \frac{\Delta t}{2} C_{inv} N^2 \max\left\{h_k^{-1}\right\} \left(\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2 + \|\tilde{E}^0\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2 + \|\tilde{E}^M\|_{L^2(\Omega)}^2 \right)$$
$$+ C_\tau^2 (N+1)(N+3) \Delta t \max\left\{h_k^{-1}\right\} \left(3 + \frac{\beta_2}{2} + \frac{\alpha + \beta_1}{2 \min\{Z_k\}} \right) \left(\|\tilde{E}^0\|_{L^2(\Omega)}^2 + \|\tilde{E}^M\|_{L^2\Omega}^2 \right)$$
$$+ C_\tau^2 (N+1)(N+3) \Delta t \max\left\{h_k^{-1}\right\} \left(3 + \frac{\beta_2}{2} + \frac{\alpha + \beta_3}{2 \min\{Y_k\}} \right) \left(\|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2 \right).$$

and so

$$\left(\min\{\underline{\varepsilon}, \underline{\mu}\} - \Delta t \max\left\{h_k^{-1}\right\} \max\{C_E, C_H\} \right) \left(\|\tilde{E}^M\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}^2 \right)$$
$$\leq \left(\max\{\bar{\varepsilon}, \bar{\mu}\} + \Delta t \max\left\{h_k^{-1}\right\} \max\{C_E, C_H\} \right) \left(\|\tilde{E}^0\|_{L^2(\Omega)}^2 + \|\tilde{H}_z^{1/2}\|_{L^2(\Omega)}^2 \right),$$

which proves the result. □

Chapter 4

Convergence Analysis

A rigorous proof of convergence is necessary for every numerical method. Error bounds for DG methods applied to Maxwell's equations were determined for upwind flux in [41], where the error of the semi-discrete approximation is estimated and for central flux in [32], where the error of the full-discrete approximation of Maxwell's equation is estimated. In this chapter we present the error estimate of fully discrete leap-frog DG scheme in 2D and analyze the convergence properties.

The core of this chapter is Theorem 4.1.5 where the spatial and temporal error of the scheme (2.32)–(2.34) is estimated. The convergence analysis of the scheme shows that in the case of Silver-Müller absorbing boundary conditions, the temporal order is reduced from two to one. In order to recover the temporal order, a predictor-corrector time integrator is proposed in that last part of Section 4.1. This idea is extended to an iterative predictor-corrector time integrator in Section 4.2. such that the scheme remains explicit. We will show that the iterative predictor-corrector scheme converges to a second order convergent in time implicit method. The stability analysis of the implicit method is also presented.

4.1 Error estimate

The main result of this section is Theorem 4.1.5 which presents the error of the fully discrete approximation of 2D Maxwell's equations. The key idea for the proof is to find a variational system for the difference between the the numerical solution and a projection of the exact solution, (E_x, E_y, H_z) onto the finite element space V_N. We estimate the error between the numerical solution and a polynomial approximation, where the polynomial approximation $(\mathscr{P}_N E_x, \mathscr{P}_N E_y, \mathscr{P}_N H_z)$ is an interpolant of (E_x, E_y, H_z) having the optimal approximation errors (A.7)–(A.8). Lemma A.2.1 in the Appendix furnishes an optimal error estimation which plays a central role in our derivation.

The spatial and temporal oder of convergence will be achieved simultaneously, while the temporal order in the case of Silver-Müller drops to first order. In order to recover, a modified method is proposed in Remark 4.1.6. This idea will be generalized and analyzed later in the next section.

To provide a proper functional setting, we need to define spaces involving time-dependent functions [29]. Let X denote a Banach space with norm $\|.\|_X$. The spaces $L^2(0,T;X)$ and $L^\infty(0,T;X)$ consist,

respectively, of all measurable functions $v : [0,T] \to X$ with

$$\|v\|_{L^2(0,T;X)} = \left(\int_0^T \|v(t)\|_X^2 \, dt \right)^{1/2} < \infty, \quad \|v\|_{L^\infty(0,T;X)} = ess \sup_{0 \le t \le T} \|v(t)\|_X < \infty.$$

In what follows, X is shorthand for any of the usual Sobolev spaces $H^p(\Omega)$ or the Banach space $L^\infty(\Omega)$.

The idea of finding an upper bound for $\|q^m - \tilde{q}^m\|_{L^2(\Omega)}$ is to split the error in two: the error of polynomial approximation and the error between of numerical and polynomial approximation where, with triangle inequality we have

$$\|q^m - \tilde{q}^m\|_{L^2(\Omega)} \le \|q^m - \mathscr{P}_N q^m\|_{L^2(\Omega)} + \|\mathscr{P}_N q^m - \tilde{q}^m\|_{L^2(\Omega)}.$$

In our DG method, we search the approximate solution in the piecewise polynomial space V_N. Let $(\mathscr{P}_N E_x, \mathscr{P}_N E_y, \mathscr{P}_N H_z) \in V_N$ be an interpolant of (E_x, E_y, H_z) having the optimal approximation errors (A.7)–(A.8). On the external boundary we define the jumps for polynomial approximation $(\mathscr{P}_N E_x, \mathscr{P}_N E_y, \mathscr{P}_N H_z)$ for different boundary conditions analogous to what we have for numerical approximation in the following way.

- For PEC:

$$[\mathscr{P}_N E_x] = 2\mathscr{P}_N E_x, \quad [\mathscr{P}_N E_y] = 2\mathscr{P}_N E_y, \quad [\mathscr{P}_N H_z] = 0. \tag{4.1}$$

- For PMC:

$$[\mathscr{P}_N E_x] = 0, \quad [\mathscr{P}_N E_y] = 0, \quad [\mathscr{P}_N H_z] = 2\mathscr{P}_N H_z. \tag{4.2}$$

- For SM-ABC:

$$[\mathscr{P}_N E_x] = \mathscr{P}_N E_x, \quad [\mathscr{P}_N E_y] = \mathscr{P}_N E_y, \quad [\mathscr{P}_N H_z] = \mathscr{P}_N H_z. \tag{4.3}$$

In order to find an estimate for $\|\mathscr{P}_N q^m - \tilde{q}^m\|_{L^2(\Omega)}$, we start by integrating (2.9) from t^m to t^{m+1} and (2.10) from $t^{m+1/2}$ to $t^{m+3/2}$. Then, multiplying the resultant by $(u_k, v_k, w_k) \in V_N$ with respect to the L^2-inner product over T_k, we obtain

$$\left(\varepsilon_{xx} \frac{E_{x_k}^{m+1} - E_{x_k}^m}{\Delta t} + \varepsilon_{xy} \frac{E_{y_k}^{m+1} - E_{y_k}^m}{\Delta t}, u_k \right)_{T_k} = \frac{1}{\Delta t} \left(\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y} dt, u_k \right)_{T_k}, \tag{4.4}$$

$$\left(\varepsilon_{yx} \frac{E_{x_k}^{m+1} - E_{x_k}^m}{\Delta t} + \varepsilon_{yy} \frac{E_{y_k}^{m+1} - E_{y_k}^m}{\Delta t}, v_k \right)_{T_k} = -\frac{1}{\Delta t} \left(\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial x} dt, v_k \right)_{T_k}, \tag{4.5}$$

$$\left(\mu\frac{H_{z_k}^{m+3/2}-H_{z_k}^{m+1/2}}{\Delta t},w_k\right)_{T_k} = \frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}}\frac{\partial E_{x_k}}{\partial y}dt,w_k\right)_{T_k}-\frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}}\frac{\partial E_{y_k}}{\partial x}dt,w_k\right)_{T_k}. \quad (4.6)$$

Recalling (2.32)–(2.34) and subtract them from (4.4)–(4.6) and using the notation

$$\xi_{x_k}^m=\mathscr{P}_N E_{x_k}^m-\tilde{E}_{x_k}^m,\qquad \rho_{x_k}^m=\mathscr{P}_N E_{x_k}^m-E_{x_k}^m,$$

$$\xi_{y_k}^m=\mathscr{P}_N E_{y_k}^m-\tilde{E}_{y_k}^m,\qquad \rho_{y_k}^m=\mathscr{P}_N E_{y_k}^m-E_{y_k}^m,$$

and

$$\eta_{z_k}^{m+1/2}=\mathscr{P}_N H_{z_k}^{m+1/2}-\tilde{H}_{z_k}^{m+1/2},\quad \phi_{z_k}^{m+1/2}=\mathscr{P}_N H_{z_k}^{m+1/2}-H_{z_k}^{m+1/2},$$

we obtain

$$\begin{aligned}
&\left(\varepsilon_{xx}\frac{\xi_{x_k}^{m+1}-\xi_{x_k}^m}{\Delta t},u_k\right)_{T_k}-\left(\varepsilon_{xx}\frac{\rho_{x_k}^{m+1}-\rho_{x_k}^m}{\Delta t},u_k\right)_{T_k}+\left(\varepsilon_{xy}\frac{\xi_{y_k}^{m+1}-\xi_{y_k}^m}{\Delta t},u_k\right)_{T_k}-\left(\varepsilon_{xy}\frac{\rho_{y_k}^{m+1}-\rho_{y_k}^m}{\Delta t},u_k\right)_{T_k}\\
&\quad=\frac{1}{\Delta t}\left(\int_{t^m}^{t^{m+1}}\frac{\partial H_{z_k}}{\partial y}dt,u_k\right)_{T_k}-\left(\frac{\partial}{\partial y}\left(\mathscr{P}_N H_{z_k}^{m+1/2}\right),u_k\right)_{T_k}+\left(\frac{\partial\eta_{z_k}^{m+1/2}}{\partial y},u_k\right)_{T_k}\\
&\qquad+\left(\frac{n_y}{Z^++Z^-}\left(Z^+[\mathscr{P}_N H_{z_k}^{m+1/2}]-\alpha\left(n_x[\mathscr{P}_N E_{y_k}^m]-n_y[\mathscr{P}_N E_{x_k}^m]\right)\right),u_k\right)_{\partial T_k}\\
&\qquad+\left(\frac{-n_y}{Z^++Z^-}\left(Z^+[\eta_{z_k}^{m+1/2}]-\alpha\left(n_x[\xi_{y_k}^m]-n_y[\xi_{x_k}^m]\right)\right),u_k\right)_{\partial T_k},
\end{aligned}\quad (4.7)$$

$$\begin{aligned}
&\left(\varepsilon_{yx}\frac{\xi_{x_k}^{m+1}-\xi_{x_k}^m}{\Delta t},v_k\right)_{T_k}-\left(\varepsilon_{yx}\frac{\rho_{x_k}^{m+1}-\rho_{x_k}^m}{\Delta t},v_k\right)_{T_k}+\left(\varepsilon_{yy}\frac{\xi_{y_k}^{m+1}-\xi_{y_k}^m}{\Delta t},v_k\right)_{T_k}-\left(\varepsilon_{yy}\frac{\rho_{y_k}^{m+1}-\rho_{y_k}^m}{\Delta t},v_k\right)_{T_k}\\
&\quad=-\frac{1}{\Delta t}\left(\int_{t^m}^{t^{m+1}}\frac{\partial H_{z_k}}{\partial x}dt,v_k\right)_{T_k}+\left(\frac{\partial}{\partial x}\left(\mathscr{P}_N H_{z_k}^{m+1/2}\right),v_k\right)_{T_k}-\left(\frac{\partial\eta_{z_k}^{m+1/2}}{\partial x},v_k\right)_{T_k}\\
&\qquad-\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\mathscr{P}_N H_{z_k}^{m+1/2}]-\alpha\left(n_x[\mathscr{P}_N E_{y_k}^m]-n_y[\mathscr{P}_N E_{x_k}^m]\right)\right),v_k\right)_{\partial T_k}\\
&\qquad+\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\eta_{z_k}^{m+1/2}]-\alpha\left(n_x[\xi_{y_k}^m]-n_y[\xi_{x_k}^m]\right)\right),v_k\right)_{\partial T_k},
\end{aligned}\quad (4.8)$$

$$
\begin{aligned}
&\left(\mu\frac{\eta_{z_k}^{m+3/2}-\eta_{z_k}^{m+1/2}}{\Delta t},w_k\right)_{T_k}-\left(\mu\frac{\phi_{z_k}^{m+3/2}-\phi_{x_k}^{m+1/2}}{\Delta t},w_k\right)_{T_k}\\
&\quad=\frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}}\frac{\partial E_{x_k}}{\partial y}dt,w_k\right)-\frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}}\frac{\partial E_{y_k}}{\partial x}dt,w_k\right)_{T_k}\\
&\quad-\left(\frac{\partial}{\partial y}\left(\mathscr{P}_N E_{x_k}^{m+1}\right),w_k\right)_{T_k}+\left(\frac{\partial\xi_{x_k}^{m+1}}{\partial y},w_k\right)_{T_k}+\left(\frac{\partial}{\partial x}\left(\mathscr{P}_N E_{y_k}^{m+1}\right),w_k\right)_{T_k}-\left(\frac{\partial\xi_{y_k}^{m+1}}{\partial x},w_k\right)_{T_k}\\
&\quad-\left(\frac{1}{Y^++Y^-}\left(Y^+(n_x[\mathscr{P}_N E_{y_k}^{m+1}]-n_y[\mathscr{P}_N E_{x_k}^{m+1}])-\alpha[\mathscr{P}_N H_{z_k}^{m+1/2}],w_k\right)\right)_{\partial T_k}\\
&\quad+\left(\frac{1}{Y^++Y^-}\left(Y^+(n_x[\xi_{y_k}^{m+1}]-n_y[\xi_{x_k}^{m+1}])-\alpha[\eta_{z_k}^{m+1/2}],w_k\right)\right)_{\partial T_k}.
\end{aligned}
\tag{4.9}
$$

Let $u_k=\xi_{x_k}^m+\xi_{x_k}^{m+1}$, $v_k=\xi_{y_k}^m+\xi_{y_k}^{m+1}$ and $w_k=\eta_{z_k}^{m+1/2}+\eta_{z_k}^{m+3/2}$ in (4.7)–(4.9). Summing from $m=0$ to $m=M-1$ and using the symmetry property of the permittivity tensor ε, we get

$$
\begin{aligned}
\left(\varepsilon\xi_k^M,\xi_k^M\right)_{T_k}+\left(\mu\eta_{z_k}^{M+1/2},\eta_{z_k}^{M+1/2}\right)_{T_k}&=\left(\varepsilon\xi_k^0,\xi_k^0\right)_{T_k}+\left(\mu\eta_{z_k}^{1/2},\eta_{z_k}^{1/2}\right)_{T_k}\\
&+\Delta t\left(\nabla\times\eta_{z_k}^{1/2},\xi_k^0\right)_{T_k}-\Delta t\left(\nabla\times\eta_{z_k}^{M+1/2},\xi_k^M\right)_{T_k}+2\Delta t\sum_{m=0}^{M-1}R_k^m,
\end{aligned}
\tag{4.10}
$$

with

$$
R_k^m=S_{1,k}^m+S_{2,k}^m+S_{3,k}^m+S_{4,k}^m,
$$

being $S_{1,k}^m$, $S_{2,k}^m$, $S_{3,k}^m$ and $S_{4,k}^m$ defined below using the average notation $\xi_{x_k}^{[m+1/2]}=(\xi_{x_k}^m+\xi_{x_k}^{m+1})/2$, $\xi_{y_k}^{[m+1/2]}=(\xi_{y_k}^m+\xi_{y_k}^{m+1})/2$ and $\eta_k^{[m+1]}=(\eta_{z_k}^{m+1/2}+\eta_{z_k}^{m+3/2})/2$,

$$
\begin{aligned}
S_{1,k}^m&=\left(\varepsilon_{xx}\frac{\rho_{x_k}^{m+1}-\rho_{x_k}^m}{\Delta t},\xi_{x_k}^{[m+1/2]}\right)_{T_k}+\left(\varepsilon_{xy}\frac{\rho_{y_k}^{m+1}-\rho_{y_k}^m}{\Delta t},\xi_{x_k}^{[m+1/2]}\right)_{T_k}\\
&+\left(\varepsilon_{yx}\frac{\rho_{x_k}^{m+1}-\rho_{x_k}^m}{\Delta t},\xi_{y_k}^{[m+1/2]}\right)_{T_k}+\left(\varepsilon_{yy}\frac{\rho_{y_k}^{m+1}-\rho_{y_k}^m}{\Delta t},\xi_{y_k}^{[m+1/2]}\right)_{T_k}\\
&+\left(\mu\frac{\phi_{z_k}^{m+3/2}-\phi_{z_k}^{m+1/2}}{\Delta t},\eta_{z_k}^{[m+1]}\right)_{T_k},
\end{aligned}
\tag{4.11}
$$

$$
\begin{aligned}
S_{2,k}^m = &-\left(\frac{\partial}{\partial y}\left(\mathscr{P}_N H_{z_k}^{m+1/2}\right), \xi_{x_k}^{[m+1/2]}\right)_{T_k} + \frac{1}{\Delta t}\left(\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y} dt, \xi_{x_k}^{[m+1/2]}\right)_{T_k} \\
&+\left(\frac{\partial}{\partial x}\left(\mathscr{P}_N H_{z_k}^{m+1/2}\right), \xi_{y_k}^{[m+1/2]}\right)_{T_k} - \frac{1}{\Delta t}\left(\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial x} dt, \xi_{y_k}^{[m+1/2]}\right)_{T_k} \\
&-\left(\frac{\partial}{\partial y}\left(\mathscr{P}_N E_{x_k}^{m+1}\right), \eta_{z_k}^{[m+1]}\right)_{T_k} + \frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}} \frac{\partial E_{x_k}}{\partial y} dt, \eta_{z_k}^{[m+1]}\right)_{T_k} \\
&+\left(\frac{\partial}{\partial x}\left(\mathscr{P}_N E_{y_k}^{m+1}\right), \eta_{z_k}^{[m+1]}\right)_{T_k} - \frac{1}{\Delta t}\left(\int_{t^{m+1/2}}^{t^{m+3/2}} \frac{\partial E_{y_k}}{\partial x} dt, \eta_{z_k}^{[m+1]}\right)_{T_k},
\end{aligned} \tag{4.12}
$$

$$
\begin{aligned}
S_{3,k}^m = &\left(\frac{n_y}{Z^+ + Z^-}\left(Z^+[\mathscr{P}_N H_{z_k}^{m+1/2}] - \alpha\left(n_x[\mathscr{P}_N E_{y_k}^m] - n_y[\mathscr{P}_N E_{x_k}^m]\right)\right), \xi_{x_k}^{[m+1/2]}\right)_{\partial T_k} \\
&-\left(\frac{n_x}{Z^+ + Z^-}\left(Z^+[\mathscr{P}_N H_{z_k}^{m+1/2}] - \alpha\left(n_x[\mathscr{P}_N E_{y_k}^m] - n_y[\mathscr{P}_N E_{x_k}^m]\right)\right), \xi_{y_k}^{[m+1/2]}\right)_{\partial T_k} \\
&-\left(\frac{1}{Y^+ + Y^-}\left(Y^+(n_x[\mathscr{P}_N E_{y_k}^{m+1}] - n_y[\mathscr{P}_N E_{x_k}^{m+1}]) - \alpha[\mathscr{P}_N H_{z_k}^{m+1/2}]\right), \eta_{z_k}^{[m+1]}\right)_{\partial T_k},
\end{aligned} \tag{4.13}
$$

and

$$
\begin{aligned}
S_{4,k}^m = &\left(\frac{-n_y}{Z^+ + Z^-}\left(Z^+[\eta_{z_k}^{m+1/2}] - \alpha\left(n_x[\xi_{y_k}^m] - n_y[\xi_{x_k}^m]\right)\right), \xi_{x_k}^{[m+1/2]}\right)_{\partial T_k} \\
&+\left(\frac{n_x}{Z^+ + Z^-}\left(Z^+[\eta_{z_k}^{m+1/2}] - \alpha\left(n_x[\xi_{y_k}^m] - n_y[\xi_{x_k}^m]\right)\right), \xi_{y_k}^{[m+1/2]}\right)_{\partial T_k} \\
&+\left(\frac{1}{Y^+ + Y^-}\left(Y^+(n_x[\xi_{y_k}^{m+1}] - n_y[\xi_{x_k}^{m+1}]) - \alpha[\eta_{z_k}^{m+1/2}], \eta_{z_k}^{[m+1]}\right)\right)_{\partial T_k} \\
&+\left(n_y \xi_{x_k}^{m+1}, \eta_{z_k}^{[m+1]}\right)_{\partial T_k} - \left(n_x \xi_{y_k}^{m+1}, \eta_{z_k}^{[m+1]}\right)_{\partial T_k}.
\end{aligned} \tag{4.14}
$$

In what follows we will derive upper bounds for $S_{1,k}^m$, $S_{2,k}^m$, $S_{3,k}^m$ and $S_{4,k}^m$ in Lemma 4.1.1 – 4.1.4. We employed Cauchy-Schwarz's and Young's inequalities frequently in our derivation.

Lemma 4.1.1. *Let $S_{1,k}^m$ be defined by (4.11). Then*

$$
\begin{aligned}
\sum_{m=0}^{M-1} \sum_{T_k \in \mathscr{T}_h} S_{1,k}^m \le \frac{Ch^{2\sigma}}{N^{2p}} &\left(\frac{\overline{\varepsilon}}{\delta}\int_0^T \left\|\frac{\partial E}{\partial t}\right\|_{H^p(\Omega)}^2 dt + \frac{\overline{\mu}}{\delta}\int_{\Delta t/2}^{T+\Delta t/2} \left\|\frac{\partial H_z}{\partial t}\right\|_{H^p(\Omega)}^2 dt\right) \\
&+\frac{\delta}{2}\left(\|\xi^0\|_{L^2(\Omega)}^2 + 2\sum_{m=1}^{M-1}\|\xi^m\|_{L^2(\Omega)}^2 + \|\xi^M\|_{L^2(\Omega)}^2\right) \\
&+\frac{\delta}{2}\left(\|\eta_z^{1/2}\|_{L^2(\Omega)}^2 + 2\sum_{m=1}^{M-1}\|\eta_z^{m+1/2}\|_{L^2(\Omega)}^2 + \|\eta_z^{M+1/2}\|_{L^2(\Omega)}^2\right),
\end{aligned} \tag{4.15}
$$

where $p \ge 0$, $\sigma = \min(p, N+1)$, C is a constant independent of $(E_{x_k}, E_{y_k}, H_{z_k})$, h and N, and δ is an arbitrary positive constant.

Proof. We start by observing that

$$\|\rho_{x_k}^{m+1}-\rho_{x_k}^{m}\|_{L^2(T_k)}^2=\left\|\int_{t^m}^{t^{m+1}}\frac{\partial\rho_{x_k}}{\partial t}dt\right\|_{L^2(T_k)}^2.$$

Applying the Cauchy-Schwarz's inequality and the approximation property of Lemma A.2.1, we get

$$\left\|\int_{t^m}^{t^{m+1}}\frac{\partial\rho_{x_k}}{\partial t}dt\right\|_{L^2(T_k)}^2\leq\Delta t\int_{t^m}^{t^{m+1}}\left\|\frac{\partial\rho_{x_k}}{\partial t}\right\|_{L^2(T_k)}^2dt\leq\Delta tC\frac{h_k^{2\sigma}}{N^{2p}}\int_{t^m}^{t^{m+1}}\left\|\frac{\partial E_{x_k}}{\partial t}\right\|_{H^p(T_k)}^2dt. \tag{4.16}$$

Using Cauchy-Schwarz's inequality, the estimate (4.16) and the Young's inequality in the form

$$ab\leq\frac{\delta}{2}a^2+\frac{1}{2\delta}b^2,$$

where δ is an arbitrary positive constant, we obtain

$$\begin{aligned}S_{1,k}^m\leq&\frac{\overline{\varepsilon}C}{\delta}\frac{h_k^{2\sigma}}{N^{2p}}\int_{t^m}^{t^{m+1}}\left\|\frac{\partial E_{x_k}}{\partial t}\right\|_{H^p(T_k)}^2dt+\frac{\delta}{2}\|\xi_{x_k}^{[m+1/2]}\|_{L^2(T_k)}^2\\&+\frac{\overline{\varepsilon}C}{\delta}\frac{h_k^{2\sigma}}{N^{2p}}\int_{t^m}^{t^{m+1}}\left\|\frac{\partial E_{y_k}}{\partial t}\right\|_{H^p(T_k)}^2dt+\frac{\delta}{2}\|\xi_{y_k}^{[m+1/2]}\|_{L^2(T_k)}^2\\&+\frac{\overline{\mu}C}{\delta}\frac{h_k^{2\sigma}}{N^{2p}}\int_{t^{m+1/2}}^{t^{m+3/2}}\left\|\frac{\partial H_{z_k}}{\partial t}\right\|_{H^p(T_k)}^2dt+\frac{\delta}{2}\|\eta_{z_k}^{[m+1]}\|_{L^2(T_k)}^2.\end{aligned}$$

Summing from $m=0$ to $M-1$, we arrive at (4.15). □

Lemma 4.1.2. *Let $S_{2,k}^m$ be defined by (4.12). Then*

$$\begin{aligned}\sum_{m=0}^{M-1}\sum_{T_k\in\mathscr{T}_h}S_{2,k}^m\leq&\frac{CMh^{2\sigma-2}}{N^{2p-2}}\left(\frac{1}{\delta}\|E\|_{L^\infty(0,T;H^p(\Omega))}^2+\frac{1}{\delta}\|H_z\|_{L^\infty(0,T;H^p(\Omega))}^2\right)\\&+C\Delta t^3\left(\frac{1}{\delta}\int_{\Delta t/2}^{T+\Delta t/2}\left|\frac{\partial^2E}{\partial t^2}\right|_{H^1(\Omega)}^2dt+\frac{1}{\delta}\int_0^T\left|\frac{\partial^2H_z}{\partial t^2}\right|_{H^1(\Omega)}^2dt\right)\\&+\delta\left(\frac{1}{2}\|\xi^0\|_{L^2(\Omega)}^2+\sum_{m=1}^{M-1}\|\xi^m\|_{L^2(\Omega)}^2+\frac{1}{2}\|\xi^M\|_{L^2(\Omega)}^2\right)\\&+\delta\left(\|\eta_z^{1/2}\|_{L^2(\Omega)}^2+2\sum_{m=1}^{M-1}\|\eta_z^{m+1/2}\|_{L^2(\Omega)}^2+\|\eta_z^{M+1/2}\|_{L^2(\Omega)}^2\right),\end{aligned} \tag{4.17}$$

where $p\geq 0$, $\sigma=\min(p,N+1)$, C is a constant independent of $(E_{x_k},E_{y_k},H_{z_k})$, h and N, and δ is an arbitrary positive constant.

Proof. It is easy to check that

$$
\begin{aligned}
&-\left(\frac{\partial}{\partial y}\left(\mathscr{P}_N H_{z_k}^{m+1/2}\right), \xi_{x_k}^{[m+1/2]}\right)_{T_k} + \frac{1}{\Delta t}\left(\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y}(s)ds, \xi_{x_k}^{[m+1/2]}\right)_{T_k} \\
&\quad = -\left(\frac{\partial \phi_{z_k}^{m+1/2}}{\partial y}, \xi_{x_k}^{[m+1/2]}\right)_{T_k} - \left(\frac{\partial H_{z_k}^{m+1/2}}{\partial y} - \frac{1}{\Delta t}\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y}(s)ds, \xi_{x_k}^{[m+1/2]}\right)_{T_k}.
\end{aligned}
$$

From Lemma A.2.1 we obtain

$$
\left\| \frac{\partial \phi_{z_k}^{m+1/2}}{\partial y} \right\|_{L^2(T_k)} \leq C \frac{h_k^{\sigma-1}}{N^{p-1}} \|H_{z_k}^{m+1/2}\|_{H^p(T_k)}, \tag{4.18}
$$

where $\sigma = \min(p, N+1)$. Applying Cauchy-Schwarz's inequality and using (4.18), we arrive at

$$
\left(\frac{\partial \phi_{z_k}^{m+1/2}}{\partial y}, \xi_{x_k}^{[m+1/2]}\right)_{T_k} \leq \frac{C}{\delta}\frac{h_k^{2\sigma-2}}{N^{2p-2}} \|H_{z_k}^{m+1/2}\|_{H^p(T_k)}^2 + \frac{\delta}{2}\|\xi_{x_k}^{[m+1/2]}\|_{L^2(T_k)}^2.
$$

Employing Bramble-Hilbert Lemma [17] yields to

$$
\left\| \frac{\partial H_{z_k}^{m+1/2}}{\partial y} - \frac{1}{\Delta t}\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y} dt \right\|_{L^2(T_k)}^2 \leq C\Delta t^3 \int_{t^m}^{t^{m+1}} \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial H_{z_k}}{\partial y}\right) \right\|_{L^2(T_k)}^2 dt,
$$

and by Cauchy-Schwarz's and Young's inequalities follows

$$
\begin{aligned}
&\left(\frac{\partial H_{z_k}^{m+1/2}}{\partial y} - \frac{1}{\Delta t}\int_{t^m}^{t^{m+1}} \frac{\partial H_{z_k}}{\partial y} dt, \xi_{x_k}^{[m+1/2]}\right)_{T_k} \\
&\qquad \leq \frac{C\Delta t^3}{\delta}\int_{t^m}^{t^{m+1}} \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial H_{z_k}}{\partial y}\right) \right\|_{L^2(T_k)}^2 dt + \frac{\delta}{2}\|\xi_{x_k}^{[m+1/2]}\|_{L^2(T_k)}^2.
\end{aligned}
$$

The other terms in $S_{2,k}^m$ can be bounded in a similar way. Therefore,

$$
\begin{aligned}
S_{2,k}^m \;\leq\;& \frac{C}{\delta}\frac{h_k^{2\sigma-2}}{N^{2p-2}}\|E_{x_k}^{m+1}\|_{H^p(T_k)}^2 + \frac{C}{\delta}\frac{h_k^{2\sigma-2}}{N^{2p-2}}\|E_{y_k}^{m+1}\|_{H^p(T_k)}^2 + \frac{2C}{\delta}\frac{h_k^{2\sigma-2}}{N^{2p-2}}\|H_{z_k}^{m+1/2}\|_{H^p(T_k)}^2 \\
&+ \frac{C\Delta t^3}{\delta}\int_{t^{m+1/2}}^{t^{m+3/2}} \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial E_{x_k}}{\partial y}\right) \right\|_{L^2(T_k)}^2 + \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial E_{y_k}}{\partial x}\right) \right\|_{L^2(T_k)}^2 dt \\
&+ \frac{C\Delta t^3}{\delta}\int_{t^m}^{t^{m+1}} \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial H_{z_k}}{\partial y}\right) \right\|_{L^2(T_k)}^2 + \left\| \frac{\partial^2}{\partial t^2}\left(\frac{\partial H_{z_k}}{\partial x}\right) \right\|_{L^2(T_k)}^2 dt \\
&+ \delta\|\xi_{x_k}^{[m+1/2]}\|_{L^2(T_k)}^2 + \delta\|\xi_{y_k}^{[m+1/2]}\|_{L^2(T_k)}^2 + 2\delta\|\eta_{z_k}^{[m+1]}\|_{L^2(T_k)}^2.
\end{aligned}
$$

Summing from $m = 0$ to $M-1$ we arrive at (4.17). □

In order to find the estimate for the terms related the outer boundary $\partial\Omega$, it is necessary to apply the boundary conditions as defined in (4.1)–(4.3). We will observe the influence of the type of boundary conditions on $S_{3,k}^m$. The terms that arise from applying SM-ABC will change the temporal order in the final estimation in Theorem 4.1.5.

Lemma 4.1.3. *Let $S^m_{3,k}$ be defined by (4.13). Then*

$$\begin{aligned}
\sum_{m=0}^{M-1}\sum_{T_k\in\mathscr{T}_h} S^m_{3,k} &\leq \frac{CMh^{2\sigma-2}}{\delta N^{2p+1}} C_\tau^2 (N+1)(N+2)\left(1+\frac{\alpha}{\min\{Z_k^2\}}\right)\|E\|^2_{L^\infty(0,T;H^p(\Omega))} \\
&+ \frac{CMh^{2\sigma-2}}{\delta N^{2p+1}} C_\tau^2 (N+1)(N+2)\left(1+\frac{\alpha}{\min\{Y_k^2\}}\right)\|H_z\|^2_{L^\infty(0,T;H^p(\Omega))} \\
&+ \frac{\beta_4 C_\tau^2 (N+1)(N+2)\Delta t}{16\delta\min\{Z_k^2\}}\int_0^{T-\Delta t/2}\left\|\frac{\partial E}{\partial t}\right\|^2_{L^\infty(\partial\Omega)} dt \\
&+ \frac{\beta_4 C_\tau^2 (N+1)(N+2)\Delta t}{32\delta\min\{Y_k^2\}}\int_{\Delta t/2}^{T}\left\|\frac{\partial H_z}{\partial t}\right\|^2_{L^\infty(\partial\Omega)} dt \\
&+ \frac{\delta}{4}(1+\beta_4)\Big(\|\xi^0\|^2_{L^2(\Omega)} + 2\sum_{m=1}^{M-1}\|\xi^m\|^2_{L^2(\Omega)} + \|\xi^M\|^2_{L^2(\Omega)}\Big) \\
&+ \frac{\delta}{4}(1+\beta_4)\Big(\|\eta_z^{1/2}\|^2_{L^2(\Omega)} + 2\sum_{m=1}^{M-1}\|\eta_z^{m+1/2}\|^2_{L^2(\Omega)} + \|\eta_z^{M+1/2}\|^2_{L^2(\Omega)}\Big), \quad (4.19)
\end{aligned}$$

where $p\geq 0$, $\sigma=\min(p,N+1)$, C and C_τ are constants independent of $(E_{x_k},E_{y_k},H_{z_k})$, h and N, and δ is an arbitrary positive constant. Moreover, $\beta_4=0$ for PEC and PMC boundary conditions and $\beta_4=1$ for Silver-Müller absorbing boundary conditions.

Proof. In order to estimate $\sum_{T_k\in\mathscr{T}_h} S^m_{3,k}$, let us write $\|[\mathscr{P}_N E^m_{x_k}]\|_{f_{ik}}$, $f_{ik}\subset F^{int}$, as

$$\begin{aligned}
\|[\mathscr{P}_N E^m_{x_k}]\|_{L^2(f_{ik})} &= \|\mathscr{P}_N E^{m-}_{x_k} - E^m_{x_k} + E^m_{x_k} - \mathscr{P}_N E^{m+}_{x_k}\|_{L^2(f_{ik})} \\
&\leq \|\mathscr{P}_N E^{m-}_{x_k} - E^m_{x_k}\|_{L^2(f_{ik})} + \|E^m_{x_k} - \mathscr{P}_N E^{m+}_{x_k}\|_{L^2(f_{ik})}.
\end{aligned}$$

By Lemma A.2.1 we deduce that

$$\|\mathscr{P}_N E^{m-}_{x_k} - E^m_{x_k}\|^2_{L^2(f_{ik})} \leq C\frac{h_k^{2\sigma-1}}{N^{2p-1}}\|E^m_{x_k}\|^2_{H^p(T_k)},$$

where $\sigma=\min(p,N+1)$ and $p>\frac{1}{2}$. In the same way, we obtain

$$\|\mathscr{P}_N E^{m+}_{x_k} - E^m_{x_k}\|^2_{L^2(f_{ik})} \leq C\frac{h_k^{2\sigma-1}}{N^{2p-1}}\|E^m_{x_k}\|^2_{H^p(T_i)}.$$

Similar estimates hold for $\|[\mathscr{P}_N E^m_{y_k}]\|^2_{L^2(f_{ik})}$ and $\|[\mathscr{P}_N H^m_{z_k}]\|^2_{L^2(f_{ik})}$.

Let us now consider the edges that belong to the external boundary $\partial\Omega$. In the case of PEC boundary condition we have $[\mathscr{P}_N H^{m+1/2}_{z_k}]=0$. Since $n_x E^m_{y_k} - n_y E^m_{x_k}=0$, then

$$n_x[\mathscr{P}_N E^m_{y_k}] - n_y[\mathscr{P}_N E^m_{x_k}] = 2(n_x(\mathscr{P}_N E^m_{y_k} - E^m_{y_k}) - n_y(\mathscr{P}_N E^m_{x_k} - E^m_{x_k})).$$

For PMC boundary conditions we have $[\mathscr{P}_N E^m_{y_k}]=0$ and $[\mathscr{P}_N E^m_{x_k}]=0$. Since $H^{m+1/2}_{z_k}=0$, then

$$[\mathscr{P}_N H^{m+1/2}_{z_k}] = 2(\mathscr{P}_N H^{m+1/2}_{z_k} - H^{m+1/2}_{z_k}).$$

Applying Cauchy-Schwarz's inequality, (A.4) of Lemma A.1.1 and Young's inequality, for the cases of PEC or PMC boundary conditions, we obtain

$$\begin{aligned}\sum_{T_k\in\mathscr{T}_h} S_{3,k}^m \;\leq\; \sum_{T_k\in\mathscr{T}_h}\Bigg(&\frac{C}{\delta}C_\tau^2(N+1)(N+2)\left(1+\frac{\alpha}{\min\{Y_k^2\}}\right)\sum_{i\in\nu_k}\frac{h_i^{2\sigma-2}}{N^{2p+1}}\|H_{z_i}^{m+1/2}\|^2_{H^p(T_i)}\\ &+\frac{C}{\delta}C_\tau^2(N+1)(N+2)\left(1+\frac{\alpha}{\min\{Z_k^2\}}\right)\sum_{i\in\nu_k}\frac{h_i^{2\sigma-2}}{N^{2p+1}}\|E_i^m\|^2_{H^p(T_i)}\\ &+\frac{\delta}{2}\|\xi_{x_k}^{[m+1/2]}\|^2_{L^2(T_k)}+\frac{\delta}{2}\|\xi_{y_k}^{[m+1/2]}\|^2_{L^2(T_k)}+\frac{\delta}{2}\|\eta_{z_k}^{[m+1]}\|^2_{L^2(T_k)}\Bigg).\end{aligned}\tag{4.20}$$

In the case of Silver-Müller absorbing boundary condition, on the edges that belong to the external boundary, we observe that

$$Z^+H_{z_k}^{m+1/2}-(n_xE_{y_k}^{m+1/2}-n_yE_{x_k}^{m+1/2})=0,\quad Y^+(n_xE_{y_k}^{m+1}-n_yE_{x_k}^{m+1})-H_{z_k}^{m+1}=0.$$

Thus,

$$\begin{aligned}Z^+[\mathscr{P}_NH_{z_k}^{m+1/2}]-(n_x[\mathscr{P}_NE_{y_k}^m]-n_y[\mathscr{P}_NE_{x_k}^m]) &= Z^+\mathscr{P}_NH_{z_k}^{m+1/2}-(n_x\mathscr{P}_NE_{y_k}^m-n_y\mathscr{P}_NE_{x_k}^m)\\ &= Z^+(\mathscr{P}_NH_{z_k}^{m+1/2}-H_{z_k}^{m+1/2})-(n_x(\mathscr{P}_NE_{y_k}^m-E_{y_k}^m)-n_y(\mathscr{P}_NE_{x_k}^m-E_{x_k}^m))\\ &\quad +n_x(E_{y_k}^{m+1/2}-E_{y_k}^m)-n_y(E_{x_k}^{m+1/2}-E_{x_k}^m),\end{aligned}$$

and

$$\begin{aligned}Y^+(n_x[\mathscr{P}_NE_{y_k}^{m+1}]-n_y[\mathscr{P}_NE_{x_k}^{m+1}])-[\mathscr{P}_NH_{z_k}^{m+1/2}] &= Y^+(n_x\mathscr{P}_NE_{y_k}^{m+1}-n_y\mathscr{P}_NE_{x_k}^{m+1})-\mathscr{P}_NH_{z_k}^{m+1/2}\\ &= Y^+(n_x(\mathscr{P}_NE_{y_k}^{m+1}-E_{y_k}^{m+1})-n_y(\mathscr{P}_NE_{x_k}^{m+1}-E_{x_k}^{m+1})\\ &\quad -(\mathscr{P}_NH_{z_k}^{m+1/2}-H_{z_k}^{m+1/2})+H_{z_k}^{m+1}-H_{z_k}^{m+1/2}.\end{aligned}$$

For $f_k\in F^{ext}$, we obtain

$$\left\|\frac{n_y}{Z^++Z^-}(n_x(E_{y_k}^{m+1/2}-E_{y_k}^m)-n_y(E_{x_k}^{m+1/2}-E_{x_k}^m))\right\|_{L^\infty(f_k)}\leq\frac{1}{2\min\{Z_k\}}\int_{t^m}^{t^{m+1/2}}\left\|\frac{\partial E}{\partial t}\right\|_{L^\infty(f_k)}dt.$$

In the same way we get the estimate

$$\left\|\frac{1}{Y^++Y^-}(H_{z_k}^{m+1}-H_{z_k}^{m+1/2})\right\|_{L^\infty(f_k)}\leq\frac{1}{2\min\{Y_k\}}\int_{t^{m+1/2}}^{t^{m+1}}\left\|\frac{\partial H_z}{\partial t}\right\|_{L^\infty(f_k)}dt.$$

Therefore, to estimate $S_{3,k}^m$ we use (A.4) of Lemma A.1.1, and we observe that we need to add the terms

$$\frac{1}{2\delta}\frac{C_\tau^2(N+1)(N+2)\Delta t}{8\min\{Z_k^2\}}\int_{t^m}^{t^{m+1/2}}\left\|\frac{\partial E}{\partial t}\right\|^2_{L^\infty(f_k)}dt,\quad\frac{\delta}{2}\|\xi_k^{[m+1/2]}\|^2_{L^2(T_k)},$$

$$\frac{1}{2\delta}\frac{C_\tau^2(N+1)(N+2)\Delta t}{16\min\{Y_k^2\}}\int_{t^{m+1/2}}^{t^{m+1}}\left\|\frac{\partial H_z}{\partial t}\right\|^2_{L^\infty(f_k)}dt,\quad\frac{\delta}{2}\|\eta_{z_k}^{[m+1]}\|^2_{L^2(T_k)},$$

to the right hand side of (4.20).

Summing from $m = 0$ to $M-1$, leads to the estimation (4.19). □

The boundary terms in (4.10) given by $S^m_{4,k}$ are similar to $\sum_{T_k\in\mathscr{T}_h} A^m_k$ (Equation (3.4)) that was estimated in Section 3.1 (by Lemma 3.1.1 and Lemma 3.1.2).

Lemma 4.1.4. *Let $S^m_{4,k}$ be defined by (4.14). Then*

$$\begin{aligned}\sum_{m=0}^{M-1}\sum_{T_k\in\mathscr{T}_h} S^m_{4,k} \le{}& \frac{1}{2}C_\tau^2(N+1)(N+2)\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+\beta_1}{2\min\{Z_k\}}\right)\|\xi^M\|^2_{L^2(\Omega)}\\ &+\frac{1}{2}C_\tau^2(N+1)(N+2)\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+2\beta_2\beta_3}{2\min\{Y_k\}}\right)\|\eta_z^{M+1/2}\|^2_{L^2(\Omega)}\\ &+\frac{1}{2}C_\tau^2(N+1)(N+2)\max\left\{h_k^{-1}\right\}(2+\beta_2)\left(\|\xi^0\|^2_{L^2(\Omega)}+\|\eta_z^{1/2}\|^2_{L^2(\Omega)}\right). \end{aligned}\tag{4.21}$$

Proof. Following the arguments we used to estimate $\sum_{T_k\in\mathscr{T}_h} A^m_k$ in Section 3.1, we can find an upper bound for $\sum_{T_k\in\mathscr{T}_h} S^m_{4,k}$ as

$$\begin{aligned}\sum_{m=0}^{M-1}\sum_{T_k\in\mathscr{T}_h} S^m_{4,k} \le{}& \frac{\alpha}{8\min\{Z_k\}}\sum_{f_{ik}\in F^{int}}\|[\xi_k^M]\|^2_{L^2(f_{ik})}+\frac{\alpha}{8\min\{Y_k\}}\sum_{f_{ik}\in F^{int}}\|[\eta_{z_k}^{M+1/2}]\|^2_{L^2(f_{ik})}\\ &+\sum_{f_{ik}\in F^{int}}\left(\|\xi_k^M\|_{L^2(f_{ik})}\|[\eta_{z_k}^{M+1/2}]\|_{L^2(f_{ik})}+\|\xi_k^0\|_{L^2(f_{ik})}\|[\eta_{z_k}^{1/2}]\|_{L^2(f_{ik})}\right)\\ &+\frac{\beta_1}{4\min\{Z_k\}}\sum_{f_k\in F^{ext}}\|\xi_k^M\|^2_{L^2(f_k)}+\frac{\beta_2\beta_3}{2\min\{Y_k\}}\sum_{f_k\in F^{ext}}\|\eta_{z_k}^{M+1/2}\|^2_{L^2(f_k)}\\ &+\beta_2\sum_{f_k\in F^{ext}}\left(\|\eta_{z_k}^{1/2}\|_{L^2(f_k)}\|\xi_k^0\|_{L^2(f_k)}+\|\eta_{z_k}^{M+1/2}\|_{L^2(f_k)}\|\xi_k^M\|_{L^2(f_k)}\right), \end{aligned}$$

where $\beta_1=\alpha,\beta_2=0$ for PEC, $\beta_1=0,\beta_2=1$, $\beta_3=\alpha$ for PMC, and $\beta_1=\beta_2=\frac{1}{2}$, $\beta_3=1$ for Silver-Müller boundary conditions. As in the proof of Theorem 3.1.3, using the inequality (A.4) of Lemma A.1.1 and the inequality (A.6) of Lemma A.1.2 (both in Appendix A) we obtain the estimate (4.21). □

The upper bounds we have already found in Lemma 4.1.1–4.1.4 give an estimate for R^m_k in (4.10). According to (4.10) it is enough to employ inverse and trace inequalities and polynomial approximations which are all stated in Appendix A, to find error of polynomial and numerical approximation.

Theorem 4.1.5. *Let us consider the leap-frog DG method (2.32)–(2.34) complemented with the discrete boundary conditions defined in Subsecection 2.2.4 and suppose that the solution of the Maxwell's equations (2.9)–(2.10) complemented by PEC, PMC or Silver-Müller absorbing boundary conditions has the following regularity:*
$E_x,E_y,H_z\in L^\infty(0,T_f;H^{s+1}(\Omega))$, $\frac{\partial E_x}{\partial t},\frac{\partial E_y}{\partial t},\frac{\partial H_z}{\partial t}\in L^2(0,T_f;H^{s+1}(\Omega)\cap L^\infty(\partial\Omega))$ and $\frac{\partial^2 E_x}{\partial t^2},\frac{\partial^2 E_y}{\partial t^2},\frac{\partial^2 H_z}{\partial t^2}\in L^2(0,T_f;H^1(\Omega))$, $s\ge 0$. If the time step Δt satisfies

$$\Delta t\le\frac{\min\{\underline{\varepsilon},\underline{\mu}\}}{\max\{C_E,C_H\}}\min\{h_k\}(1-\delta),\quad 0<\delta<1,\tag{4.22}$$

where C_E and C_H are the constants defined in Theorem 3.1.3, then, for the case of PEC and PMC boundary conditions, holds

$$\max_{1\leq m\leq M}\left(\|E^m-\tilde{E}^m\|_{L^2(\Omega)}+\|H_z^{m+1/2}-\tilde{H}_z^{m+1/2}\|_{L^2(\Omega)}\right)\leq C(\Delta t^2+h^{\min\{s,N\}})$$
$$+C\left(\|E^0-\tilde{E}^0\|_{L^2(\Omega)}+\|H_z^{1/2}-\tilde{H}_z^{1/2}\|_{L^2(\Omega)}\right),$$

and, for the case of Silver-Müller absorbing boundary conditions, holds

$$\max_{1\leq m\leq M}\left(\|E^m-\tilde{E}^m\|_{L^2(\Omega)}+\|H_z^{m+1/2}-\tilde{H}_z^{m+1/2}\|_{L^2(\Omega)}\right)\leq C(\Delta t+h^{\min\{s,N\}})$$
$$+C\left(\|E^0-\tilde{E}^0\|_{L^2(\Omega)}+\|H_z^{1/2}-\tilde{H}_z^{1/2}\|_{L^2(\Omega)}\right),$$

where C is a generic constant independent of Δt and the mesh size h.

Proof. From (4.10) and taking into account Lemma A.1.2 and the estimates from previous lemmata, we obtain

$$\begin{aligned}
\min\{\underline{\varepsilon},\underline{\mu}\}&\left(\|\xi^M\|^2_{L^2(\Omega)}+\|\eta_z^{M+1/2}\|^2_{L^2(\Omega)}\right)\leq \max\{\bar{\varepsilon},\bar{\mu}\}\left(\|\xi^0\|^2_{L^2(\Omega)}+\|\eta_z^{1/2}\|^2_{L^2(\Omega)}\right)\\
&+\frac{\Delta t}{2}C_{inv}N^2\max\left\{h_k^{-1}\right\}\left(\|\eta_z^{1/2}\|^2_{L^2(\Omega)}+\|\xi^0\|^2_{L^2(\Omega)}+\|\eta_z^{M+1/2}\|^2_{L^2(\Omega)}+\|\xi^M\|^2_{L^2(\Omega)}\right)\\
&+\Delta t\,\delta(7+\beta_4)\left(\sum_{m=1}^{M-1}\|\xi^m\|^2_{L^2(\Omega)}+\sum_{m=1}^{M-1}\|\eta_z^{m+1/2}\|^2_{L^2(\Omega)}\right)\\
&+\frac{\Delta t\,\delta}{2}(7+\beta_4)\left(\|\xi^0\|^2_{L^2(\Omega)}+\|\xi^M\|^2_{L^2(\Omega)}+\|\eta_{z_k}^{1/2}\|^2_{L^2(\Omega)}+\|\eta_{z_k}^{M+1/2}\|^2_{L^2(\Omega)}\right)\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+\beta_1}{2\min\{Z_k\}}\right)\|\xi^M\|^2_{L^2(\Omega)}\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(2+\beta_2+\frac{\alpha+2\beta_2\beta_3}{2\min\{Y_k\}}\right)\|\eta_z^{M+1/2}\|^2_{L^2(\Omega)}\\
&+C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}(2+\beta_2)\left(\|\xi^0\|^2_{L^2(\Omega)}+\|\eta_z^{1/2}\|^2_{L^2(\Omega)}\right)\\
&+\frac{Ch^{2\sigma}\Delta t}{\delta N^{2p}}\left(\overline{\varepsilon}\int_0^T\left\|\frac{\partial E}{\partial t}\right\|^2_{H^p(\Omega)}dt+\overline{\mu}\int_{\Delta t/2}^{T+\Delta t/2}\left\|\frac{\partial H_z}{\partial t}\right\|^2_{H^p(\Omega)}dt\right)\\
&+\frac{2C\Delta t^4}{\delta}\left(\int_{\Delta t/2}^{T+\Delta t/2}\left|\frac{\partial^2 E}{\partial t^2}\right|^2_{H^1(\Omega)}dt+\int_0^T\left|\frac{\partial^2 H_z}{\partial t^2}\right|^2_{H^1(\Omega)}dt\right)\\
&+\frac{2CTh^{2\sigma-2}}{\delta N^{2p+1}}\left(N^3+\left(1+\frac{\alpha}{\min\{Z_k^2\}}\right)C_\tau^2(N+1)(N+2)\right)\|E\|^2_{L^\infty(0,T;H^p(\Omega))}\\
&+\frac{2CTh^{2\sigma-2}}{\delta N^{2p+1}}\left(N^3+\left(1+\frac{\alpha}{\min\{Y_k^2\}}\right)C_\tau^2(N+1)(N+2)\right)\|H_z\|^2_{L^\infty(0,T;H^p(\Omega))}\\
&+\frac{\beta_4C_\tau^2(N+1)(N+2)\Delta t^2}{8\delta\min\{Z_k^2\}}\int_0^{T-\Delta t/2}\left\|\frac{\partial E}{\partial t}\right\|^2_{L^\infty(\partial\Omega)}dt\\
&+\frac{\beta_4C_\tau^2(N+1)(N+2)\Delta t^2}{16\delta\min\{Y_k^2\}}\int_{\Delta t/2}^{T}\left\|\frac{\partial H_z}{\partial t}\right\|^2_{L^\infty(\partial\Omega)}dt,
\end{aligned}$$

where $p \geq 0$, $\sigma = \min(p, N+1)$, C and C_τ are constants independent of $(E_{x_k}, E_{y_k}, H_{z_k})$, h_k and N, δ is an arbitrary positive constant, $\beta_4 = 0$ for *PEC* and *PMC* boundary conditions and $\beta_4 = 1$ for Silver-Müller absorbing boundary conditions.

If (4.22) holds, using the discrete Gronwall's Lemma (Lemma A.3.1 in Appendix A) we obtain

$$\begin{aligned}
\|\xi^M\|^2_{L^2(\Omega)} + \|\eta_z^{M+1/2}\|^2_{L^2(\Omega)} \leq C(\varepsilon,\mu,N)\Big(& \|\xi^0\|^2_{L^2(\Omega)} + \|\eta_z^{1/2}\|^2_{L^2(\Omega)} \\
& + \Delta t h^{2\sigma} \int_0^T \left\| \frac{\partial E}{\partial t} \right\|^2_{H^p(\Omega)} dt + \Delta t h^{2\sigma} \int_{\Delta t/2}^{T+\Delta t/2} \left\| \frac{\partial H_z}{\partial t} \right\|^2_{H^p(\Omega)} dt \\
& + \Delta t^4 \int_0^T \left| \frac{\partial^2 E}{\partial t^2} \right|^2_{H^1(\Omega)} dt + \Delta t^4 \int_{\Delta t/2}^{T+\Delta t/2} \left| \frac{\partial^2 H_z}{\partial t^2} \right|^2_{H^1(\Omega)} dt \\
& + h^{2\sigma-2} \|E^m\|^2_{L^\infty(0,T;H^p(\Omega))} + h^{2\sigma-2} \|H_z^{m+1/2}\|^2_{L^\infty(0,T;H^p(\Omega))} \\
& + \beta_4 \Delta t^2 \int_0^{T-\Delta t/2} \left\| \frac{\partial E}{\partial t} \right\|^2_{L^\infty(\partial\Omega)} dt + \beta_4 \Delta t^2 \int_{\Delta t/2}^{T} \left\| \frac{\partial H_z}{\partial t} \right\|^2_{L^\infty(\partial\Omega)} dt \Big).
\end{aligned}$$

We complete the proof by using the triangle inequality and the *hp* approximation properties of Lemma A.2.1 to estimate $\|\mathscr{P}_N E^M_{x_k} - E^M_{x_k}\|_{L^2(\Omega)}$, $\|\mathscr{P}_N E^M_{y_k} - E^M_{y_k}\|_{L^2(\Omega)}$ and $\|\mathscr{P}_N H^{M+1/2}_{z_k} - H^{M+1/2}_{z_k}\|_{L^2(\Omega)}$.

□

Remark 4.1.6. We want to remark that in the case of Silver-Müller absorbing boundary conditions we only get first order convergence in time. A possible way to recover second order convergence is to consider a locally implicit time scheme (see *e.g.* [32]). In order to keep efficiency, we propose an alternative which is explicit and second order convergent in time:
For each time step solve (2.32)–(2.34) and save the solution in the variables $(\tilde{\tilde{E}}^{m+1}_{x_k}, \tilde{\tilde{E}}^{m+1}_{y_k}, \tilde{\tilde{H}}^{m+3/2}_{z_k})$. Then the numerical solution $(\tilde{E}^{m+1}_{x_k}, \tilde{E}^{m+1}_{y_k}, \tilde{H}^{m+3/2}_{z_k})$ is computed replacing in (2.32)–(2.34) the numerical flux by the following expression

$$\begin{pmatrix}
\frac{-n_y}{Z^+ + Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}] - \alpha\left(n_x \frac{[\tilde{E}_y^m]+[\tilde{\tilde{E}}_y^{m+1}]}{2} - n_y \frac{[\tilde{E}_x^m]+[\tilde{\tilde{E}}_x^{m+1}]}{2} \right)\right) \\
\frac{n_x}{Z^+ + Z^-}\left(Z^+[\tilde{H}_z^{m+1/2}] - \alpha\left(n_x \frac{[\tilde{E}_y^m]+[\tilde{\tilde{E}}_y^{m+1}]}{2} - n_y \frac{[\tilde{E}_x^m]+[\tilde{\tilde{E}}_x^{m+1}]}{2} \right)\right) \\
\frac{1}{Y^+ + Y^-}\left(Y^+\left(n_x[\tilde{E}_y^{m+1}] - n_y[\tilde{E}_x^{m+1}]\right) - \alpha \frac{[\tilde{H}_z^{m+1/2}]+[\tilde{\tilde{H}}_z^{m+3/2}]}{2} \right)
\end{pmatrix}.$$

We will determine the order of temporal convergence with this modification in time integrator later in next chapter.

4.2 Implicit and explicit iterative time integrators

The temporal order of convergence of the scheme (2.32)–(2.34) is reduced from second order to first order due to the terms that arise from Silver-Müller absorbing boundary conditions. In Remark 4.1.6 we propose a predictor-corrector step as time integrator in order to recover second order convergence in time. In this section we study an iterative scheme and its relation with a second order convergent

implicit method. The scheme in Remark 4.1.6 can be viewed as a particular case of the iterative technique, where only the first iteration is computed.

The iterative method is defined in Subsection 4.2.2 and we will show, in Theorem 4.2.4, that its solution converges to the solution of the implicit scheme defined in Subsection 4.2.1. In this way, we can compute the numerical solution with a fully explicit iterative scheme expecting similar convergence properties to the implicit method.

4.2.1 Implicit method

Instead of backward approximations used in the explicit leap-frog DG scheme (2.32)–(2.34), in order to define an implicit leap-frog DG method, we use average approximations $\bar{E}^{[m+1/2]}$ for $\bar{E}^{m+1/2}$ and $\bar{H}^{[m+1]}$ for $\bar{H}^{m+1}$ where

$$\bar{E}^{[m+1/2]} = \frac{\bar{E}^m + \bar{E}^{m+1}}{2}, \quad \bar{H}^{[m+1]} = \frac{\bar{H}^{m+1/2} + \bar{H}^{m+3/2}}{2}. \tag{4.23}$$

The scheme defined in the following way: given an initial approximation $(\bar{E}^0_{x_k}, \bar{E}^0_{y_k}, \bar{H}^{1/2}_{z_k}) \in V_N$, for each $m = 0, 1, \ldots, M-1$, we compute $(\bar{E}^{m+1}_{x_k}, \bar{E}^{m+1}_{y_k}, \bar{H}^{m+1/2}_{z_k}) \in V_N$ such that, $\forall (u_k, v_k, w_k) \in V_N$,

$$\begin{aligned}
&\left(\varepsilon_{xx}\frac{\bar{E}^{m+1}_{x_k} - \bar{E}^m_{x_k}}{\Delta t} + \varepsilon_{xy}\frac{\bar{E}^{m+1}_{y_k} - \bar{E}^m_{y_k}}{\Delta t}, u_k\right)_{T_k} = \left(\partial_y \bar{H}^{m+1/2}_{z_k}, u_k\right)_{T_k} \\
&\quad + \left(\frac{-n_y}{Z^+ + Z^-}\left(Z^+[\bar{H}^{m+1/2}_z] - \alpha\left(n_x[\bar{E}^{[m+1/2]}_y] - n_y[\bar{E}^{[m+1/2]}_x]\right)\right), u_k\right)_{\partial T_k}, \qquad (4.24) \\
&\left(\varepsilon_{yx}\frac{\bar{E}^{m+1}_{x_k} - \bar{E}^m_{x_k}}{\Delta t} + \varepsilon_{yy}\frac{\bar{E}^{m+1}_{y_k} - \bar{E}^m_{y_k}}{\Delta t}, v_k\right)_{T_k} = -\left(\partial_x \bar{H}^{m+1/2}_{z_k}, v_k\right)_{T_k} \\
&\quad + \left(\frac{n_x}{Z^+ + Z^-}\left(Z^+[\bar{H}^{m+1/2}_z] - \alpha\left(n_x[\bar{E}^{[m+1/2]}_y] - n_y[\bar{E}^{[m+1/2]}_x]\right)\right), v_k\right)_{\partial T_k}, \qquad (4.25) \\
&\left(\mu\frac{\bar{H}^{m+3/2}_{z_k} - \bar{H}^{m+1/2}_{z_k}}{\Delta t}, w_k\right)_{T_k} = \left(\partial_y \bar{E}^{m+1}_{x_k} - \partial_x \bar{E}^{m+1}_{y_k}, w_k\right)_{T_k} \\
&\quad + \left(\frac{1}{Y^+ + Y^-}\left(Y^+(n_x[\bar{E}^{m+1}_y] - n_y[\bar{E}^{m+1}_x]) - \alpha[\bar{H}^{[m+1]}_z]\right), w_k\right)_{\partial T_k}. \qquad (4.26)
\end{aligned}$$

The boundary conditions are considered as for the explicit method but using (4.23). We note that equations (4.24)–(4.26) are defined implicitly, since the upwind fluxes involve the unknowns $\bar{E}^{m+1}_{x_k}$, $\bar{E}^{m+1}_{y_k}$ and $\bar{H}^{m+3/2}_{z_k}$. In the following, we will provide a sufficient condition for the L^2-stability of the implicit leap-frog DG method (4.24)–(4.26) for the case of SM-ABC.

Choosing $u_k = \Delta t \bar{E}^{[m+1/2]}_{x_k}$, $v_k = \Delta t \bar{E}^{[m+1/2]}_{y_k}$ and $w_k = \Delta t \bar{H}^{[m+1]}_{z_k}$, we have

$$\begin{aligned}
&\left(\varepsilon\bar{E}_k^{m+1},\bar{E}_k^{m+1}\right)_{T_k}-\left(\varepsilon\bar{E}_k^{m},\bar{E}_k^{m}\right)_{T_k}=2\Delta t\left(\nabla\times\bar{H}_{z_k}^{m+1/2},\bar{E}_k^{[m+1/2]}\right)_{T_k}\\
&+2\Delta t\left(\frac{-n_y}{Z^++Z^-}\left(Z^+[\bar{H}_z^{m+1/2}]-\alpha\left(n_x[\bar{E}_y^{[m+1/2]}]-n_y[\bar{E}_x^{[m+1/2]}]\right)\right),\bar{E}_{x_k}^{[m+1/2]}\right)_{\partial T_k}\\
&+2\Delta t\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\bar{H}_z^{m+1/2}]-\alpha\left(n_x[\bar{E}_y^{[m+1/2]}]-n_y[\bar{E}_x^{[m+1/2]}]\right)\right),\bar{E}_{y_k}^{[m+1/2]}\right)_{\partial T_k},
\end{aligned}\tag{4.27}$$

$$\begin{aligned}
&\left(\mu\bar{H}_{z_k}^{m+3/2},\bar{H}_{z_k}^{m+3/2}\right)_{T_k}-\left(\mu\bar{H}_{z_k}^{m+1/2},\bar{H}_{z_k}^{m+1/2}\right)_{T_k}=-2\Delta t\left(\text{curl }\bar{E}_k^{m+1},\bar{H}_{z_k}^{[m+1]}\right)_{T_k}\\
&+2\Delta t\left(\frac{1}{Y^++Y^-}\left(Y^+\left(n_x[\bar{E}_y^{m+1}]-n_y[\bar{E}_x^{m+1}]\right)-\alpha[\bar{H}_z^{[m+1]}]\right),\bar{H}_{z_k}^{[m+1]}\right)_{\partial T_k}.
\end{aligned}\tag{4.28}$$

Using the identity,

$$\left(\text{curl }\bar{E}_k^{m+1},\bar{H}_{z_k}^{[m+1]}\right)_{T_k}=\left(\nabla\times\bar{H}_{z_k}^{[m+1]},\bar{E}_k^{m+1}\right)_{T_k}+\left(n_x\bar{E}_{y_k}^{m+1}-n_y\bar{E}_{x_k}^{m+1},\bar{H}_{z_k}^{[m+1]}\right)_{\partial T_k},$$

summing (4.27) and (4.28) from $m=0$ to $m=M-1$, and integrating by parts, we get

$$\begin{aligned}
&\left(\varepsilon\bar{E}_k^{M},\bar{E}_k^{M}\right)_{T_k}+\left(\mu\bar{H}_{z_k}^{M+1/2},\bar{H}_{z_k}^{M+1/2}\right)_{T_k}=\left(\varepsilon\bar{E}_k^{0},\bar{E}_k^{0}\right)_{T_k}+\left(\mu\bar{H}_{z_k}^{1/2},\bar{H}_{z_k}^{1/2}\right)_{T_k}\\
&+\Delta t\left(\nabla\times\bar{H}_{z_k}^{1/2},\bar{E}_k^{0}\right)_{T_k}-\Delta t\left(\nabla\times\bar{H}_{z_k}^{M+1/2},\bar{E}_k^{M}\right)_{T_k}+2\Delta t\sum_{m=0}^{M-1}A_k^m,
\end{aligned}\tag{4.29}$$

where

$$\begin{aligned}
A_k^m \;=\;&\left(\frac{-n_y}{Z^++Z^-}\left(Z^+[\bar{H}_z^{m+1/2}]-\alpha\left(n_x[\bar{E}_y^{[m+1/2]}]-n_y[\bar{E}_x^{[m+1/2]}]\right)\right),\bar{E}_{x_k}^{[m+1/2]}\right)_{\partial T_k}\\
&+\left(\frac{n_x}{Z^++Z^-}\left(Z^+[\bar{H}_z^{m+1/2}]-\alpha\left(n_x[\bar{E}_y^{[m+1/2]}]-n_y[\bar{E}_x^{[m+1/2]}]\right)\right),\bar{E}_{y_k}^{[m+1/2]}\right)_{\partial T_k}\\
&+\left(\frac{1}{Y^++Y^-}\left(Y^+\left(n_x[\bar{E}_y^{m+1}]-n_y[\bar{E}_x^{m+1}]\right)-\alpha[\bar{H}_z^{[m+1]}]\right),\bar{H}_{z_k}^{[m+1]}\right)_{\partial T_k}\\
&-\left(n_x\bar{E}_{y_k}^{m+1}-n_y\bar{E}_{x_k}^{m+1},\bar{H}_{z_k}^{[m+1]}\right)_{\partial T_k}.
\end{aligned}$$

Summing over all elements $T_k\in\mathscr{T}_h$ we obtain

$$\sum_{T_k\in\mathscr{T}_h}A_k^m=B_1^m+B_2^m,$$

where B_1 contains the terms related to internal edges and could be written as

$$B_1^m=B_{11}^m+B_{12}^m+B_{13}^m,$$

where

$$B_{11}^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(-\frac{(n_y)_{ki}}{Z_i+Z_k}\left(Z_i[\bar{H}_{z_k}^{m+1/2}] - \alpha\left((n_x)_{ki}[\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki}[\bar{E}_{x_k}^{[m+1/2]}]\right)\right)\bar{E}_{x_k}^{[m+1/2]}$$
$$-\frac{(n_y)_{ik}}{Z_i+Z_k}\left(Z_k[\bar{H}_{zi}^{m+1/2}] - \alpha\left((n_x)_{ik}[\bar{E}_{yi}^{[m+1/2]}] - (n_y)_{ik}[\bar{E}_{xi}^{[m+1/2]}]\right)\right)\bar{E}_{xi}^{[m+1/2]}$$
$$-\frac{Y_i(n_y)_{ki}}{Y_i+Y_k}[\bar{E}_{x_k}^{m+1}]\bar{H}_{z_k}^{[m+1]} - \frac{Y_k(n_y)_{ik}}{Y_i+Y_k}[\bar{E}_{xi}^{m+1}]\bar{H}_{zi}^{[m+1]}$$
$$+(n_y)_{ki}\bar{E}_{x_k}^{m+1}\bar{H}_{z_k}^{[m+1]} + (n_y)_{ik}\bar{E}_{xi}^{m+1}\bar{H}_{zi}^{[m+1]}\Bigg)\,ds, \tag{4.30}$$

$$B_{12}^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(\frac{(n_x)_{ki}}{Z_i+Z_k}\left(Z_i[\bar{H}_{z_k}^{m+1/2}] - \alpha\left((n_x)_{ki}[\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki}[\bar{E}_{x_k}^{[m+1/2]}]\right)\right)\bar{E}_{y_k}^{[m+1/2]}$$
$$+\frac{(n_x)_{ik}}{Z_i+Z_k}\left(Z_k[\bar{H}_{zi}^{m+1/2}] - \alpha\left((n_x)_{ik}[\bar{E}_{yi}^{[m+1/2]}] - (n_y)_{ik}[\bar{E}_{xi}^{[m+1/2]}]\right)\right)\bar{E}_{yi}^{[m+1/2]}$$
$$+\frac{Y_i(n_x)_{ki}}{Y_i+Y_k}[\bar{E}_{y_k}^{m+1}]\bar{H}_{z_k}^{[m+1]} + \frac{Y_k(n_x)_{ik}}{Y_i+Y_k}[\bar{E}_{yi}^{m+1}]\bar{H}_{zi}^{[m+1]}$$
$$-(n_x)_{ki}\bar{E}_{y_k}^{m+1}\bar{H}_{z_k}^{[m+1]} - (n_x)_{ik}\bar{E}_{yi}^{m+1}\bar{H}_{zi}^{[m+1]}\Bigg)\,ds, \tag{4.31}$$

and

$$B_{13}^m = -\sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \left(\frac{\alpha}{Y_i+Y_k}[\bar{H}_{z_k}^{[m+1]}]\bar{H}_{z_k}^{[m+1]} + \frac{\alpha}{Y_i+Y_k}[\bar{H}_{zi}^{[m+1]}]\bar{H}_{zi}^{[m+1]}\right) ds, \tag{4.32}$$

and B_2^m are the terms of the outer boundary

$$B_2^m = \sum_{f_k\in F^{ext}} \int_{f_k} \Bigg(\frac{-(n_y)_k}{2Z_k}\left(Z_k[\bar{H}_{z_k}^{m+1/2}] - \alpha\left((n_x)_k[\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_k[\bar{E}_{x_k}^{[m+1/2]}]\right)\right)\bar{E}_{x_k}^{[m+1/2]}$$
$$+\frac{(n_x)_k}{2Z_k}\left(Z_k[\bar{H}_{z_k}^{m+1/2}] - \alpha\left((n_x)_k[\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_k[\bar{E}_{x_k}^{[m+1/2]}]\right)\right)\bar{E}_{y_k}^{[m+1/2]}$$
$$+\frac{1}{2Y_k}\left(Y_k\left((n_x)_k[\bar{E}_{y_k}^{m+1}] - (n_y)_k[\bar{E}_{x_k}^{m+1}]\right) - \alpha[\bar{H}_{z_k}^{[m+1]}]\right)\bar{H}_{z_k}^{[m+1]}$$
$$-\left((n_x)_k\bar{E}_{y_k}^{m+1} - (n_y)_k\bar{E}_{x_k}^{m+1}\right)\bar{H}_{z_k}^{[m+1]}\Bigg)\,ds. \tag{4.33}$$

We find an upper bound for B_1^m in Lemma 4.2.1 and for B_2^m in Lemma 4.2.2.

Lemma 4.2.1. *Let B_{11}^m, B_{12}^m and B_{13}^m be defined by (4.30), (4.31) and (4.32), respectively and $B_1^m = B_{11}^m + B_{12}^m + B_{13}^m$. Then*

$$\sum_{m=0}^{M-1} B_1^m \leq \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{1}{2(Z_i+Z_k)} \Bigg(- (n_y)_{ki} \left(Z_i \bar{E}_{x_k}^0 + Z_k \bar{E}_{xi}^0\right) [\bar{H}_{z_k}^{1/2}] + (n_x)_{ki} \left(Z_i \bar{E}_{y_k}^0 + Z_k \bar{E}_{yi}^0\right) [\bar{H}_{z_k}^{1/2}] + (n_y)_{ki} \left(Z_i \bar{E}_{x_k}^M + Z_k \bar{E}_{xi}^M\right) [\bar{H}_{z_k}^{M+1/2}] - (n_x)_{ki} \left(Z_i \bar{E}_{y_k}^M + Z_k \bar{E}_{yi}^M\right) [\bar{H}_{z_k}^{M+1/2}] \Bigg) ds.$$

Proof. Considering the equalities (3.9) and (3.10) for B_{11}^m we have

$$\begin{aligned} B_{11}^m = \frac{1}{2} \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \Bigg(& - \frac{(n_y)_{ki}}{Z_i+Z_k} \left(Z_i [\bar{H}_{z_k}^{m+1/2}] - \alpha \left((n_x)_{ki} [\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki} [\bar{E}_{x_k}^{[m+1/2]}]\right)\right) \bar{E}_{x_k}^m \\ & - \frac{(n_y)_{ki}}{Z_i+Z_k} \left(-\alpha \left((n_x)_{ki} [\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki} [\bar{E}_{x_k}^{[m+1/2]}]\right)\right) \bar{E}_{x_k}^{m+1} \\ & - \frac{(n_y)_{ik}}{Z_i+Z_k} \left(Z_k [\bar{H}_{zi}^{m+1/2}] - \alpha \left((n_x)_{ik} [\bar{E}_{y_i}^{[m+1/2]}] - (n_y)_{i_k} [\bar{E}_{xi}^{[m+1/2]}]\right)\right) \bar{E}_{xi}^m \\ & - \frac{(n_y)_{ik}}{Z_i+Z_k} \left(-\alpha \left((n_x)_{ik} [\bar{E}_{y_i}^{[m+1/2]}] - (n_y)_{ik} [\bar{E}_{x_i}^{[m+1/2]}]\right)\right) \bar{E}_{xi}^{m+1} \\ & - \frac{Y_i (n_y)_{ki}}{Y_i+Y_k} [\bar{E}_{x_k}^{m+1}] \bar{H}_{z_k}^{m+3/2} - \frac{Y_k (n_y)_{ik}}{Y_i+Y_k} [\bar{E}_{xi}^{m+1}] \bar{H}_{zi}^{m+3/2} \\ & + (n_y)_{ki} \bar{E}_{x_k}^{m+1} \bar{H}_{z_k}^{m+3/2} + (n_y)_{ik} \bar{E}_{xi}^{m+1} \bar{H}_{z_i}^{m+3/2} \Bigg) ds. \end{aligned}$$

Summing from $m = 0$ to $m = M-1$ we conclude that

$$\sum_{m=0}^{M-1} B_{11}^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{(n_y)_{ki}}{2(Z_i+Z_k)} \Bigg(- \left(Z_i \bar{E}_{x_k}^0 + Z_k \bar{E}_{xi}^0\right) [\bar{H}_{z_k}^{1/2}] + \left(Z_i \bar{E}_{x_k}^M + Z_k \bar{E}_{xi}^M\right) [\bar{H}_{z_k}^{M+1/2}] + 2\alpha \sum_{m=0}^{M-1} \left((n_x)_{ki} [\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki} [\bar{E}_{x_k}^{[m+1/2]}]\right) [\bar{E}_{x_k}^{[m+1/2]}] \Bigg) ds.$$

In the same way, for B_{12}^m we have

$$\sum_{m=0}^{M-1} B_{12}^m = \sum_{f_{ik}\in F^{int}} \int_{f_{ik}} \frac{(n_x)_{ki}}{2(Z_i+Z_k)} \Bigg(\left(Z_i \bar{E}_{y_k}^0 + Z_k \bar{E}_{yi}^0\right) [\bar{H}_{z_k}^{1/2}] - \left(Z_i \bar{E}_{y_k}^M + Z_k \bar{E}_{yi}^M\right) [\bar{H}_{z_k}^{M+1/2}] - 2\alpha \sum_{m=0}^{M-1} \left((n_x)_{ki} [\bar{E}_{y_k}^{[m+1/2]}] - (n_y)_{ki} [\bar{E}_{x_k}^{[m+1/2]}]\right) [\bar{E}_{y_k}^{[m+1/2]}] \Bigg) ds.$$

Then

$$\sum_{m=0}^{M-1}(B_{11}^m+B_{12}^m)=\sum_{f_{ik}\in F^{int}}\int_{f_{ik}}\frac{1}{2(Z_i+Z_k)}\Bigg(-(n_y)_{ki}\left(Z_i\bar{E}_{x_k}^0+Z_k\bar{E}_{xi}^0\right)[\bar{H}_{z_k}^{1/2}]+(n_x)_{ki}\left(Z_i\bar{E}_{y_k}^0+Z_k\bar{E}_{yi}^0\right)[\bar{H}_{z_k}^{1/2}]$$
$$-2\alpha\sum_{m=0}^{M-1}\left((n_x)_{ki}[\bar{E}_{y_k}^{[m+1/2]}]-(n_y)_{ki}[\bar{E}_{x_k}^{[m+1/2]}]\right)^2$$
$$+(n_y)_{ki}\left(Z_i\bar{E}_{x_k}^M+Z_k\bar{E}_{xi}^M\right)[\bar{H}_{z_k}^{M+1/2}]$$
$$-(n_x)_{ki}\left(Z_i\bar{E}_{y_k}^M+Z_k\bar{E}_{yi}^M\right)[\bar{H}_{z_k}^{M+1/2}]\Bigg)ds.$$

For B_{13}^m we have

$$\sum_{m=0}^{M-1}B_{13}^m\leq 0,$$

which concludes the proof. □

Let us now analyze the term on the boundary.

Lemma 4.2.2. *Let B_2^m be defined by (4.33). Then*

$$\sum_{m=0}^{M-1}B_2^m\leq\sum_{f_k\in F^{ext}}\int_{f_k}\frac{1}{4}\left(\bar{H}_{z_k}^{1/2}\left((n_x)_k\bar{E}_{y_k}^0-(n_y)_k\bar{E}_{x_k}^0\right)\bar{H}_{z_k}^{M+1/2}\left((n_x)_k\bar{E}_{y_k}^M-(n_y)_k\bar{E}_{x_k}^M\right)\right)ds.$$

Proof. We have

$$B_2^m=\frac{1}{2}\sum_{f_k\in F^{ext}}\int_{f_k}\Bigg(\left(-(n_y)_k\bar{H}_{z_k}^{m+1/2}+\frac{(n_y)_k}{Z_k}\left((n_x)_k\bar{E}_{y_k}^{[m+1/2]}-(n_y)_k\bar{E}_{x_k}^{[m+1/2]}\right)\right)\bar{E}_{x_k}^{[m+1/2]}$$
$$+\left((n_x)_k\bar{H}_{z_k}^{m+1/2}-\frac{(n_x)_k}{Z_k}\left((n_x)_k\bar{E}_{y_k}^{[m+1/2]}-(n_y)_k\bar{E}_{x_k}^{[m+1/2]}\right)\right)\bar{E}_{y_k}^{[m+1/2]}$$
$$-\left(\frac{1}{Y_k}\bar{H}_{z_k}^{[m+1]}+(n_x)_k\bar{E}_{y_k}^{m+1}-(n_y)_k\bar{E}_{x_k}^{m+1}\right)\bar{H}_{z_k}^{[m+1]}\Bigg)ds.$$

Summing from $m=0$ to $m=M-1$, we deduce that

$$\sum_{m=0}^{M-1}B_2^m\leq\sum_{f_k\in F^{ext}}\int_{f_k}\frac{1}{4}\left(\bar{H}_{z_k}^{1/2}\left((n_x)_k\bar{E}_{y_k}^0-(n_y)_k\bar{E}_{x_k}^0\right)-\bar{H}_{z_k}^{M+1/2}\left((n_x)_k\bar{E}_{y_k}^M-(n_y)_k\bar{E}_{x_k}^M\right)\right)ds.$$

□

Theorem 4.2.3. *Let us consider the leap-frog DG method (4.24)–(4.26) with SM-ABC. If the time step Δt is such that*

$$\Delta t<\frac{\min\{\underline{\varepsilon},\underline{\mu}\}}{C_N}\min\{h_k\},\tag{4.34}$$

where

$$C_N = \frac{1}{2}C_{inv}N^2 + 2C_\tau^2(N+1)(N+2),$$

with C_τ defined by (A.4) of Lemma A.1.1 and C_{inv} defined by (A.6) of Lemma A.1.2, then the method is stable.

Proof. From (4.29) and the previous lemmata, considering the Cauchy-Schwarz's inequality and taking into account that $Z_i/(Z_i+Z_k) < 1$, we obtain

$$\begin{aligned}
\sum_{T_k\in\mathscr{T}_h}\left((\varepsilon\bar{E}_k^M,\bar{E}_k^M)_{T_k}+(\mu\bar{H}_{z_k}^{M+1/2},\bar{H}_{z_k}^{M+1/2})_{T_k}\right) &\leq \sum_{T_k\in\mathscr{T}_h}\left((\varepsilon\bar{E}_k^0,\bar{E}_k^0)_{T_k}+(\mu\bar{H}_{z_k}^{1/2},\bar{H}_{z_k}^{1/2})_{T_k}\right)\\
&+\Delta t\sum_{T_k\in\mathscr{T}_h}\left(\|\nabla\times\bar{H}_{z_k}^{1/2}\|_{L^2(T_k)}\|\bar{E}_k^0\|_{L^2(T_k)}+\|\nabla\times\bar{H}_{z_k}^{M+1/2}\|_{L^2(T_k)}\|\bar{E}_k^M\|_{L^2(T_k)}\right)\\
&+\Delta t\sum_{f_{ik}\in F^{int}}\left((\|\bar{E}_k^M\|_{L^2(f_{ik})}+\|\bar{E}_i^M\|_{L^2(f_{ik})})\|[\bar{H}_{z_k}^{M+1/2}]\|_{L^2(f_{ik})}+(\|\bar{E}_k^0\|_{L^2(f_{ik})}+\|\bar{E}_k^0\|_{L^2(f_{ik})})\|[\bar{H}_{z_k}^{1/2}]\|_{L^2(f_{ik})}\right)\\
&+\Delta t\sum_{f_k\in F^{ext}}\left(\|\bar{H}_{z_k}^{1/2}\|_{L^2(f_k)}\|\bar{E}_k^0\|_{L^2(f_k)}+\|\bar{H}_{z_k}^{M+1/2}\|_{L^2(f_k)}\|\bar{E}_k^M\|_{L^2(f_k)}\right).
\end{aligned}$$

Using the inequality (A.4) of Lemma A.1.1 and the inequality (A.6) of Lemma A.1.2 (both in Appendix), we get

$$\begin{aligned}
\min\{\underline{\varepsilon},\underline{\mu}\}\left(\|\bar{E}^M\|_{L^2(\Omega)}^2+\|\bar{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\right) &\leq \max\{\bar{\varepsilon},\bar{\mu}\}\left(\|\bar{E}^0\|_{L^2(\Omega)}^2+\|\bar{H}_z^{1/2}\|_{L^2(\Omega)}^2\right)\\
&+\frac{\Delta t}{2}C_{inv}N^2\max\left\{h_k^{-1}\right\}\left(\|\bar{H}_z^{1/2}\|_{L^2(\Omega)}^2+\|\bar{E}^0\|_{L^2(\Omega)}^2+\|\bar{H}_z^{M+1/2}\|_{L^2(\Omega)}^2+\|\bar{E}^M\|_{L^2(\Omega)}^2\right)\\
&+2C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(\|\bar{E}^M\|_{L^2(\Omega)}^2+\|\bar{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\right)\\
&+2C_\tau^2(N+1)(N+2)\Delta t\max\left\{h_k^{-1}\right\}\left(\|\bar{E}^0\|_{L^2(\Omega)}^2+\|\bar{H}_z^{1/2}\|_{L^2(\Omega)}^2\right).
\end{aligned}$$

and so, taking $C_0 = \frac{1}{2}C_{inv}N^2 + 2C_\tau^2(N+1)(N+2)$,

$$\begin{aligned}
\left(\min\{\underline{\varepsilon},\underline{\mu}\}-\Delta t\max\left\{h_k^{-1}\right\}C_N\right)\left(\|\bar{E}^M\|_{L^2(\Omega)}^2+\|\bar{H}_z^{M+1/2}\|_{L^2(\Omega)}^2\right) &\leq\\
\left(\max\{\bar{\varepsilon},\bar{\mu}\}+\Delta t\max\left\{h_k^{-1}\right\}C_0\right)\left(\|\bar{E}^0\|_{L^2(\Omega)}^2+\|\bar{H}_z^{1/2}\|_{L^2(\Omega)}^2\right),
\end{aligned}$$

which concludes the proof. □

Although the implicit method (4.24)–(4.26) is conditionable stable, the stability condition (4.34) is less restrictive when compared with the stability condition (3.14) for the explicit method (2.32)–(2.34). Moreover, following the proof of Theorem 4.1.5, we can easily derive that the implicit method is second order convergent in time even for the case of SM-ABC. Note that the drawback of this method is its computational efficiency.

4.2.2 Iterative explicit method

The equations (4.24)–(4.26) are implicit and thus the computational effort to compute the numerical solution is higher when compared with explicit schemes. To avoid this, we will consider an iterative

process being each iteration explicit. This process starts, as before, with an approximation to the initial data which we denote now by $(\hat{E}^0_{x_k}, \hat{E}^0_{y_k}, \hat{H}^{1/2}_{z_k}) \in V_N$. For each $m = 0, 1, \ldots, M-1$, we initialize the iterative process by

$$\hat{E}^{m+1,0}_{x_k} = \hat{E}^m_{x_k}, \quad \hat{E}^{m+1,0}_{y_k} = \hat{E}^m_{y_k}, \quad \hat{H}^{m+1/2,0}_{z_k} = \hat{H}^{m+1/2}_{z_k}.$$

The $n+1^{\text{th}}$ inner iteration of the iterative scheme, for $n = 0, 1, 2, \ldots$, is: find $(\hat{E}^{m+1,n+1}_{x_k}, \hat{E}^{m+1,n+1}_{y_k}, \hat{H}^{m+1/2,n+1}_{z_k}) \in V_N$ such that, $\forall (u_k, v_k, w_k) \in V_N$ we have

$$\begin{aligned}
&\left(\varepsilon_{xx}\frac{\hat{E}^{m+1,n+1}_{x_k} - \hat{E}^m_{x_k}}{\Delta t} + \varepsilon_{xy}\frac{\hat{E}^{m+1,n+1}_{y_k} - \hat{E}^m_{y_k}}{\Delta t}, u_k\right)_{T_k} = \left(\partial_y \hat{H}^{m+1/2}_{z_k}, u_k\right)_{T_k} \\
&+ \left(\frac{-n_y}{Z^+ + Z^-}\left(Z^+[\hat{H}^{m+1/2}_z] - \alpha\left(n_x[\hat{E}^{[m+1/2,n]}_y] - n_y[\hat{E}^{[m+1/2,n]}_x]\right)\right), u_k\right)_{\partial T_k},
\end{aligned} \tag{4.35}$$

$$\begin{aligned}
&\left(\varepsilon_{yx}\frac{\hat{E}^{m+1,n+1}_{x_k} - \hat{E}^m_{x_k}}{\Delta t} + \varepsilon_{yy}\frac{\hat{E}^{m+1,n+1}_{y_k} - \hat{E}^m_{y_k}}{\Delta t}, v_k\right)_{T_k} = -\left(\partial_x \hat{H}^{m+1/2}_{z_k}, v_k\right)_{T_k} \\
&+ \left(\frac{n_x}{Z^+ + Z^-}\left(Z^+[\hat{H}^{m+1/2}_z] - \alpha\left(n_x[\hat{E}^{[m+1/2,n]}_y] - n_y[\hat{E}^{[m+1/2,n]}_x]\right)\right), v_k\right)_{\partial T_k},
\end{aligned} \tag{4.36}$$

$$\begin{aligned}
&\left(\mu\frac{\hat{H}^{m+3/2,n+1}_{z_k} - \hat{H}^{m+1/2}_{z_k}}{\Delta t}, w_k\right)_{T_k} = \left(\partial_y \hat{E}^{m+1}_{x_k} - \partial_x \hat{E}^{m+1}_{y_k}, w_k\right)_{T_k} \\
&+ \left(\frac{1}{Y^+ + Y^-}\left(Y^+(n_x[\hat{E}^{m+1}_y] - n_y[\hat{E}^{m+1}_x]) - \alpha[\hat{H}^{[m+1,n]}_z]\right), w_k\right)_{\partial T_k},
\end{aligned} \tag{4.37}$$

where

$$\hat{E}^{[m+1/2,n]} = \frac{\hat{E}^m + \hat{E}^{m+1,n}}{2}, \quad \hat{H}^{[m+1,n]} = \frac{\hat{H}^{m+1/2} + \hat{H}^{m+3/2,n}}{2}.$$

When following convergence criterion is satisfied

$$\|\hat{E}^{m+1,n+1} - \hat{E}^{m+1,n}\|_{L^2(\Omega)} < \text{tol}, \quad \|\hat{H}^{m+3/2,n+1}_z - \hat{H}^{m+3/2,n}_z\|_{L^2(\Omega)} < \text{tol},$$

for some pre-defined small constant tol, then the current time step $m+1$ is terminated and the correspondent numerical solution is denoted by $(\hat{E}^{m+1}_{x_k}, \hat{E}^{m+1}_{y_k}, \hat{H}^{m+1/2}_{z_k})$. If we only perform the iteration $n = 0$ we obtain the same method as in Remark 4.1.6.

We will show that the solution of the iterative predictor-corrector scheme (4.35)–(4.37) converges to the solution of the method (4.24)–(4.26) under certain stability conditions. Let define the difference

between two successive numeric values of electromagnetic fields by

$$\begin{aligned}\delta_n \hat{E}_{x_k}^{m+1} &= \hat{E}_{x_k}^{m+1,n+1} - \hat{E}_{x_k}^{m+1,n},\\ \delta_n \hat{E}_{y_k}^{m+1} &= \hat{E}_{y_k}^{m+1,n+1} - \hat{E}_{y_k}^{m+1,n},\\ \delta_n \hat{H}_{z_k}^{m+3/2} &= \hat{H}_{z_k}^{m+3/2,n+1} - \hat{H}_{z_k}^{m+3/2,n},\end{aligned}$$

for $n = 0, 1, 2, \cdots$. We will find an upper bound for $\delta_n \hat{E}_{x_k}^{m+1}$, $\delta_n \hat{E}_{y_k}^{m+1}$ and $\delta_n \hat{H}_{z_k}^{m+3/2}$ in the following theorem.

Theorem 4.2.4. *The the solution of the iterative predictor-corrector scheme (4.35)–(4.37) converges to the solution of the method (4.24)–(4.26) provided that (3.14) is satisfied.*

Proof. The condition (3.14) ensures that $\|\delta_0 \hat{E}^{m+1}\|_{L^2(\Omega)}$ and $\|\delta_0 \hat{H}_z^{m+3/2}\|_{L^2(\Omega)}$ are bounded. Taking the difference of (4.35)–(4.37) between two successive iterations, $n+1$ and n, and replacing u_k, v_k and w_k by, respectively, $\delta_n \hat{E}_{x_k}^{m+1}$, $\delta_n \hat{E}_{y_k}^{m+1}$ and $\delta_n \hat{H}_{z_k}^{m+3/2}$ we obtain

$$\begin{aligned}\sum_{T_k \in \mathscr{T}_h} \left(\varepsilon \delta_n \hat{E}_k^{m+1}, \delta_n \hat{E}_k^{m+1}\right)_{T_k} = \frac{\Delta t}{2} \sum_{f_{ik} \in F^{int}} \int_{f_{ik}} \Bigg(& \frac{(n_y)_{ki}}{Z_i + Z_k} \left((n_x)_{ki} \delta_{n-1}[\hat{E}_{y_k}^{m+1}] - (n_y)_{ki} \delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right) \delta_n \hat{E}_{x_k}^{m+1} \\ & + \frac{(n_y)_{ik}}{Z_i + Z_k} \left((n_x)_{ik} \delta_{n-1}[\hat{E}_{yi}^{m+1}] - (n_y)_{ik} \delta_{n-1}[\hat{E}_{xi}^{m+1}]\right) \delta_n \hat{E}_{x_i}^{m+1} \Bigg) ds \\ - \frac{\Delta t}{2} \sum_{f_{ik} \in F^{int}} \int_{f_{ik}} \Bigg(& \frac{(n_x)_{ki}}{Z_i + Z_k} \left((n_x)_{ki} \delta_{n-1}[\hat{E}_{y_k}^{m+1}] - (n_y)_{ki} \delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right) \delta_n \hat{E}_{y_k}^{m+1} \\ & + \frac{(n_x)_{ik}}{Z_i + Z_k} \left((n_x)_{ik} \delta_{n-1}[\hat{E}_{yi}^{m+1}] - (n_y)_{ik} \delta_{n-1}[\hat{E}_{xi}^{m+1}]\right) \delta_n \hat{E}_{y_i}^{m+1} \Bigg) ds \\ + \frac{\Delta t}{2} \sum_{f_k \in F^{ext}} \int_{f_k} \Bigg(& \frac{(n_y)_k}{2Z_k} \left((n_x)_k \delta_{n-1}[\hat{E}_{y_k}^{m+1}] - (n_y)_k \delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right) \delta_n \hat{E}_{x_k}^{m+1} \\ & - \frac{(n_x)_k}{2Z_k} \left((n_x)_k \delta_{n-1}[\hat{E}_{y_k}^{m+1}] - (n_y)_k \delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right) \delta_n \hat{E}_{y_k}^{m+1} \Bigg) ds,\end{aligned}$$

$$\begin{aligned}&\sum_{T_k \in \mathscr{T}_h} \left(\mu \delta_n \hat{H}_{z_k}^{m+3/2}, \delta_n \hat{H}_{z_k}^{m+3/2}\right)_{T_k} = \\ &\quad - \frac{\Delta t}{2} \sum_{f_{ik} \in F^{int}} \int_{f_{ik}} \left(\frac{1}{Y_i + Y_k} \delta_{n-1}[\hat{H}_{z_k}^{m+3/2}] \delta_n \hat{H}_{z_k}^{m+3/2} + \frac{1}{Y_i + Y_k} \delta_{n-1}[\hat{H}_{zi}^{m+3/2}] \delta_n \hat{H}_{zi}^{m+3/2} \right) ds \\ &\quad - \frac{\Delta t}{2} \sum_{f_k \in F^{ext}} \int_{f_k} \left(\frac{1}{2Y_k} \delta_{n-1}[\hat{H}_{z_k}^{m+3/2}] \delta_n \hat{H}_{z_k}^{m+3/2} \right) ds.\end{aligned}$$

Then

$$\sum_{T_k\in\mathscr{T}_h}\left(\varepsilon\delta_n\hat{E}_k^{m+1},\delta_n\hat{E}_k^{m+1}\right)_{T_k}=\frac{\Delta t}{2}\sum_{f_{ik}\in F^{int}}\int_{f_{ik}}\left(\frac{(n_y)_{ki}}{Z_i+Z_k}\left((n_x)_{ki}\delta_{n-1}[\hat{E}_{y_k}^{m+1}]-(n_y)_{ki}\delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right)\delta_n[\hat{E}_{x_k}^{m+1}]\right)ds$$
$$-\frac{\Delta t}{2}\sum_{f_{ik}\in F^{int}}\int_{f_{ik}}\left(\frac{(n_x)_{ki}}{Z_i+Z_k}\left((n_x)_{ki}\delta_{n-1}[\hat{E}_{y_k}^{m+1}]-(n_y)_{ki}\delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right)\delta_n[\hat{E}_{y_k}^{m+1}]\right)ds$$
$$+\frac{\Delta t}{2}\sum_{f_k\in F^{ext}}\int_{f_k}\left(\frac{(n_y)_k}{2Z_k}\left((n_x)_k\delta_{n-1}[\hat{E}_{y_k}^{m+1}]-(n_y)_k\delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right)\delta_n\hat{E}_{x_k}^{m+1}\right.$$
$$\left.-\frac{(n_x)_k}{2Z_k}\left((n_x)_k\delta_{n-1}[\hat{E}_{y_k}^{m+1}]-(n_y)_k\delta_{n-1}[\hat{E}_{x_k}^{m+1}]\right),\delta_n\hat{E}_{y_k}^{m+1}\right)ds,$$

and

$$\sum_{T_k\in\mathscr{T}_h}\left(\mu\,\delta_n\hat{H}_{z_k}^{m+3/2},\delta_n\hat{H}_{z_k}^{m+3/2}\right)_{T_k}=-\frac{\Delta t}{2}\sum_{f_{ik}\in F^{int}}\int_{f_{ik}}\left(\frac{1}{Y_i+Y_k}\delta_{n-1}[\hat{H}_{z_k}^{m+3/2}]\delta_n[\hat{H}_{z_k}^{m+3/2}]\right)ds$$
$$-\frac{\Delta t}{2}\sum_{f_k\in F^{ext}}\int_{f_k}\left(\frac{1}{2Y_k}\delta_{n-1}[\hat{H}_{z_k}^{m+3/2}]\delta_n\hat{H}_{z_k}^{m+3/2}\right)ds.$$

So

$$\sum_{T_k\in\mathscr{T}_h}\left(\varepsilon\delta_n\hat{E}_k^{m+1},\delta_n\hat{E}_k^{m+1}\right)_{T_k}\le\frac{\Delta t}{4\min\{Z_k\}}\sum_{f_{ik}\in F^{int}}\|\delta_{n-1}[\hat{E}_k^{m+1}]\|_{L^2(f_{ik})}\|\delta_n[\hat{E}_k^{m+1}]\|_{L^2(f_{ik})}$$
$$+\frac{\Delta t}{4\min\{Z_k\}}\sum_{f_k\in F^{ext}}\|\delta_{n-1}[\hat{E}_k^{m+1}]\|_{L^2(f_k)}\|\delta_n[\hat{E}_k^{m+1}]\|_{L^2(f_k)},$$

$$\sum_{T_k\in\mathscr{T}_h}\left(\mu\,\delta_n\hat{H}_{z_k}^{m+3/2},\delta_n\hat{H}_{z_k}^{m+3/2}\right)_{T_k}\le\frac{\Delta t}{4\min\{Y_k\}}\sum_{f_{ik}\in F^{int}}\|\delta_{n-1}[\hat{H}_{z_k}^{m+3/2}]\|_{L^2(f_{ik})}\|\delta_n[\hat{H}_{z_k}^{m+3/2}]\|_{L^2(f_{ik})}$$
$$+\frac{\Delta t}{4\min\{Y_k\}}\sum_{f_k\in F^{ext}}\|\delta_{n-1}[\hat{H}_{z_k}^{m+3/2}]\|_{L^2(f_k)}\|\delta_n\hat{H}_{z_k}^{m+3/2}\|_{L^2(f_k)}.$$

Consequently, we obtain

$$\underline{\varepsilon}\|\delta_n\hat{E}^{m+1}\|_{L^2(\Omega)}\le\frac{\Delta t}{\min\{Z_k\}}C_\tau^2(N+1)(N+2)\max\left\{h_k^{-1}\right\}\|\delta_{n-1}\hat{E}^{m+1}\|_{L^2(\Omega)},$$

$$\underline{\mu}\ \|\delta_n\hat{H}_{z_k}^{m+3/2}\|_{L^2(\Omega)}\le\frac{\Delta t}{\min\{Y_k\}}C_\tau^2(N+1)(N+2)\max\left\{h_k^{-1}\right\}\|\delta_{n-1}\hat{H}_z^{m+3/2}\|_{L^2(\Omega)}.$$

Taking condition (3.14) into account, we conclude the proof. □

Chapter 5

Numerical Results

This chapter is dedicated to a detailed numerical evaluation of the proposed leap-frog DG methods on triangular meshes for solving two dimensional TE Maxwell's equations. We assess the stability condition and the error estimate of the method through numerical experiments for 2D wave propagation in anisotropic media with PEC and Silver-Müller boundary conditions (SM-ABC). Results for PMC boundary conditions were also obtained but not presented since they are similar to the equivalent ones for the PEC case.

In our simulations we consider a non diagonal and possibly spatially-varying symmetric and positive definite permittivity tensor. While it is sometimes possible to find a rotated coordinate system in which this permittivity tensor is diagonal [5, 49], it is not clear how to apply such a rotated reference frame in general. In particular, when the simulation domain contains two or more anisotropic regions with different orientations, it would be impossible to diagonalize for all materials under the same rotation. Furthermore, if the permittivity tensor contains imaginary off-diagonal elements, as in a magneto-optic material [31], the permittivity tensor cannot be diagonalized by a simple rotation. Here, we avoid any rotation.

In order to have a suitable analytic solution for computing the error, corresponding to each of boundary conditions and each type of anisotropic permittivity tensor (constant or space-dependent), source terms are added to the system of Maxwell's equations. The Matlab codes in [39] are the inception codes we use for set up and are updated to deal with our particular numerical scheme and model specifications.

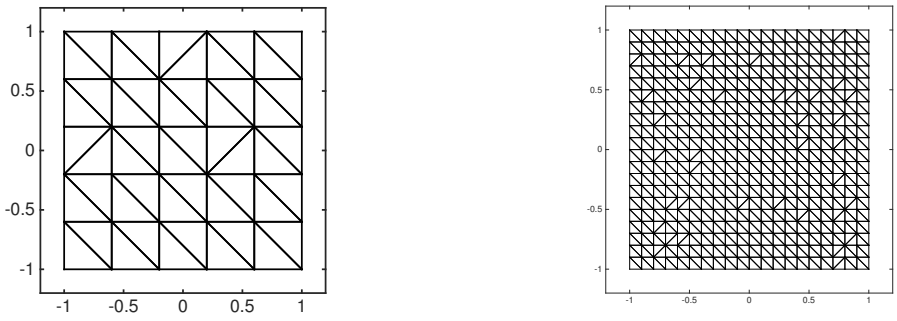
This chapter starts with the definition of the simulation setting. The structure of domain discretization and the characteristics of meshes we use in our tests are stated in Section 5.1. The sharpness of the stability conditions obtained in Theorem 3.1.3 is checked in Section 5.2 for both central and upwind fluxes while considering PEC and SM-ABC. The convergence result achieved in Theorem 4.1.5 is illustrated in Section 5.3 for the same set of experiments considered for the stability analysis. The efficiency of the proposed predictor-corrector method to recover the temporal convergence order is checked in the last part of the same section. All the stability and convergence results are obtain for two permittivity tensors: constant and space-dependent tensor. We back to the motivation behind this work in Section 5.4 and discuss the scattering through eye's structure.

K	32	50	200	800	3200
Nv	25	36	121	441	1681
h_{min}	0.7071	0.5657	0.2828	0.1414	0.0707

Table 5.1 The number of triangle elements and vertices in computational meshes used in the computations. h_{min} denotes the shortest distance between two vertices in the mesh.

5.1 Simulation setting

In the following we consider that two-dimension TE mode of Maxwell's equations (5.11)–(5.13) are space discretized using discontinuous Galerkin method on triangular mesh. The computation domain is considered as the square $\Omega = (-1,1)^2$, which is triangulated with K non-overlapping straight-sided triangles. The characteristics of the meshes used in the computations are summarised in Table 5.1. Two examples of meshes on the computational domain are presented in Figure 5.1. On each triangle



(a) $K = 50$ (b) $K = 800$

Fig. 5.1 Examples of computational mesh on a square domain used in 2D computations.

we define

$$N_p = \frac{(N+1)(N+2)}{2}$$

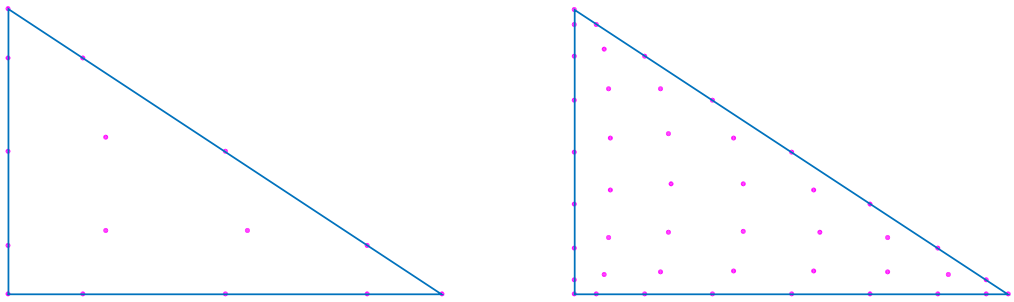
nodal points, where N is the order of polynomial approximation. In the simulations we consider the warp-blend points [85]. Figure 5.2 shows the distribution of these points on a sample triangle element for $N = 4$ and $N = 8$.

In our experiments we consider a non diagonal symmetric positive definite permittivity tenor first as a constant tensor

$$\varepsilon = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}, \tag{5.1}$$

and then as a space-dependent tensor

$$\varepsilon(x,y) = \begin{pmatrix} 4x^2+y^2+1 & \sqrt{x^2+y^2} \\ \sqrt{x^2+y^2} & x^2+1 \end{pmatrix}. \tag{5.2}$$



(a) $N = 4$ (b) $N = 8$

Fig. 5.2 Distribution of warp-blend nodal points in a triangular element.

In all the simulations we consider isotropic permeability which is set as $\mu = 1$.

Let us consider the Maxwell's equations defined on the square $\Omega = (-1,1)^2$, complemented with initial and boundary conditions. For the case of PEC boundary condition we consider the initial conditions

$$E_x(x,y,0) = 0, \tag{5.3}$$
$$E_y(x,y,0) = 0, \tag{5.4}$$
$$H_z(x,y,\Delta t/2) = \cos(\pi x)\cos(\pi y)\cos(\omega \Delta t/2), \tag{5.5}$$

where

$$\omega = \pi\sqrt{\frac{1}{\varepsilon_{xx}} + \frac{1}{\varepsilon_{yy}}}, \tag{5.6}$$

and for SM-ABC we consider the initial conditions

$$E_x(x,y,0) = 0, \tag{5.7}$$
$$E_y(x,y,0) = 0, \tag{5.8}$$
$$H_z(x,y,\Delta t/2) = \sin(\pi \Delta t/2)\sin(\pi xy). \tag{5.9}$$

The sharpness of stability result as well as the spatial order of convergence for different boundary conditions, for different degree of polynomial approximation and for both central and upwind fluxes will be illustrated while refining the mesh according to Table 5.1. The temporal order of convergence will be analyzed in a space grid with $K = 800$ and $N = 8$. In all of the experiments the simulation time is fixed at $T = 1$.

5.2 Stability condition

We will check numerically that (3.14) defines a sharp stability condition, in terms of the influence of N, the order of the polynomial approximation, and h_{min}, the minimum triangle diameter in the mesh. In our experiments, we computed C that satisfies

$$\Delta t_{max} = \frac{C}{(N+1)(N+2)} h_{min}, \tag{5.10}$$

where Δt_{max} is the maximum observed value of Δt such that the method is stable.

In Table 5.2 and Table 5.3 the results are computed for different mesh sizes, considering respectively central and upwind fluxes in the DG method, for the case of PEC boundary conditions, while in Table 5.4 and Table 5.5, the results are computed for the case of SM-ABC.

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.17	1.80	0.1	2.12	0.065	2.30	0.044	2.33	0.032	2.37
0.2828	0.088	1.87	0.05	2.12	0.031	2.20	0.021	2.23	0.016	2.37
0.1414	0.044	1.87	0.024	2.04	0.015	2.12	0.01	2.12	0.0078	2.32
0.0707	0.021	1.78	0.012	2.04	0.0078	2.20	0.0054	2.30	0.0038	2.26
0.0354	0.01	1.70	0.006	2.04	0.0039	2.20	0.0027	2.30	0.0019	2.26
0.0177	0.0054	1.83	0.003	2.04	0.0019	2.15	0.0013	2.21	0.00095	2.26

Table 5.2 Δt_{max} such that the method is stable and C computed by (5.10) for PEC boundary conditions, central flux and constant permittivity tensor (5.1).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.10	1.06	0.056	1.19	0.034	1.20	0.023	1.22	0.016	1.19
0.2828	0.047	1.00	0.026	1.10	0.016	1.13	0.011	1.17	0.0081	1.20
0.1414	0.023	0.98	0.012	1.02	0.008	1.13	0.0054	1.15	0.0039	1.16
0.0707	0.011	0.93	0.0062	1.05	0.0039	1.10	0.0026	1.10	0.0019	1.13
0.0354	0.0055	0.93	0.003	1.02	0.0019	1.07	0.0013	1.10	0.0009	1.07
0.0177	0.0027	0.92	0.0015	1.02	0.0009	1.02	0.0006	1.02	0.0004	0.95

Table 5.3 Δt_{max} such that the method is stable and C computed by (5.10) for PEC boundary conditions, upwind flux and constant permittivity tensor (5.1).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.18	1.91	0.1	2.12	0.064	2.26	0.044	2.33	0.031	2.30
0.2828	0.092	1.95	0.05	2.12	0.031	2.19	0.021	2.02	0.015	2.23
0.1414	0.044	1.87	0.024	2.04	0.015	2.12	0.01	2.12	0.0079	2.35
0.0707	0.021	1.78	0.012	2.04	0.0077	2.18	0.0053	2.25	0.0038	2.26
0.0354	0.01	1.70	0.006	2.04	0.0038	2.15	0.0026	2.21	0.0019	2.26
0.0177	0.0053	1.80	0.003	2.04	0.0018	2.04	0.0012	2.04	0.00095	2.26

Table 5.4 Δt_{max} such that the method is stable and C computed by (5.10) for SM-ABC, central flux and constant permittivity tensor (5.1).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.11	1.17	0.057	1.21	0.035	1.24	0.023	1.22	0.016	1.19
0.2828	0.051	1.08	0.026	1.10	0.016	1.13	0.011	1.17	0.008	1.19
0.1414	0.023	0.98	0.012	1.02	0.008	1.13	0.0054	1.15	0.0039	1.16
0.0707	0.011	0.93	0.0061	1.04	0.0039	1.10	0.0026	1.10	0.0019	1.13
0.0354	0.0055	0.93	0.003	1.02	0.0018	1.07	0.0013	1.10	0.00097	1.15
0.0177	0.0027	0.92	0.0015	1.02	0.00097	1.10	0.00065	1.10	0.00045	1.07

Table 5.5 Δt_{max} such that the method is stable and C computed by (5.10) for SM-ABC, upwind flux and constant permittivity tensor (5.1).

As expected from the condition (3.14), the numerical examples in Table 5.2 and Table 5.3 show that the stability regions corresponding to central fluxes are slightly bigger when compared to the regions obtained using upwind fluxes. From all the examples presented, we may deduce that the right hand side of (3.14) is a sharp bound for Δt_{max}. Moreover, we can also conclude that Δt_{max} is directly proportional to h_{min} and inversely proportional to $(N+1)(N+2)$.
In the same framework, we consider the space-dependent tensor (5.2). The experiments are repeated. The collected data is summarized in Table 5.6–5.9 for PEC and SM-ABC boundary conditions for both central and upwind flux.

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.11	1.27	0.065	1.38	0.04	1.41	0.027	1.43	0.02	1.48
0.2828	0.056	1.19	0.031	1.32	0.019	1.34	0.013	1.38	0.0099	1.47
0.1414	0.027	1.15	0.015	1.27	0.0095	1.34	0.0066	1.40	0.0048	1.42
0.0707	0.013	1.10	0.0074	1.26	0.0047	1.33	0.0033	1.40	0.0024	1.43
0.0354	0.0067	1.14	0.0037	1.25	0.0023	1.30	0.0016	1.36	0.0012	1.43
0.0177	0.0033	1.12	0.0018	1.22	0.0011	1.24	0.00079	1.34	0.00059	1.40

Table 5.6 Δt_{max} such that the method is stable and C computed by (5.10) for PEC boundary conditions, central flux and space-dependent permittivity tensor (5.2).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.064	0.68	0.033	0.70	0.021	0.74	0.014	0.74	0.01	0.74
0.2828	0.029	0.62	0.016	0.68	0.01	0.70	0.0069	0.73	0.0049	0.73
0.1414	0.014	0.59	0.0076	0.65	0.0048	0.68	0.0033	0.70	0.0024	0.71
0.0707	0.0068	0.58	0.0037	0.63	0.0023	0.65	0.0016	0.68	0.0011	0.65
0.0354	0.0033	0.56	0.0018	0.61	0.0011	0.62	0.00081	0.69	0.0057	0.68
0.0177	0.0017	0.58	0.0088	0.60	0.00054	0.61	0.0004	0.68	0.00028	0.67

Table 5.7 Δt_{max} such that the method is stable and C computed by (5.10) for PEC boundary conditions, upwind flux and space-dependent permittivity tensor (5.2).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.12	1.27	0.065	1.38	0.04	1.41	0.027	1.43	0.02	1.48
0.2828	0.057	1.20	0.03	1.27	0.019	1.34	0.013	1.38	0.0098	1.46
0.1414	0.027	1.15	0.014	1.19	0.0095	1.34	0.0065	1.38	0.0048	1.42
0.0707	0.013	1.10	0.0073	1.24	0.0047	1.33	0.0032	1.35	0.0024	1.43
0.0354	0.0067	1.14	0.0036	1.22	0.0023	1.30	0.0016	1.36	0.0012	1.43
0.0177	0.0033	1.12	0.0018	1.22	0.0011	1.24	0.00079	1.34	0.00059	1.40

Table 5.8 Δt_{max} such that the method is stable and C computed by (5.10) for SM-ABC, central flux and space-dependent permittivity tensor (5.2).

h_{min}	$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C	Δt_{max}	C
0.5657	0.068	0.72	0.035	0.74	0.021	0.74	0.014	0.74	0.01	0.74
0.2828	0.03	0.64	0.016	0.68	0.01	0.70	0.0069	0.73	0.005	0.74
0.1414	0.014	0.59	0.0076	0.65	0.0048	0.68	0.0033	0.70	0.0024	0.71
0.0707	0.0068	0.58	0.0037	0.63	0.0024	0.68	0.0016	0.68	0.0012	0.72
0.0354	0.0033	0.56	0.0018	0.61	0.0011	0.62	0.00081	0.69	0.0058	0.69
0.0177	0.0017	0.58	0.0088	0.60	0.00054	0.61	0.0004	0.68	0.00028	0.67

Table 5.9 Δt_{max} such that the method is stable and C computed by (5.10) for SM-ABC, upwind flux and space-dependent permittivity tensor (5.2).

5.3 Order of convergence

In this section, we will illustrate the theoretical results of convergence. We consider the model problem

$$\varepsilon_{xx}\frac{\partial E_x}{\partial t}+\varepsilon_{xy}\frac{\partial E_y}{\partial t}=\frac{\partial H_z}{\partial y}+P(x,y,t), \tag{5.11}$$

$$\varepsilon_{yx}\frac{\partial E_x}{\partial t}+\varepsilon_{yy}\frac{\partial E_y}{\partial t}=-\frac{\partial H_z}{\partial x}+Q(x,y,t), \tag{5.12}$$

$$\mu\frac{\partial H_z}{\partial t}=-\frac{\partial E_y}{\partial x}+\frac{\partial E_x}{\partial y}+R(x,y,t), \tag{5.13}$$

defined in the square $\Omega=(-1,1)^2$. The source terms $P(x,y,t)$, $Q(x,y,t)$ and $R(x,y,t)$ are introduced in order to make it easier to find examples with known exact solution and consequently with the possibility to compute the error of the numerical solution. The problem is complemented with initial and boundary conditions in the same way as in the previous section.

We will analyze the behaviour of the proposed scheme by computing the following L^2-errors in each numerical example:

$$\text{Error } \tilde{E}_x=\|E_x^M-\tilde{E}_x^M\|_{L^2(\Omega)}, \tag{5.14}$$

$$\text{Error } \tilde{E}_y=\|E_y^M-\tilde{E}_y^M\|_{L^2(\Omega)}, \tag{5.15}$$

$$\text{Error } \tilde{H}_z=\|H_z^{M+1/2}-\tilde{H}_z^{M+1/2}\|_{L^2(\Omega)}. \tag{5.16}$$

For the computation of convergence rates in space, we herein use

$$\text{Order}=\frac{\log(\text{Error } \tilde{U}_{h,\Delta t}/\text{Error } \tilde{U}_{h^*,\Delta t})}{\log(h/h^*)}, \tag{5.17}$$

where $U_{h,\Delta t}$ and $U_{h^*,\Delta t}$ denote, respectively, the exact solution and the numerical solutions computed for time step Δt and two consecutive meshes of diameters h and h^*. For the computation of convergence rates in time we proceed in a correspondent way computing the errors of the numerical solutions for

two different values of Δt a fixed value for the meshe of diameter h, i.e.,

$$\text{Order} = \frac{\log(\text{Error } \tilde{U}_{h,\Delta t}/\text{Error } \tilde{U}_{h,\Delta t^*})}{\log(\Delta t/\Delta t^*)}. \tag{5.18}$$

5.3.1 PEC boundary condition

Let us consider the equations (5.11)–(5.13) complemented with PEC boundary condition and initial conditions (5.3)–(5.5). The source terms P, Q and R are obtained such that the problem has the exact solution

$$\begin{aligned}
E_x(x,y,t) &= \frac{-\pi}{\omega\varepsilon_{xx}}\cos(\pi x)\sin(\pi y)\sin(\omega t),\\
E_y(x,y,t) &= \frac{\pi}{\omega\varepsilon_{yy}}\sin(\pi x)\cos(\pi y)\sin(\omega t),\\
H_z(x,y,t) &= \cos(\pi x)\cos(\pi y)\cos(\omega t),
\end{aligned}$$

where ω is given by (5.6). For this type of boundary conditions, the theoretical convergence analysis presented in the previous chapter for the proposed DG leap-frog integrator (2.32)–(2.34) showed that the order of convergence in space and time is $\mathscr{O}(h^N) + \mathscr{O}(\Delta t^2)$.

To illustrate the order of convergence in space, the mesh is refined while the time step is fixed at $\Delta t = 10^{-5}$. For the first test, we consider the constant permittivity tensor (5.1). The source terms P, Q and R that complete (5.11)–(5.13) depend on the permittivity tensor and, for this case, they are given by

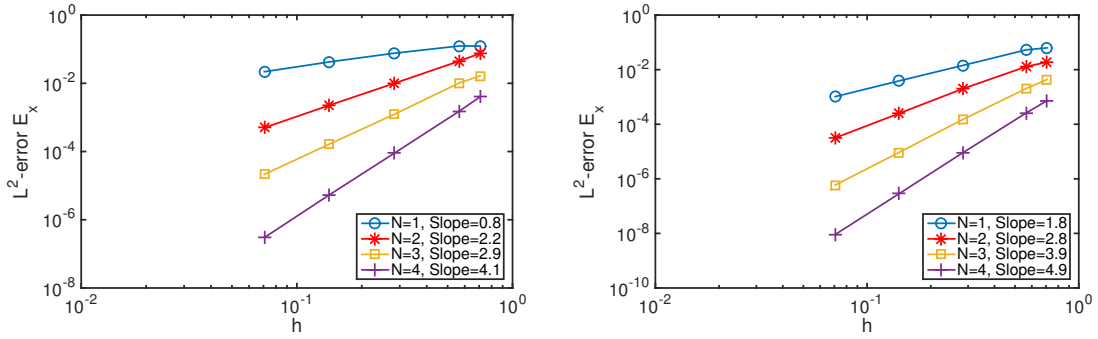
$$P(x,y,t) = \frac{\varepsilon_{xy}}{\varepsilon_{yy}}\pi\sin(\pi x)\cos(\pi y)\cos(\omega t), \tag{5.19}$$

$$Q(x,y,t) = -\frac{\varepsilon_{yx}}{\varepsilon_{xx}}\pi\cos(\pi x)\sin(\pi y)\cos(\omega t), \tag{5.20}$$

$$R(x,y,t) = 0. \tag{5.21}$$

In Table 5.10 we present the L^2-errors (5.14)–(5.16) as well as the spatial order of convergence computed according to (5.17). The results were obtained while refining the mesh according to Table 5.1, the degree of the polynomial approximation varies from $N = 1$ to $N = 4$ and considering both central and upwind fluxes. In Figure 5.3 we plot the discrete L^2-error of the $\tilde{E}_x$ component of electric field (5.14) for the same set of parameters given in Table 5.10. We plot of the error depending on the maximum element diameter for each mesh, where both the vertical and horizontal axis are scaled logarithmically. The numerical order of convergence is approximated by the slope of the linear regression line.

As we may see, for central flux the numerical convergence rate is close to the value estimated in Theorem 4.1.5, $\mathscr{O}(h^N)$, while for upwind flux we observe higher order of convergence, up to $\mathscr{O}(h^{N+1})$ in some cases.



(a) Central flux. (b) Upwind flux.

Fig. 5.3 L^2-error for $\tilde{E}_x$ (5.14) *versus* h, for constant permittivity tensor (5.1) and PEC boundary conditions.

We now consider the case were the permittivity tensor is space-dependent and given by (5.2). The source terms P, Q and R in (5.11)–(5.13) are given by

$$P(x,y,t) = \frac{\varepsilon_{xy}}{\varepsilon_{yy}}\pi\sin(\pi x)\cos(\pi y)\cos(\omega t) + \frac{\partial\omega}{\partial y}t\cos(\pi x)\cos(\pi y)\sin(\omega t), \tag{5.22}$$

$$Q(x,y,t) = -\frac{\varepsilon_{xy}}{\varepsilon_{xx}}\pi\cos(\pi x)\sin(\pi y)\cos(\omega t) - \frac{\partial\omega}{\partial x}t\cos(\pi x)\cos(\pi y)\sin(\omega t), \tag{5.23}$$

$$\begin{aligned} R(x,y,t) = {} & \frac{\pi t}{\omega\varepsilon_{yy}}\frac{\partial\omega}{\partial x}\sin(\pi x)\cos(\pi y)\cos(\omega t) - \frac{\pi}{\omega^2\varepsilon_{yy}^2}\left(\frac{\partial\omega}{\partial x}\varepsilon_{yy} + \omega\frac{\partial\varepsilon_{yy}}{\partial x}\right)\sin(\pi x)\cos(\pi y)\sin(\omega t) \\ & + \frac{\pi t}{\omega\varepsilon_{xx}}\frac{\partial\omega}{\partial y}\cos(\pi x)\sin(\pi y)\cos(\omega t) - \frac{\pi}{\omega^2\varepsilon_{xx}^2}\left(\varepsilon_{xx}\frac{\partial\omega}{\partial y} + \omega\frac{\partial\varepsilon_{xx}}{\partial y}\right)\cos(\pi x)\sin(\pi y)\sin(\omega t). \end{aligned} \tag{5.24}$$

The source terms are changed when compared to the previous case due the space dependency of tensor elements and ω given by (5.6). Note that the source terms (5.19)–(5.21) for constant permittivity tensor are special cases of (5.22)–(5.24).

The same set of parameters used for the constant tensor case are consider and the experiments repeated. The collected data with this permittivity tensor is summarised in Table 5.11. The results for the spatial convergence are plotted in Figure 5.4. As for the previous test, for the central flux, the order of convergence is near $\mathscr{O}(h^N)$, and for upwind flux we observe higher order.

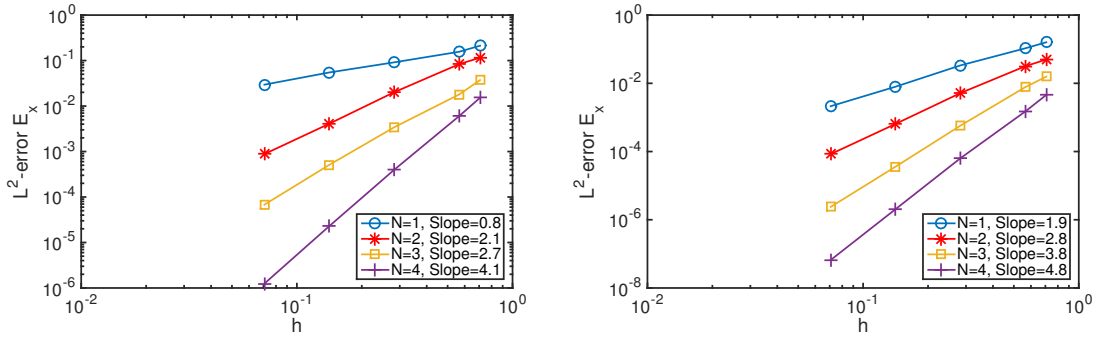
To visualize the convergence in time, the polynomials degree and the number of elements have been set to $N = 8$ and $K = 800$, respectively. The L^2-errors of the electromagnetic fields (5.14)–(5.16) are computed while decreasing the time step. The collected data is summarised in Table 5.12 and Table 5.13 for the constant and space-dependent tensors respectively. The results plotted in Figure 5.5 illustrate the second order of convergency in time for PEC boundary condition established by Theorem 4.1.5. These results correspond to upwind flux and similar results are observed for central flux.

	N	K	h	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Central flux	1	32	7.07E-01	1.23E-01		2.29E-01		3.57E-01	
		50	5.66E-01	1.24E-01	-0.02	1.65E-01	1.46	2.13E-01	2.32
		200	2.83E-01	7.63E-02	0.70	8.58E-02	0.94	5.29E-02	2.01
		800	1.41E-01	4.24E-02	0.85	4.90E-02	0.81	2.03E-02	1.38
		3200	7.07E-02	2.19E-02	0.95	2.54E-02	0.95	4.04E-03	2.33
	2	32	7.07E-01	7.53E-02		8.62E-02		9.63E-02	
		50	5.66E-01	4.48E-02	2.33	6.15E-02	1.52	4.60E-02	3.31
		200	2.83E-01	9.98E-03	2.17	1.17E-02	2.39	8.19E-03	2.49
		800	1.41E-01	2.27E-03	2.14	2.45E-03	2.26	1.07E-03	2.94
		3200	7.07E-02	5.00E-04	2.18	5.43E-04	2.17	1.23E-04	3.12
	3	32	7.07E-01	1.65E-02		1.95E-02		2.11E-02	
		50	5.66E-01	1.02E-02	2.16	1.06E-02	2.71	8.99E-03	3.81
		200	2.83E-01	1.23E-03	3.05	1.45E-03	2.87	5.24E-04	4.10
		800	1.41E-01	1.65E-04	2.90	1.94E-04	2.91	2.11E-05	4.64
		3200	7.07E-02	2.14E-05	2.94	2.53E-05	2.94	1.96E-06	3.43
	4	32	7.07E-01	4.01E-03		4.72E-03		2.43E-03	
		50	5.66E-01	1.52E-03	4.36	1.92E-03	4.02	9.88E-04	4.03
		200	2.83E-01	8.91E-05	4.09	1.06E-04	4.19	3.41E-05	4.86
		800	1.41E-01	5.37E-06	4.05	6.09E-06	4.12	1.05E-06	5.02
		3200	7.07E-02	3.01E-07	4.16	3.44E-07	4.15	3.66E-08	4.84
Upwind flux	1	32	7.07E-01	6.27E-02		1.24E-01		3.18E-01	
		50	5.66E-01	5.34E-02	0.72	8.69E-02	1.58	2.04E-01	1.98
		200	2.83E-01	1.42E-02	1.91	2.32E-02	1.90	4.85E-02	2.07
		800	1.41E-01	3.87E-03	1.88	5.68E-03	2.03	1.10E-02	2.14
		3200	7.07E-02	1.03E-03	1.91	1.53E-03	1.89	2.60E-03	2.08
	2	32	7.07E-01	1.89E-02		3.71E-02		6.61E-02	
		50	5.66E-01	1.30E-02	1.68	2.25E-02	2.25	3.43E-02	2.94
		200	2.83E-01	1.99E-03	2.71	3.06E-03	2.88	4.66E-03	2.88
		800	1.41E-01	2.48E-04	3.00	3.83E-04	3.00	6.05E-04	2.95
		3200	7.07E-02	3.16E-05	2.98	4.85E-05	2.98	7.74E-05	2.97
	3	32	7.07E-01	4.35E-03		7.59E-03		1.09E-02	
		50	5.66E-01	2.05E-03	3.37	3.60E-03	3.35	5.13E-03	3.37
		200	2.83E-01	1.49E-04	3.79	2.14E-04	4.07	3.33E-04	3.95
		800	1.41E-01	9.23E-06	4.01	1.39E-05	3.95	2.14E-05	3.96
		3200	7.07E-02	5.91E-07	3.97	9.11E-07	3.93	1.35E-06	3.98
	4	32	7.07E-01	7.28E-04		1.16E-03		1.71E-03	
		50	5.66E-01	2.62E-04	4.58	4.43E-04	4.32	6.44E-04	4.36
		200	2.83E-01	8.96E-06	4.87	1.45E-05	4.93	2.07E-05	4.96
		800	1.41E-01	2.83E-07	4.98	4.58E-07	4.98	6.56E-07	4.98
		3200	7.07E-02	9.09E-09	4.96	1.44E-08	4.99	2.11E-08	4.96

Table 5.10 The L^2-errors (5.14)–(5.16) and the spatial order for PEC boundary condition and constant permittivity tensor (5.1).

	N	K	h	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Central flux	1	32	7.07E-01	2.12E-01		2.05E-01		4.60E-01	
		50	5.66E-01	1.56E-01	1.37	1.82E-01	0.55	2.79E-01	2.23
		200	2.83E-01	9.15E-02	0.77	9.55E-02	0.93	8.94E-02	1.64
		800	1.41E-01	5.48E-02	0.74	5.56E-02	0.78	2.29E-02	1.97
		3200	7.07E-02	2.95E-02	0.89	2.91E-02	0.93	5.49E-03	2.06
	2	32	7.07E-01	1.16E-01		1.36E-01		1.44E-01	
		50	5.66E-01	8.35E-02	1.47	8.07E-02	2.33	8.56E-02	2.33
		200	2.83E-01	2.02E-02	2.05	2.00E-02	2.01	1.07E-02	3.00
		800	1.41E-01	4.09E-03	2.30	4.03E-03	2.31	1.40E-03	2.93
		3200	7.07E-02	8.80E-04	2.22	8.58E-04	2.23	1.79E-04	2.97
	3	32	7.07E-01	3.73E-02		3.77E-02		2.45E-02	
		50	5.66E-01	1.80E-02	3.27	1.94E-02	2.97	1.07E-02	3.69
		200	2.83E-01	3.42E-03	2.39	3.41E-03	2.51	7.46E-04	3.85
		800	1.41E-01	5.09E-04	2.75	4.73E-04	2.85	4.89E-05	3.93
		3200	7.07E-02	6.79E-05	2.91	6.05E-05	2.97	3.06E-06	4.00
	4	32	7.07E-01	1.52E-02		1.38E-02		4.92E-03	
		50	5.66E-01	6.02E-03	4.17	5.29E-03	4.30	2.10E-03	3.81
		200	2.83E-01	3.98E-04	3.92	3.68E-04	3.85	9.10E-05	4.53
		800	1.41E-01	2.30E-05	4.11	2.12E-05	4.12	2.76E-06	5.04
		3200	7.07E-02	1.25E-06	4.20	1.17E-06	4.18	8.89E-08	4.96
Upwind flux	1	32	7.07E-01	1.60E-01		1.13E-01		3.73E-01	
		50	5.66E-01	1.06E-01	1.81	1.11E-01	0.05	2.46E-01	1.87
		200	2.83E-01	3.35E-02	1.67	3.57E-02	1.64	5.86E-02	2.07
		800	1.41E-01	7.96E-03	2.08	8.79E-03	2.02	1.25E-02	2.23
		3200	7.07E-02	2.13E-03	1.90	2.27E-03	1.95	3.05E-03	2.04
	2	32	7.07E-01	4.99E-02		5.26E-02		8.08E-02	
		50	5.66E-01	3.18E-02	2.02	3.34E-02	2.03	4.78E-02	2.35
		200	2.83E-01	5.18E-03	2.62	5.35E-03	2.64	7.30E-03	2.71
		800	1.41E-01	6.46E-04	3.00	6.92E-04	2.95	9.66E-04	2.92
		3200	7.07E-02	8.38E-05	2.95	8.87E-05	2.96	1.22E-04	2.98
	3	32	7.07E-01	1.57E-02		1.88E-02		2.28E-02	
		50	5.66E-01	7.76E-03	3.17	7.46E-03	4.15	8.99E-03	4.18
		200	2.83E-01	5.75E-04	3.75	5.82E-04	3.68	6.93E-04	3.70
		800	1.41E-01	3.60E-05	4.00	3.68E-05	3.98	4.43E-05	3.97
		3200	7.07E-02	2.38E-06	3.92	2.39E-06	3.95	2.82E-06	3.97
	4	32	7.07E-01	4.58E-03		3.61E-03		4.60E-03	
		50	5.66E-01	1.50E-03	4.99	1.53E-03	3.83	1.73E-03	4.39
		200	2.83E-01	6.39E-05	4.56	6.11E-05	4.65	7.16E-05	4.59
		800	1.41E-01	2.05E-06	4.96	1.95E-06	4.97	2.30E-06	4.96
		3200	7.07E-02	6.59E-08	4.96	6.27E-08	4.96	7.43E-08	4.96

Table 5.11 The L^2-errors (5.14)–(5.16) and the spatial order for PEC boundary condition and space-dependent permittivity tensor (5.2).

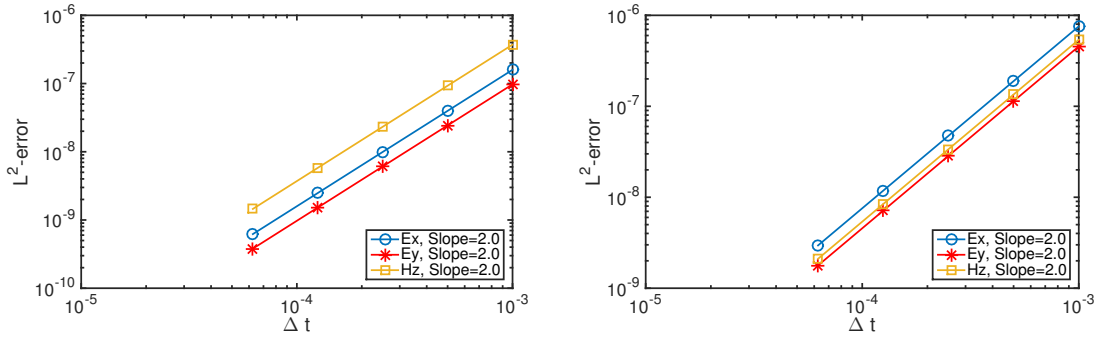


(a) Central flux. (b) Upwind flux.

Fig. 5.4 L^2-error for $\tilde{E}_x$ (5.14) *versus* h, for PEC boundary conditions and space-dependent permittivity tensor (5.2).

Δt	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
1.00E-03	1.59E-07		9.73E-08		3.72E-07	
5.00E-04	3.98E-08	2.00	2.43E-08	2.00	9.29E-08	2.00
2.50E-04	9.94E-09	2.00	6.08E-09	2.00	2.32E-08	2.00
1.25E-04	2.48E-09	2.00	1.52E-09	2.00	5.81E-09	2.00
6.25E-05	6.21E-10	2.00	3.80E-10	2.00	1.45E-09	2.00

Table 5.12 The L^2-errors (5.14)–(5.16) and the temporal order for PEC boundary condition and constant permittivity tensor (5.1).



(a) Constant permittivity tensor. (b) Space-dependente permittivity tensor.

Fig. 5.5 L^2-errors (5.14)–(5.16) *versus* Δt for PEC boundary conditions and upwind flux.

Δt	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
1.00E-03	7.56E-07		4.56E-07		5.40E-07	
5.00E-04	1.89E-07	2.00	1.14E-07	2.00	1.35E-07	2.00
2.50E-04	4.73E-08	2.00	2.85E-08	2.00	3.37E-08	2.00
1.25E-04	1.18E-08	2.00	7.12E-09	2.00	8.42E-09	2.00
6.25E-05	2.95E-09	2.00	1.78E-09	2.00	2.10E-09	2.00

Table 5.13 The L^2-error (5.14)–(5.16) and the temporal order for PEC boundary condition and space-dependent permittivity tensor (5.2).

5.3.2 Silver-Müller absorbing boundary condition

We now consider the test problem (5.11)–(5.13) with SM-ABC and initial conditions (5.3)–(5.5). The source terms P, Q and R are obtained such that the problem has the exact solution

$$\begin{aligned} E_x(x,y,t) &= -\sqrt{\frac{\varepsilon_{yy}}{\det(\varepsilon)}}\sin(\pi t)\sin(\pi x),\\ E_y(x,y,t) &= \sqrt{\frac{\varepsilon_{xx}}{\det(\varepsilon)}}\sin(\pi t)\sin(\pi y),\\ H_z(x,y,t) &= \sin(\pi t)\sin(\pi xy). \end{aligned}$$

For this type of boundary conditions, the theoretical convergence analysis established by Theorem 4.1.5 showed that the convergence order of the leap-frog DG scheme (2.32)–(2.34) in space and time is $\mathscr{O}(h^N)+\mathscr{O}(\Delta t)$.

The set of experiments is the same as in the case of PEC boundary conditions. The mesh is refined according to Table 5.1 for different degrees for the polynomial approximation from $N=1$ to $N=4$ and both central and upwind flux. To illustrate the order of convergence in space, the mesh is refined while the time step is fixed at $\Delta t = 10^{-5}$, except while the degree for the polynomial approximation is $N=4$, where we consider $\Delta t = 10^{-6}$.

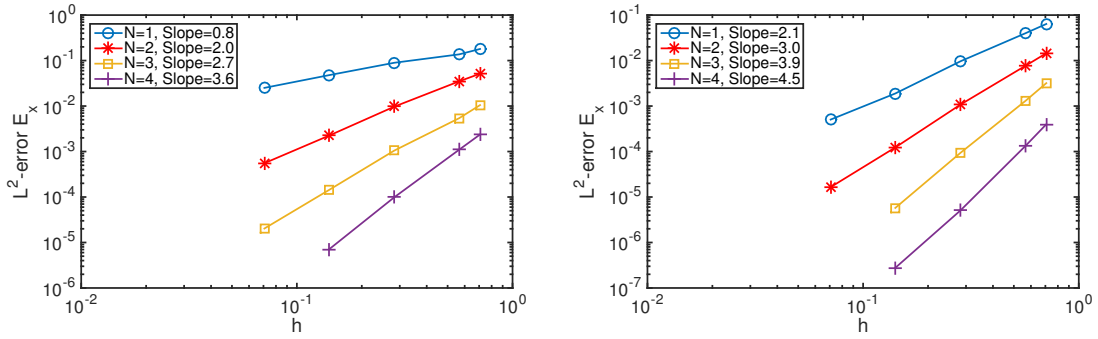
We first consider the constant permittivity tensor (5.1). For this case, the source terms in (5.11)–(5.13) are given by

$$P(x,y,t) = -\pi\varepsilon_{xx}\sqrt{\frac{\varepsilon_{yy}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi x) + \pi\varepsilon_{xy}\sqrt{\frac{\varepsilon_{xx}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi y) - \pi x\sin(\pi t)\cos(\pi xy), \tag{5.25}$$

$$Q(x,y,t) = -\pi\varepsilon_{yx}\sqrt{\frac{\varepsilon_{yy}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi x) + \pi\varepsilon_{yy}\sqrt{\frac{\varepsilon_{xx}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi y) + \pi y\sin(\pi t)\cos(\pi xy), \tag{5.26}$$

$$R(x,y,t) = \pi\mu\cos(\pi t)\sin(\pi xy). \tag{5.27}$$

In Table 5.14 we present the L^2-errors (5.14)–(5.16) as well as the spatial order of convergence computed according to (5.17). In Figure 5.6 we plot the discrete L^2-error of the $\tilde{E}_x$ component of electric field (5.14) for the same set of parameters given in Table 5.14. We plot of the error depending on the maximum element diameter for each mesh, where both the vertical and horizontal axis are



(a) Central flux. (b) Upwind flux.

Fig. 5.6 L^2-error for $\tilde{E}_x$ (5.14) *versus* h, for SM-ABC and constant permittivity tensor (5.1).

scaled logarithmically. The numerical order of convergence is approximated by the slope of the linear regression line.

As we may see, for central flux the numerical convergence rate is close to the value estimated in Theorem 4.1.5, $\mathscr{O}(h^N)$, while for upwind flux we observe higher order of convergence, up to $\mathscr{O}(h^{N+1})$ in some cases.

We now consider the case where permittivity tensor is space-dependent and given by (5.2). The source terms P, Q and R in (5.11)–(5.13) are changed due to space dependency of the tensor elements and are given by

$$\begin{aligned} P(x,y,t) &= -\pi\varepsilon_{xx}\sqrt{\frac{\varepsilon_{yy}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi x) + \pi\varepsilon_{xy}\sqrt{\frac{\varepsilon_{xx}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi y) \\ &\quad - \pi x\sin(\pi t)\cos(\pi xy), \end{aligned} \tag{5.28}$$

$$\begin{aligned} Q(x,y,t) &= -\pi\varepsilon_{xy}\sqrt{\frac{\varepsilon_{yy}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi x) + \pi\varepsilon_{yy}\sqrt{\frac{\varepsilon_{xx}}{\det(\varepsilon)}}\cos(\pi t)\sin(\pi y) \\ &\quad + \pi y\sin(\pi t)\cos(\pi xy), \end{aligned} \tag{5.29}$$

$$\begin{aligned} R(x,y,t) &= \pi\mu\cos(\pi t)\sin(\pi xy) + \frac{\frac{\partial\varepsilon_{xx}}{\partial y}\det(\varepsilon) + \frac{\partial\det(\varepsilon)}{\partial y}\varepsilon_{xx}}{2\det(\varepsilon)^2}\sqrt{\frac{\det(\varepsilon)}{\varepsilon_{xx}}}\sin(\pi t)\sin(\pi y) \\ &\quad - \frac{\frac{\partial\varepsilon_{yy}}{\partial x}\det(\varepsilon) + \frac{\partial\det(\varepsilon)}{\partial x}\varepsilon_{yy}}{2\det(\varepsilon)^2}\sqrt{\frac{\det(\varepsilon)}{\varepsilon_{xx}}}\sin(\pi t)\sin(\pi x). \end{aligned} \tag{5.30}$$

The L^2-error sand the order of convergence in space are computed as in the previous case. All these data is collected and summarised in Table 5.15 and illustrated in Figure 5.7. As for the constant tensor case, we conclude that, for central flux the numerical convergence rate is close to the value estimated in Theorem 4.1.5, $\mathscr{O}(h^N)$, while for upwind flux we observe higher order of convergence, up to $\mathscr{O}(h^{N+1})$ in some cases.

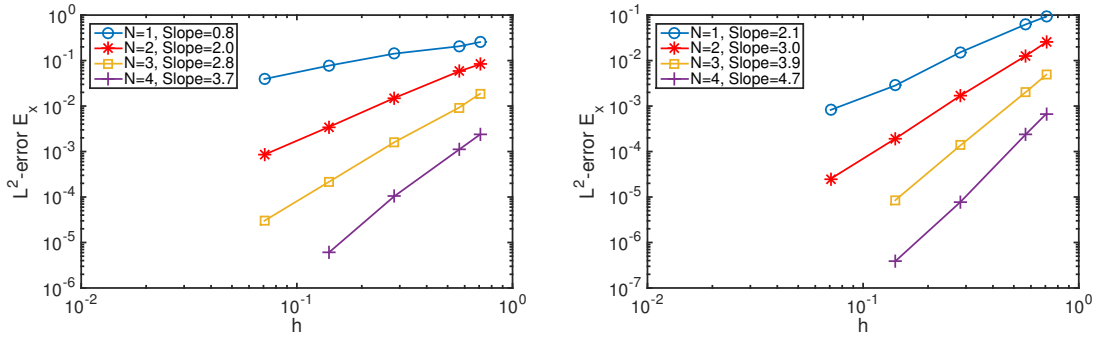
To visualize the convergence in time, the approximation polynomial degree and the number of elements have been set to $N = 8$ and $K = 800$, respectively. For SM-ABC the data plotted in Figure 5.8 illustrates the first order of convergency established by Theorem 4.1.5.

	N	K	h	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Central flux	1	32	7.07E-01	1.80E-01		2.20E-01		2.51E-01	
		50	5.66E-01	1.37E-01	1.23	1.59E-01	1.47	1.64E-01	1.91
		200	2.83E-01	8.95E-02	0.61	1.02E-01	0.63	2.40E-02	2.77
		8000	1.41E-01	4.78E-02	0.91	5.49E-02	0.90	3.80E-03	2.66
		3200	7.07E-02	2.51E-02	0.93	2.91E-02	0.92	7.67E-04	2.31
	2	32	7.07E-01	5.15E-02		5.80E-02		5.12E-02	
		50	5.66E-01	3.50E-02	1.74	3.92E-02	1.75	2.24E-02	3.70
		200	2.83E-01	9.71E-03	1.85	1.05E-02	1.90	1.45E-03	3.95
		800	1.41E-01	2.24E-03	2.12	2.40E-03	2.13	1.37E-04	3.41
		3200	7.07E-02	5.54E-04	2.01	5.89E-04	2.03	6.47E-06	4.40
	3	32	7.07E-01	1.05E-02		1.20E-02		3.77E-03	
		50	5.66E-01	5.42E-03	2.95	6.24E-03	2.93	1.83E-03	3.24
		200	2.83E-01	1.05E-03	2.37	1.20E-03	2.37	7.85E-05	4.55
		800	1.41E-01	1.43E-04	2.88	1.67E-04	2.85	4.46E-06	4.14
		3200	7.07E-02	2.03E-05	2.81	2.42E-05	2.79	7.24E-07	2.63
	4	32	7.07E-01	2.37E-03		2.85E-03		8.78E-04	
		50	5.66E-01	1.10E-03	3.42	1.31E-03	3.50	2.11E-04	6.40
		200	2.83E-01	9.88E-05	3.48	1.11E-04	3.55	6.25E-06	5.07
		800	1.41E-01	7.07E-06	3.80	7.83E-06	3.83	9.42E-07	2.73
		3200	7.07E-02	7.43E-07	3.25	8.57E-07	3.19	7.97E-07	0.24
Upwind flux	1	32	7.07E-01	6.38E-02		7.57E-02		1.60E-01	
		50	5.66E-01	3.99E-02	2.11	4.95E-02	1.90	8.80E-02	2.68
		200	2.83E-01	9.66E-03	2.05	1.23E-02	2.01	1.70E-02	2.37
		800	1.41E-01	1.89E-03	2.35	2.52E-03	2.28	2.52E-03	2.75
		3200	7.07E-02	5.04E-04	1.91	6.15E-04	2.04	5.16E-04	2.29
	2	32	7.07E-01	1.45E-02		1.85E-02		2.77E-02	
		50	5.66E-01	7.68E-03	2.86	9.21E-03	3.13	7.58E-03	5.81
		200	2.83E-01	1.08E-03	2.83	1.45E-03	2.67	4.94E-04	3.94
		800	1.41E-01	1.23E-04	3.13	1.79E-04	3.01	2.56E-05	4.27
		3200	7.07E-02	1.65E-05	2.90	2.41E-05	2.90	2.50E-06	3.35
	3	32	7.07E-01	3.19E-03		4.17E-03		3.52E-03	
		50	5.66E-01	1.30E-03	4.01	1.69E-03	4.05	9.96E-04	5.66
		200	2.83E-01	9.39E-05	3.79	1.26E-04	3.75	4.21E-05	4.56
		800	1.41E-01	5.75E-06	4.03	7.74E-06	4.03	3.24E-06	3.70
		3200	7.07E-02	1.71E-06	1.75	2.07E-06	1.90	2.05E-06	0.66
	4	32	7.07E-01	3.97E-04		5.10E-04		4.90E-04	
		50	5.66E-01	1.33E-04	4.90	1.74E-04	4.83	1.09E-04	6.72
		200	2.83E-01	5.09E-06	4.71	6.99E-06	4.64	1.91E-06	5.84
		800	1.41E-01	2.77E-07	4.20	3.59E-07	4.28	2.85E-07	2.75

Table 5.14 The L^2-errors (5.14)–(5.16) and the spatial order for SM-ABC and constant permittivity tensor (5.1).

	N	K	h	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Central flux	1	32	7.07E-01	2.55E-01		2.61E-01		2.54E-01	
		50	5.66E-01	2.06E-01	0.96	2.18E-01	0.82	1.51E-01	2.33
		200	2.83E-01	1.42E-01	0.54	1.46E-01	0.58	2.38E-02	2.66
		800	1.41E-01	7.75E-02	0.87	8.00E-02	0.86	5.00E-03	2.25
		3200	7.07E-02	3.93E-02	0.98	4.17E-02	0.94	1.15E-03	2.12
	2	32	7.07E-01	8.28E-02		8.16E-02		4.53E-02	
		50	5.66E-01	5.81E-02	1.58	5.58E-02	1.70	1.68E-02	4.44
		200	2.83E-01	1.50E-02	1.96	1.50E-02	1.90	1.04E-03	4.01
		800	1.41E-01	3.48E-03	2.10	3.50E-03	2.10	1.08E-04	3.26
		3200	7.07E-02	8.52E-04	2.03	8.49E-04	2.04	4.83E-06	4.49
	3	32	7.07E-01	1.86E-02		1.87E-02		2.94E-03	
		50	5.66E-01	9.28E-03	3.11	9.16E-03	3.20	1.20E-03	4.03
		200	2.83E-01	1.60E-03	2.54	1.65E-03	2.47	6.07E-05	4.30
		800	1.41E-01	2.16E-04	2.89	2.26E-04	2.87	3.64E-06	4.06
		3200	7.07E-02	3.00E-05	2.85	3.20E-05	2.82	6.55E-07	2.48
	4	32	7.07E-01	2.37E-03		2.85E-03		8.78E-04	
		50	5.66E-01	1.11E-03	3.40	1.28E-03	3.59	2.23E-04	6.14
		200	2.83E-01	1.04E-04	3.41	1.19E-04	3.42	3.81E-06	5.87
		800	1.41E-01	6.06E-06	4.11	6.84E-06	4.13	2.55E-07	3.90
Upwind flux	1	32	7.07E-01	9.29E-02		9.86E-02		1.48E-01	
		50	5.66E-01	6.20E-02	1.81	6.60E-02	1.80	7.97E-02	2.78
		200	2.83E-01	1.53E-02	2.02	1.62E-02	2.02	1.52E-02	2.39
		800	1.41E-01	2.83E-03	2.44	3.40E-03	2.26	2.43E-03	2.65
		3200	7.07E-02	8.21E-04	1.78	8.88E-04	1.94	4.71E-04	2.36
	2	32	7.07E-01	2.54E-02		2.34E-02		1.52E-02	
		50	5.66E-01	1.26E-02	3.15	1.16E-02	3.14	6.01E-03	4.16
		200	2.83E-01	1.70E-03	2.89	1.82E-03	2.67	4.47E-04	3.75
		800	1.41E-01	1.92E-04	3.15	2.16E-04	3.08	2.35E-05	4.25
		3200	7.07E-02	2.51E-05	2.93	2.88E-05	2.90	2.48E-06	3.24
	3	32	7.07E-01	4.87E-03		5.33E-03		2.64E-03	
		50	5.66E-01	2.00E-03	3.99	2.11E-03	4.14	9.01E-04	4.82
		200	2.83E-01	1.37E-04	3.87	1.57E-04	3.75	3.97E-05	4.50
		800	1.41E-01	8.49E-06	4.01	9.41E-06	4.06	3.18E-06	3.64
		3200	7.07E-02	2.28E-06	1.90	2.25E-06	2.07	2.07E-06	0.62
	4	32	7.07E-01	6.75E-04		6.60E-04		4.07E-04	
		50	5.66E-01	2.35E-04	4.73	2.35E-04	4.62	1.06E-04	6.03
		200	2.83E-01	7.82E-06	4.91	8.86E-06	4.73	1.81E-06	5.87
		800	1.41E-01	3.87E-07	4.34	4.08E-07	4.44	2.83E-07	2.67

Table 5.15 The L^2-errors (5.14)–(5.16) and the spatial order for SM-ABC and space-dependent permittivity tensor (5.2).

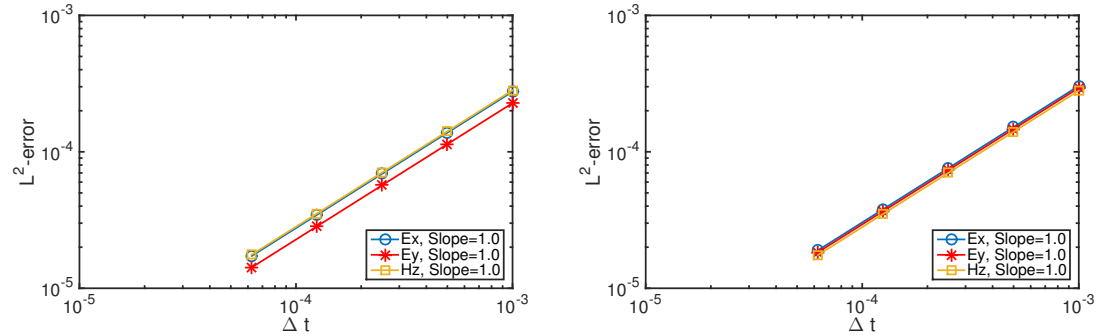


(a) Central flux. (b) Upwind flux.

Fig. 5.7 L^2-error for $\tilde{E}_x$ (5.14) *versus* h, for SM-ABC and space dependent permittivity tensor (5.2).

Δt	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Leap-frog time integrator						
1.00E-03	2.76E-04		2.28E-04		2.82E-04	
5.00E-04	1.38E-04	1.00	1.14E-04	1.00	1.41E-04	1.00
2.50E-04	6.90E-05	1.00	5.69E-05	1.00	7.06E-05	1.00
1.25E-04	3.45E-05	1.00	2.84E-05	1.00	3.53E-05	1.00
6.25E-05	1.73E-05	1.00	1.42E-05	1.00	1.77E-05	1.00
Predictor-corrector time integrator						
1.00E-03	4.18E-05		2.73E-05		4.76E-05	
5.00E-04	9.71E-06	2.11	6.46E-06	2.08	1.06E-05	2.16
2.50E-04	2.35E-06	2.05	1.57E-06	2.04	2.54E-06	2.07
1.25E-04	5.77E-07	2.02	3.88E-07	2.02	6.21E-07	2.03
6.25E-05	1.43E-07	2.01	9.64E-08	2.01	1.54E-07	2.02

Table 5.16 The L^2-errors (5.14)–(5.16) and the temporal order for SM-ABC and constant permittivity tensor (5.1) when the predictor-corrector method is considered.



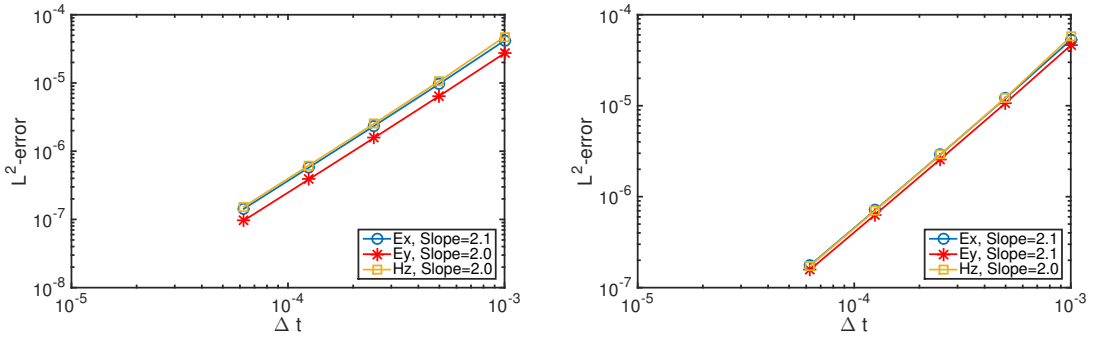
(a) Constant permittivity tensor. (b) Space-dependent permittivity tensor.

Fig. 5.8 L^2-errors (5.14)–(5.16) *versus* Δt, for SM-ABC and upwind flux.

Δt	Error $\tilde{E}_x$	Order	Error $\tilde{E}_y$	Order	Error $\tilde{H}_z$	Order
Leap-frog time integrator						
1.00E-03	3.04E-04		2.94E-04		2.81E-04	
5.00E-04	1.52E-04	1.00	1.47E-04	1.00	1.40E-04	1.00
2.50E-04	7.60E-05	1.00	7.35E-05	1.00	7.01E-05	1.00
1.25E-04	3.80E-05	1.00	3.67E-05	1.00	3.51E-05	1.00
6.25E-05	1.90E-05	1.00	1.84E-05	1.00	1.75E-05	1.00
Predictor-corrector time integrator						
1.00E-03	5.28E-05		4.62E-05		5.74E-05	
5.00E-04	1.21E-05	2.13	1.07E-05	2.11	1.23E-05	2.23
2.50E-04	2.91E-06	2.06	2.59E-06	2.05	2.90E-06	2.08
1.25E-04	7.13E-07	2.03	6.36E-07	2.02	7.05E-07	2.04
6.25E-05	1.76E-07	2.01	1.58E-07	2.01	1.74E-07	2.02

Table 5.17 The L^2-errors (5.14)–(5.16) and the temporal order for SM-ABC and space-dependent permittivity tensor (5.2) when the predictor-corrector method is considered.

As we seen theoretically and numerically, when we consider the SM-ABC for our fully explicit DG scheme (2.32)–(2.34), the temporal order of convergence is not what we expected from leap-frog time integrator. The temporal order reduces from two to one. In order to recover the order in time and stay explicit in defining the flux, we proposed an iterative predictor-corrector time integration scheme (4.35)–(4.37). As we may see in Table 5.16 and Figure 5.9 the second temporal order is perceived in our results.



(a) Constant permittivity tensor. (b) Space-dependent permittivity tensor.

Fig. 5.9 L^2-errors (5.14)–(5.16) *versus* Δt, for SM-ABC and upwind flux when the predictor-corrector method is considered.

5.4 Modeling scattered electromagnetic wave's propagation through eye's structures

As mentioned in the Introduction, the research that lead to this dissertation was developed in the framework of a more general project that aims to achieve a computational model to simulate the

electromagnetic wave's propagation through the eye's structures in order to create a virtual optical coherence tomography (OCT) scan [73].

As OCT standard techniques only provide structural information [74], it is necessary to expand OCT data analysis to account for both structural and functional information. A mathematical modeling of OCT data could expand the information provides by OCT from structural to functional information. The functional information may provide a means for optical biopsy. For instance, the variation of retinal nerve fiber layer (RNFL) birefringence may provide early detection of subcellular changes in glaucoma and other diseases affecting the optic nerve [75].

Simulating the full complexity of the retina, in particular the variation of the size and shape of each structure, distance between them and the respective refractive indexes, requires a rigorous approach that can be achieved by solving Maxwell's equations. As the interest is to acquire the backscattered light intensity, we start this section by the scattered field formulation. Then we build up a two dimensional model which tries to represent a single nucleus of the outer nuclear layer (ONL) of the retina. The efficiency of our method is examined by simulating the light scattering in this 2D domain. The evolution of scattering field intensity in time is obtained using the predictor-corrector DG method.

5.4.1 The scattered field formulation

We exploit the linearity of the Maxwell's equations (5.11)-(5.13) in order to separating the electromagnetic fields (E, H) into incident fields (E^i, H^i) and scattered components (E^s, H^s), *i.e.*,

$$E = E^s + E^i \qquad \text{and} \qquad H = H^s + H^i. \tag{5.31}$$

Assuming that the incident field is also a solution of the Maxwell's equations we obtain in the same way as in [82], the scattered field formulation,

$$\varepsilon_{xx}\frac{\partial E_x^s}{\partial t} + \varepsilon_{xy}\frac{\partial E_y^s}{\partial t} = \frac{\partial H_z^s}{\partial y} + P \tag{5.32}$$

$$\varepsilon_{yx}\frac{\partial E_x^s}{\partial t} + \varepsilon_{yy}\frac{\partial E_y^s}{\partial t} = -\frac{\partial H_z^s}{\partial x} + Q \tag{5.33}$$

$$\mu\frac{\partial H_z^s}{\partial t} = -\frac{\partial E_y^s}{\partial x} + \frac{\partial E_x^s}{\partial y} + R \qquad \text{in} \quad \Omega\times(0,T], \tag{5.34}$$

with the source terms

$$P(x,y,t) = (\varepsilon^i - \varepsilon_{xx})\frac{\partial E_x^i}{\partial t} - \varepsilon_{xy}\frac{\partial E_y^i}{\partial t}, \tag{5.35}$$

$$Q(x,y,t) = -\varepsilon_{yx}\frac{\partial E_x^i}{\partial t} + (\varepsilon^i - \varepsilon_{yy})\frac{\partial E_y^i}{\partial t}, \tag{5.36}$$

$$R(x,y,t) = (\mu^i - \mu)\frac{\partial H_z^i}{\partial t}, \tag{5.37}$$

where ε^i and μ^i represent, respectively, the relative permittivity and permeability of the medium in which the incident field propagates in the absence of scatterers (in the background medium).

Additionally, using this formulation it is very easy to specify an incident wave using an analytic formula.

5.4.2 Light scattering in outer nuclear layer

In order to apply our method on a real model problem the outer nuclear layer is chosen among the retina's layers. This layer was chosen as it consistently presents the characteristics of diabetic macular edema [21] and because spherical scatterers can adequately model it, which helps to simplify the simulation.

The outer nuclear layer is mostly populated by the cells bodies of light sensitive photoreceptor cells (rods and cons). The nucleus is the biggest organelle in the photoreceptor cell's soma and presents a high refractive index difference to the surrounding medium. Thus the main contribution to light scattering in this layer comes from the nucleus [76]. The outer nuclear layer could be modeled as a population of spherical nuclei in an homogenous medium.

As a proof of concept we present a simple simulation in a two dimensional square domain which contains a circle that aims to represent the single nucleus in the ONL. This domain is presented in Figure 5.10 while the difference between the permittivity in the circle and the background domain is shown with ε'.

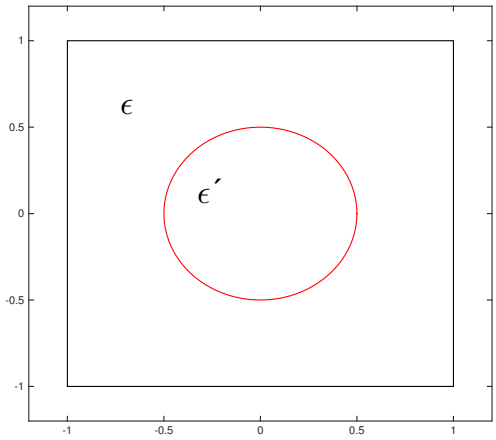


Fig. 5.10 Square computational domain which contains a circle that aims to represent the single nucleus in ONL.

Let us consider equations (5.32)–(5.33), in $\Omega = (-1,1)^2$, complemented with Silver-Müller absorbing boundary condition and null initial condition. The absorbing boundary condition is chosen for the model as the non absorbing boundary conditions provide undesirable reflections that invade the computational domain.

In the experiments the magnetic permeability and relative permittivity are considered as constants, $\varepsilon^i = 1$ and $\mu = 1$. The permittivity ε is considered as a diagonal matrix with $\varepsilon_{xx}(x,y) = \varepsilon_{yy}(x,y) = 1.2$ for (x,y) such that $\sqrt{x^2+y^2} < 0.5$ and $\varepsilon_{xx}(x,y) = \varepsilon_{yy}(x,y) = 1$ otherwise. For the incident wave we consider the planar wave $E_y^i(x,t) = \cos(10(x-t))$.

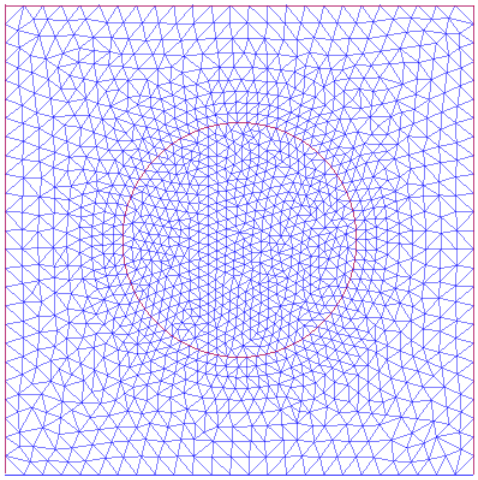
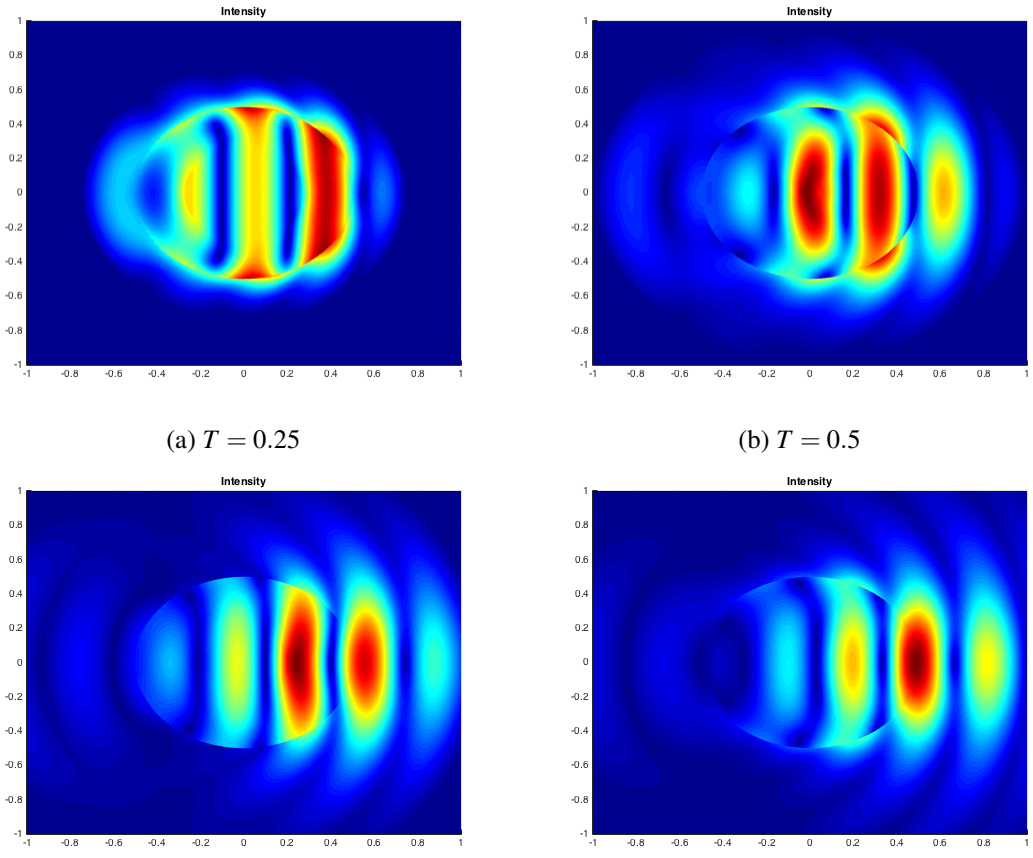


Fig. 5.11 Computational domain and triangular mesh.



(a) $T = 0.25$ (b) $T = 0.5$

(c) $T = 0.75$ (d) $T = 1$

Fig. 5.12 Evolution of the scattered field intensity $I^s = \sqrt{(E_x^s)^2 + (E_y^s)^2}$ with time.

The simulation is done with predictor-detector DG method which is more efficient in the case of SM-ABC [6]. The scattered field intensity

$$I^s = \sqrt{(E_x^s)^2 + (E_y^s)^2}$$

were obtained with the predictor-corrector DG method defined on the mesh plotted in Figure 5.11. In simulations we consider $\alpha = 0$ (central flux) and the approximation polynomial degree as $N = 4$. The time step is chosen as $\Delta t = 0.002$ and the final simulation time is $T = 1$. The evolution of scattered field intensity with time is plotted in Figure 5.12.

Chapter 6

Conclusion

The target of this dissertation was to formulate a fully explicit DGTD solution of Maxwell's equations. The nodal discontinuous Galerkin method was employed for space discretization and coupled with a leap-frog time integrator yielding a fully explicit scheme which is capable to deal with different formulations for the numerical flux, anisotropic materials and different types of boundary conditions. The proposed leap-frog discontinuous Galerkin scheme was analyzed rigorously and the numerical results supporting the achieved theoretical results were provided.

In this chapter, we first summarize the accomplishment of each part in Section 6.1 and then end with some final comments and perspectives for future work in Section 6.2.

6.1 Summary

The first part of this dissertation was devoted to the background on computational electromagnetic with Maxwell's equations. We first presented Maxwell's equations and some of their basic properties, like the constitutive relations between the electromagnetic fields, their behavior at interfaces and boundary conditions. We then introduced the leap-frog discontinuous Galerkin method which was our choice model in the numerical integration of Maxwell's equations in anisotropic materials.

In order to reduce the number of equations of the 3D model, the transverse electric mode of Maxwell's equations was considered as our 2D model. Our model included the most common boundary conditions PEC and PMC as well as the so-called first order Silver-Müller absorbing boundary condition in order to truncate the unbounded domain to a bounded domain.

The main ingredient of the DG methods is the numerical flux. Our formulation of the numerical flux had been developed in a general framework which unified different flux-evaluation schemes and included the treatment of anisotropic materials. The leap-frog time integration method had been applied to the DG semi-discrete scheme to obtain the leap-frog DG scheme. We used backward approximation in upwind flux terms which yielded a fully explicit scheme.

The second part of this work dealt with the analysis of leap-frog DG scheme. The sufficient condition of stability was firstly analyzed for the two dimensional problem and the bound of the stability region was deduced in detail which revealed the influence of the mesh size, the choice of numerical flux and the degree of the polynomials used in the construction of the finite element space

making possible to balance accuracy and computational efficiency. Moreover, the stability analysis was further extended to the three dimensional problem and the stability region was also presented.

We rigorously analyzed the convergence property of the leap-frog DG scheme in 2D. We proved that under the stability condition the scheme is arbitrary high-order in space. The error estimate analysis of the scheme demonstrated the second order of temporal convergence for PEC and PMC boundary conditions but only the first order in the case of Silver-Müller absorbing boundary condition. We proposed a predictor-corrector time integrating scheme which recovered the time order to two in the case of SM-ABC. This idea was developed further and we defined an iterative predictor-corrector time integrator. The iterative predictor-corrector DG method was defined such that the scheme remained fully explicit and converged to a second order implicit method. The stability analysis of the implicit method was also presented.

Finally, in a set of numerical experiments we provided the numerical results supporting the achieved theoretical results. The sharpness of the stability region was checked and the temporal and spatial order of convergence were confirmed for both central and upwind fluxes and different boundary conditions. Moreover, the efficiency of the predictor-corrector time integrator to recover the time order was checked. In the last part, we present the results of simulation with our method in the framework of our application of interest. The light scattering was simulated in a 2D domain which aims to represent a single nucleus in the outer nuclear layer of retina.

6.2 Outlook

Numerical analysis is a crossroad of several disciplines and it is crucial on developing efficient and accurate solutions to real-world problems, while maintaining a solid theoretical base. Realistic models are usually very intricate. Frequently they involve coupled systems of time dependent partial differential equations, whose mathematical analysis, that is often quite complex, requires sophisticated mathematical tools. To establish mathematical models to simulate the behavior and dynamics of those systems and to provide the mathematical foundations of the numerical methods, to analyze their theoretical properties, namely stability and accuracy, is then a challenging problem of paramount importance.

The perspectives of the research following the work of the present thesis, comprises contributions in the field of numerical analysis and also in the field of biomathematics.

The leap-frog discontinuous Galerkin method that was implemented and analyzed in this dissertation is an efficient method. However, the method used for the time integration has a great impact on the accuracy of the numerical solution. The leap-frog DG scheme is of arbitrary high-order convergent in space while the temporal order is restricted to two. The first straightforward future work could be on improving the accuracy of the time integration method. A possible choice would be explore the time integration by high order methods like, for instance, an high order explicit Runge-Kutta method, and study their quantitative and qualitative properties.

Other alternative could be use other less exploited potentials of the discontinuous Galerkin methods. In particular, DG methods can be used in a space-time approach, giving an effective framework for high-order accurate methods. In this technique, time is considered as an extra dimension and it is treated in the same way as the spatial coordinates. Space-time DG methods while allowing for

discontinuities in the temporal discretization, combine the well known advantages of the DG methods, such as of flexibility for local mesh refinement, adjustment of the polynomial order in each element and excellent performance on parallel computers. The main drawback of the space-time DG methods is their implicit nature.

Another direction of future research, following some preliminary results of simulation presented on this thesis, is to model the light scattering in the retina aiming to mimicking the OCT imaging system. The idea is to integrate the time-dependent Maxwell's equations to numerically solve local scattering effects within the retina, considering OCT wavelengths and different realistic settings for each layer.

References

[1] Abdallah, S. S., Ramahi, O., and Bizheva, K. (2007). FDTD simulation of electromagnetic wave scattering from retina cells. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1639–1642. IEEE.

[2] Alvarez, J., Angulo, L. D., Bretones, A. R., and Garcia, S. G. (2012). A spurious-free discontinuous Galerkin time-domain method for the accurate modeling of microwave filters. *IEEE Transactions on Microwave Theory and Techniques*, 60(8):2359–2369.

[3] Alvarez, J., Angulo, L. D., Pantoja, M. F., Bretones, A. R., and Garcia, S. G. (2010). Source and boundary implementation in vector and scalar DGTD. *IEEE Transactions. Antennas Propagation*, 58(6):1997–2003.

[4] Álvarez González, J. (2014). *A Discontinuous Galerkin Finite Element Method for the Time-Domain Solution of Maxwell Equations*. PhD thesis, Universidad de Granada.

[5] Araújo, A., Barbeiro, S., and Ghalati, M. K. (2015). A discontinuous Galerkin scheme for solving 2D wave propagation in anisotropic materials. In *Bioengineering (ENBENG), 2015 IEEE 4th Portuguese Meeting on*, pages 1–4. IEEE.

[6] Araújo, A., Barbeiro, S., and Ghalati, M. K. (2016). Convergence of a leap-frog discontinuous Galerkin method for time-domain Maxwell's equations in anisotropic materials. In *ECMI book series of Mathematics in Industry*, (Accepted for publication). Springer.

[7] Araújo, A., Barbeiro, S., Pinto, L., Caramelo, F., Correia, A. L., Morgado, M., Serranho, P., Silva, A. S. C., and Bernardes, R. (2013). Maxwell's equations to model electromagnetic wave's propagation through eye's structures. *Proceedings of the 13th International Conference on Computational and Mathematical Methods in Science and Engineering,CMMSE 2013, Almeria, Ian Hamilton and Jesús Vigo-Aguiar Eds*, 1:121–129.

[8] Babuška, L. and Suri, M. (1987). The optimal convergence rate of the p-version of the finite element method. *SIAM Journal on Numerical Analysis*, 24(4):750–776.

[9] Berenger, J.-P. (1994). A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200.

[10] Berenger, J.-P. (1996). Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 127(2):363–379.

[11] Born, M. and Wolf, E. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, Cambridge.

[12] Busch, K., König, M., and Niegemann, J. (2011). Discontinuous Galerkin methods in nanophotonics. *Laser & Photonics Reviews*, 5(6):773–809.

[13] Cangellaris, A. C. and Wright, D. B. (1991). Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in fdtd simulations of electromagnetic phenomena. *IEEE Transactions on Antennas and Propagation*, 39(10):1518–1525.

[14] Cense, B., Chen, T. C., Park, B. H., Pierce, M. C., and De Boer, J. F. (2004). Thickness and birefringence of healthy retinal nerve fiber layer tissue measured with polarization-sensitive optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 45(8):2606–2612.

[15] Chen, Q. and Babuška, I. (1995). Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Computer Methods in Applied Mechanics and Engineering*, 128(3):405–417.

[16] Chun, S. and Hesthaven, J. S. (2009). High-order accurate thin layer approximations for time-domain electromagnetics. Part I: General metal backed coatings. *Journal of Computational and Applied Mathematics*, 231(2):598–611.

[17] Ciarlet, P. G. (1978). *The Finite Element Method for Elliptic Problems*. North-Holland, The Netherlands.

[18] Ciulla, T. A., Amador, A. G., and Zinman, B. (2003). Diabetic retinopathy and diabetic macular edema pathophysiology, screening, and novel therapies. *Diabetes Care*, 26(9):2653–2664.

[19] Cockburn, B. and Shu, C.-W. (1991). The Runge-Kutta local projection p^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO-Modélisation Mathématique et Analyse Numérique*, 25(3):337–361.

[20] Cohen, G., Ferrieres, X., and Pernet, S. (2006). A spatial high-order hexahedral discontinuous Galerkin method to solve Maxwell's equations in time domain. *Journal of Computational Physics*, 217(2):340–363.

[21] Correia, A., Pinto, L., Araújo, A., Barbeiro, S., Caramelo, F., Menezes, P., Morgado, M., Serranho, P., and Bernardes, R. (2014). Monte Carlo simulation of diabetic macular edema changes on optical coherence tomography data. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 724–727. IEEE.

[22] Courant, R., Friedrichs, K., and Lewy, H. (1967). On the partial difference equations of mathematical physics. *IBM Journal*, 11(2):215–234.

[23] Cunha-Vaz, J. and Bernardes, R. (2005). Nonproliferative retinopathy in diabetes type 2. initial stages and characterization of phenotypes. *Progress in Retinal and Eye Research*, 24(3):355–377.

[24] Dobson, D. C. (1999). An efficient method for band structure calculations in 2D photonic crystals. *Journal of Computational Physics*, 149(2):363 – 376.

[25] Dosopoulos, S. and Lee, J.-F. (2010). Interior penalty discontinuous Galerkin finite element method for the time-dependent first order Maxwell's equations. *Antennas and Propagation, IEEE Transactions on*, 58(12):4085–4090.

[26] Dunn, A. and Richards-Kortum, R. (1996). Three-dimensional computation of light scattering from cells. *IEEE Journal of Selected Topics in Quantum Electronics*, 2(4):898–905.

[27] Emmrich, E. (2005). Stability and error of the variable two-step BDF for semilinear parabolic problems. *Journal of Applied Mathematics and Computing*, 19(1-2):33–55.

[28] Enoch, J. M., Tobey, F. L., and Bedell, H. E. (1981). *Vertebrate Photoreceptor Optics*. Springer-Verlag, Berlin.

[29] Evans, L. C. (1998). *Partial Differential Equations*. American Mathematical Society, Providence, RI.

[30] Fahs, H. (2009). Development of a hp-like discontinuous Galerkin time-domain method on non-conforming simplicial meshes for electromagnetic wave propagation. *International Journal of Numerical Analysis and Modeling*, 6(2):193–216.

[31] Fallahkhair, A. B., Li, K. S., and Murphy, T. E. (2008). Vector finite difference modesolver for anisotropic dielectric waveguides. *Journal of Lightwave Technology*, 26(11):1423–1431.

[32] Fezoui, L., Lanteri, S., Lohrengel, S., and Piperno, S. (2005). Convergence and stability of a discontinuous Galerkin time-domain method for the 3D heterogeneous Maxwell's equations on unstructured meshes. *ESAIM: Mathematical Modeling and Numerical Analysis*, 39(6):1149–1176.

[33] Fujimoto, J. G., Brezinski, M. E., Tearney, G. J., Boppart, S. A., Bouma, B., Hee, M. R., Southern, J. F., and Swanson, E. A. (1995). Optical biopsy and imaging using optical coherence tomography. *Nature Medicine*, 1(9):970–972.

[34] Fujimoto, J. G., Puliafito, C., Margolis, R., Oseroff, A., De Silvestri, S., and Ippen, E. (1986). Femtosecond optical ranging in biological systems. *Optics Letters*, 11(3):150–152.

[35] Fumeaux, C., Baumann, D., Leuchtmann, P., and Vahldieck, R. (2004). A generalized local time-step scheme for efficient FVTD simulations in strongly inhomogeneous meshes. *IEEE Transactions on Microwave Theory and Techniques*, 52(3):1067–1076.

[36] Georgoulis, E. H. (2008). Inverse-type estimates on hp-finite element spaces and applications. *Mathematics of Computation*, 77(261):201–219.

[37] Harosi, F. (1981). *Microspectrophotometry and Optical Phenomena: Birefringence, Dichroism and Anomalous Dispersion*. Springer-Verlag, Berlin.

[38] Hemenger, R. P. (1989). Birefringence of a medium of tenuous parallel cylinders. *Applied Optics*, 28(18):4030–4034.

[39] Hesthaven, J. and Warburton, T. (2008). *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer-Verlag, New York.

[40] Hesthaven, J. S. (1998). From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM Journal on Numerical Analysis*, 35(2):655–676.

[41] Hesthaven, J. S. and Warburton, T. (2002). Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations. *Journal of Computational Physics*, 181(1):186–221.

[42] Hesthaven, J. S. and Warburton, T. (2004). Discontinuous Galerkin methods for the time-domain Maxwell's equations: An introduction. *ACES Newsletter*, 19:10–29.

[43] Huang, X.-R., Bagga, H., Greenfield, D. S., and Knighton, R. W. (2004). Variation of peripapillary retinal nerve fiber layer birefringence in normal human subjects. *Investigative Ophthalmology & Visual Science*, 45(9):3073–3080.

[44] Jackson, J. D. (1999). *Classical Electrodynamics*. Wiley, New York.

[45] Jin, J.-M. (2002). *The Finite Element Method in Electromagnetics*. Wiley, New York, 2nd edition.

[46] JURGENS, T., Taflove, A., Umashankar, K., and MOORE, T. (1992). Finite-difference time-domain modeling of curved surfaces. *IEEE Transactions on Antennas and Propagation*, 40(4):357–366.

[47] Kanevsky, A., Carpenter, M. H., Gottlieb, D., and Hesthaven, J. S. (2007). Application of implicit-explicit high order Runge-Kutta methods to discontinuous-Galerkin schemes. *Journal of Computational Physics*, 225(2):1753–1781.

[48] Klöckner, A., Warburton, T., Bridge, J., and Hesthaven, J. S. (2009). Nodal discontinuous Galerkin methods on graphics processors. *Journal of Computational Physics*, 228(21):7863–7882.

[49] König, M., Busch, K., and Niegemann, J. (2010). The discontinuous Galerkin time-domain method for Maxwell's equations with anisotropic materials. *Photonics and Nanostructures-Fundamentals and Applications*, 8(4):303–309.

[50] Landau, L. and Lifshitz, E. (1984). *Electrodynamics of Continuous Media*. Pergamon Press, Oxford.

[51] Lee, J.-F., Lee, R., and Cangellaris, A. (1997). Time-domain finite-element methods. *IEEE Transactions on Antennas and Propagation*, 45(3):430–442.

[52] Leis, R. (1986). *Maxwell's Equations*. Vieweg+Teubner Verlag, Wiesbaden.

[53] Leonhardt, U. and Tyc, T. (2009). Broadband invisibility by non-Euclidean cloaking. *Science*, 323(5910):110–112.

[54] Lesaint, P. and Raviart, P.-A. (1974). On a finite element method for solving the neutron transport equation. *Mathematical Aspects of Finite Elements in Partial Differential Equations*, (33):89–123.

[55] LeVeque, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, Cambridge.

[56] Li, J., Waters, J. W., and Machorro, E. A. (2012). An implicit leap-frog discontinuous Galerkin method for the time-domain Maxwell's equations in metamaterials. *Computer Methods in Applied Mechanics and Engineering*, 223:43–54.

[57] Liebman, P. A. (1975). Birefringence, dichroism and rod outer segment structure. In *Photoreceptor Optics*, pages 199–214. Springer.

[58] Limeres, J., Calvo, M. L., Enoch, J. M., and Lakshminarayanan, V. (2003). Light scattering by an array of birefringent optical waveguides: Theoretical foundations. *Journal of Optical Society of America B*, 20(7):1542–1549.

[59] Love, A. E. H. (1892). *A Treatise on the Mathematical Theory of Elasticity*. Cambridge University Press, Cambridge.

[60] Lu, T., Zhang, P., and Cai, W. (2004). Discontinuous Galerkin methods for dispersive and lossy Maxwell's equations and PML boundary conditions. *Journal of Computational Physics*, 200(2):549–580.

[61] Madsen, N. K. and Ziolkowski, R. W. (1990). A three-dimensional modified finite volume technique for Maxwell's equations. *Electromagnetics*, 10(1-2):147–161.

[62] Mikhailov, S. A. and Ziegler, K. (2007). New electromagnetic mode in graphene. *Physical Review Letters*, 99:016803.

[63] Mohammadian, A. H., Shankar, V., and Hall, W. F. (1991). Computation of electromagnetic scattering and radiation using a time-domain finite-volume discretization procedure. *Computer Physics Communications*, 68(1):175–196.

[64] Monk, P. (2003). *Finite Element Methods for Maxwell's Equations*. Oxford University Press, Oxford.

[65] Monk, P. and Süli, E. (1994). A convergence analysis of Yee's scheme on nonuniform grids. *SIAM Journal on Numerical Analysis*, 31(2):393–412.

[66] Montseny, E., Pernet, S., Ferrières, X., and Cohen, G. (2008). Dissipative terms and local time-stepping improvements in a spatial high order discontinuous Galerkin scheme for the time-domain Maxwell's equations. *Journal of Computational Physics*, 227(14):6795–6820.

[67] Piket-May, M. J., Troy, J. B., and Taflove, A. (1993). Electrodynamics of visible-light interactions with the vertebrate retinal rod. *Optics Letters*, 18(8):568–570.

[68] Piperno, S. (2000). L^2-stability of the upwind first order finite volume scheme for the Maxwell equations in two and three dimensions on arbitrary unstructured meshes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 34(1):139–158.

[69] Reed, W. H. and Hill, T. (1973). Triangular mesh methods for the neutron transport equation. *Los Alamos Report LA-UR-73-479*.

[70] Riviére, B. (2008). *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[71] Riviére, B., Wheeler, M. F., and Girault, V. (2001). A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(3):902–931.

[72] Saleh, B. E. and Teich, M. C. (1991). *Fundamentals of Photonics*. Wiley, New York.

[73] Santos, M., Araújo, A., Barbeiro, S., Caramelo, F., Correia, A., Marques, M. I., Pinto, L., Serranho, P., Bernardes, R., and Morgado, M. (2015). Simulation of cellular changes on optical coherence tomography of human retina. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 8147–8150.

[74] Schliesser, J. A., Gallimore, G., Kunjukunju, N., Sabates, N. R., Koulen, P., and Sabates, F. N. (2014). Clinical application of optical coherence tomography in combination with functional diagnostics: advantages and limitations for diagnosis and assessment of therapy outcome in central serous chorioretinopathy. *Clinical Ophthalmology (Auckland, NZ)*, 8:2337–2345.

[75] Schuman, J. S., Pedut-Kloizman, T., Pakter, H., Wang, N., Guedes, V., Huang, L., Pieroth, L., Scott, W., Hee, M., Fujimoto, J., Ishikawa, H., Bilonick, R., Kagemann, L., and Wollstein, G. (2007). Optical coherence tomography and histologic measurements of nerve fiber layer thickness in normal and glaucomatous monkey eyes. *Investigative Ophthalmology and Visual Science*, 48(8):3645–3654.

[76] Seet, K. Y., Nieminen, T. A., and Zvyagin, A. V. (2009). Refractometry of melanocyte cell nuclei using optical scatter images recorded by digital fourier microscopy. *Journal of Biomedical Optics*, 14(4):044031.

[77] Senior, T. B. and Volakis, J. L. (1995). *Approximate Boundary Conditions in Electromagnetics*. IEE, London.

[78] Serranho, P., Morgado, M., and Bernardes, R. (2012). Optical coherence tomography: a concept review in: Optical coherence tomography: A clinical and technical update. *Springer-Verlag*, pages 139–156.

[79] Shankar, V., Mohammadian, A. H., and Hall, W. F. (1990). A time-domain, finite-volume treatment for the Maxwell equations. *Electromagnetics*, 10(1-2):127–145.

[80] Sidman, R. L. (1957). The structure and concentration of solids in photoreceptor cells studied by refractometry and interference microscopy. *The Journal of Biophysical and Biochemical Cytology*, 3(1):15–30.

[81] Srinivasan, V., Chen, Y., Duker, J., and Fujimoto, J. (2009). In vivo functional imaging of intrinsic scattering changes in the human retina with high-speed ultrahigh resolution oct. *Optics Express*, 17(5):3861–3877.

[82] Taflove, A. (1995). *Computational Electrodynamics: The Finite-difference Time-domain Method*. Artech House, Boston.

[83] Tsao, P. H. (1993). Derivation and implication of the symmetry property of the permittivity tensor. *American Journal of Physics*, 61(9):823–825.

[84] van de Hulst, H. C. (1981). *Light Scattering by Small Particles*. Dover Publications, New York.

[85] Warburton, T. (2006). An explicit construction of interpolation nodes on the simplex. *Journal of Engineering Mathematics*, 56(3):247–262.

[86] Wiener, O. H. (1912). *Die Theorie des Mischkörpers für das Feld der Stationären Strömung. 1. Abhandlung: Die Mittelwertsätze für Kraft, Polarisation und Energie*. BG Teubner.

[87] Wloka, J. (1987). *Partial Differential Equations*. Cambridge University Press, Cambridge.

[88] Yariv, A. and Yeh, P. (1984). *Optical Waves in Crystals*. Wiley, New York.

[89] Yee, K. S. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transaction on Antennas and Propagation*, 14(3):302–307.

[90] Yeh, P. and Gu, C. (2009). *Optics of Liquid Crystal Displays*. Wiley Publishing, 2nd edition.

[91] Zhou, Q. and Knighton, R. W. (1997). Light scattering and form birefringence of parallel cylindrical arrays that represent cellular organelles of the retinal nerve fiber layer. *Applied Optics*, 36(10):2273–2285.

Appendix A

Technical lemmata

The lemmata included this chapter are technical tools needed to derive the stability condition and the convergence estimates. In the first section the trace and inverse inequalities that we mostly used is stated. This inequalities turn out to be very useful for analyzing DG methods. The polynomial approximations is stated in Section A.2 and at the end the discrete Gronwall's lemma that we used in our analysis is stated.

A.1 Inverse and trace inequalities

Inverse and trace inequalities are very important tools in the analysis of DG methods. We consider the following trace inequalities (see *e.g.* [70]).

Lemma A.1.1. *Let T_k be an element of $\mathscr{T}_h$ with diameter h_k and let f_k be an edge or a face of T_k. There exists a positive constant C independent of h_k such that, for any $u \in H^1(T_k)$,*

$$\|u\|_{L^2(f_k)} \leq C\sqrt{\frac{|f_k|}{|T_k|}}\left(\|u\|_{L^2(T_k)} + h_k\|\nabla u\|_{L^2(T_k)}\right). \tag{A.1}$$

Moreover, if u is a polynomials of degree less than or equal to N, there exists a positive constant C_{trace} independent of h_k and u but dependent on the polynomials degree N, such that

$$\|u\|_{L^2(f_k)} \leq C_{trace}\sqrt{\frac{|f_k|}{|T_k|}}\|u\|_{L^2(T_k)}.$$

An exact expression for the constant C_{trace} can be given as a function of the polynomials degree, and the following inequality holds for any $u \in P_N(T_k)$

$$\textit{in } 2D: \quad \|u\|_{L^2(f_k)} \leq \sqrt{\frac{(N+1)(N+2)}{2}\frac{|f_k|}{|T_k|}}\|u\|_{L^2(T_k)}, \tag{A.2}$$

$$\textit{in } 3D: \quad \|u\|_{L^2(f_k)} \leq \sqrt{\frac{(N+1)(N+3)}{3}\frac{|f_k|}{|T_k|}}\|u\|_{L^2(T_k)}. \tag{A.3}$$

Consequently, there exists a positive constant C_τ independent of h_k and N but depent on the shape-regularity h_k/τ_k, where τ_k is the diameter of the largest inscribed ball contained in T_k (see (2.14)), such that, for any $u \in P_N(T_k)$,

$$\text{in } 2D: \quad \|u\|_{L^2(\partial T_k)} \leq C_\tau \sqrt{(N+1)(N+2)} h_k^{-1/2} \|u\|_{L^2(T_k)}, \tag{A.4}$$

$$\text{in } 3D: \quad \|u\|_{L^2(\partial T_k)} \leq C_\tau \sqrt{(N+1)(N+3)} h_k^{-1/2} \|u\|_{L^2(T_k)}. \tag{A.5}$$

The next result is an inverse-type estimate ([17, 36]), where we present explicitly the dependence of the constant on the polynomials degree.

Lemma A.1.2. *Let us consider $T_k \in \mathscr{T}_h$ with diameter h_k. There exists a positive constant C_{inv} independent of h_k and N such that, for any $u \in P_N(T_k)$,*

$$\|u\|_{H^q(T_k)} \leq C_{inv} N^{2q} h_k^{-q} \|u\|_{L^2(T_k)}, \tag{A.6}$$

where $q \geq 0$.

Note that C_{inv} depends on the shape-regularity h_k/τ_k, where τ_k is the diameter of the largest inscribed ball contained in T_k. A sharper estimate reads

$$\forall u \in P_N(T_k), \quad \|u\|_{H^q(T_k)} \leq \tilde{C}_{inv} N^{2q} \tau_k^{-q} \|u\|_{L^2(T_k)}.$$

A.2 Polynomial approximation

The following lemma gives polynomial approximation with Inequality (A.7) and polynomial approximation on mesh faces with Inequality (A.8). The reader can refer to [8] or [71] for the following approximation properties.

Lemma A.2.1. *Let $T_k \in \mathscr{T}_h$ and $u \in H^p(T_k)$. Then there exits a constant C depending on p and on the shape-regularity of T_k but independent of u, h_k and N and a sequence $\mathscr{P}_N u \in P_N(T_k)$, $N = 1, 2, \ldots$, such that, for any $0 \leq q \leq p$*

$$\|u - \mathscr{P}_N u\|_{H^q(T_k)} \leq C \frac{h_k^{\sigma-q}}{N^{p-q}} \|u\|_{H^p(T_k)}, \quad p \geq 0, \tag{A.7}$$

$$\|u - \mathscr{P}_N u\|_{L^2(f_k)} \leq C \frac{h_k^{\sigma-1/2}}{N^{p-1/2}} \|u\|_{H^p(T_k)}, \quad p > \frac{1}{2}, \tag{A.8}$$

where $\sigma = \min(p, N+1)$ and f_k is an edge of T_k.

A.3 Discrete Gronwall's lemma

We use the following version of discrete Gronwall's lemma in our analysis. For the proof see *e.g.* [27, 87].

Lemma A.3.1. *Let a_n, b_n, c_n $\lambda_n \neq 0$ with $\{c_n\}$ being monodically increasing. Then*

$$a_n + b_n \leq \sum_{j=2}^{n-1} \lambda_j a_j + c_n, \quad n = 2,3,\cdots$$

implies for $n = 2,3,\cdots$

$$a_n + b_n \leq c_n \prod_{j=2}^{n-1}(1+\lambda_j) \leq c_n \exp(\sum_{j=2}^{n-1} \lambda_j).$$