

CONVERGENCE AND STABILITY OF A DISCONTINUOUS GALERKIN TIME-DOMAIN METHOD FOR THE 3D HETEROGENEOUS MAXWELL EQUATIONS ON UNSTRUCTURED MESHES

LOULA FEZOU¹, STÉPHANE LANTERI¹, STÉPHANIE LOHRENGEL² AND SERGE PIPERNO¹

Abstract. A Discontinuous Galerkin method is used for to the numerical solution of the time-domain Maxwell equations on unstructured meshes. The method relies on the choice of local basis functions, a centered mean approximation for the surface integrals and a second-order leap-frog scheme for advancing in time. The method is proved to be stable for cases with either metallic or absorbing boundary conditions, for a large class of basis functions. A discrete analog of the electromagnetic energy is conserved for metallic cavities. Convergence is proved for \mathbb{P}_k Discontinuous elements on tetrahedral meshes, as well as a discrete divergence preservation property. Promising numerical examples with low-order elements show the potential of the method.

Mathematics Subject Classification. 65M12, 65M60, 78-08, 78A40.

Received: August 17, 2004. Revised: July 2, 2005.

INTRODUCTION

The modeling of systems involving electromagnetic waves has known a kind of reinvention [24] through the resolution of the time-domain Maxwell equations on space grids. Although the finite difference time-domain (FDTD) methods based on Yee's scheme [27] are still prominent, many different types of methods have been proposed, like finite element time-domain (FETD) methods, which are based on unstructured meshes and can deal with complex geometries [4]. However, they induce heavy computations or lumping of mass matrices and significant works are still being devoted to the construction of edge elements allowing accurate and efficient mass lumping [7, 14]. Similarly, mimetic methods [13] have proven properties for Maxwell equations which make them close to the classical edge elements in the unstructured case.

Gathering many advantages, finite volume time-domain (FVTD) methods can also be based on unstructured meshes and get rid of differential operators (and finite element mass matrices) using Green's formula for the integration over finite volumes. FVTD methods (*i.e.* piecewise constant, discontinuous, Galerkin-type finite element approximation) have been developed on body-fitted coordinates [23], on unstructured finite element triangulations [4] or on totally destructured meshes [1]. First-order conservative upwind schemes, for which stability [19], convergence [8] and L^1 error estimates of $h^{1/2}$ [25] were proven, are too dissipative to be used for the

Keywords and phrases. Electromagnetics, finite volume methods, discontinuous Galerkin methods, centered fluxes, leap-frog time scheme, L^2 stability, unstructured meshes, absorbing boundary condition, convergence, divergence preservation.

¹ CERMICS, INRIA, BP93, 06902 Sophia-Antipolis Cedex, France. Serge.Piperno@cermics.enpc.fr

² Dieudonné Lab., UNSA, UMR CNRS 6621, Parc Valrose, 06108 Nice Cedex 2, France.

numerical simulation of electromagnetic waves propagation, since the numerical diffusion induced by upwinding makes long-run computations (at least over several periods) very inaccurate. Remaki has proposed a non-dissipative FVTD method in heterogeneous media [22], which is based on a second-order leap-frog time scheme and on second-order centered numerical fluxes, and yields a conservative finite volume method with no numerical diffusion, in the sense that some L^2 -type, quadratic, discrete, electromagnetic energy is conserved, even on finite volumes of arbitrary shape. Still, the convergence of such a FVTD method is limited by the piecewise constant approach (see [17] for a convergence result under some mesh regularity assumption). The discontinuous Galerkin methods enjoy a renewed favor nowadays and are now used in many and various applications [5] as people (re)discover the abilities of these methods to handle complicated geometries and meshes, to achieve a high order of accuracy by simply choosing suitable basis functions, including spectral elements in order to reach high-order accuracy [15], with different approaches either on tetrahedral meshes using Lagrange polynomials [10–12] or ongoing work on hexahedral meshes using products of Legendre polynomials. The existing software are mostly based on upwind or partially upwind fluxes, along with multi-step low-storage Runge-Kutta time-schemes (less often on a leap-frog scheme and centered fluxes) and remain highly parallelizable. One can notice here that, although Runge-Kutta schemes are robust and stable, no stability proof is available for these kind of schemes on arbitrary meshes.

At the same time, one of the most important properties which should be aimed at is the conservation of a discrete analog of the electromagnetic energy (if there is no electric conductivity). This cannot be obtained with discontinuous Galerkin methods based on upwind fluxes [4, 15, 26], although upwind fluxes lead to more robust codes, particularly for frequency-domain computations (low-frequency stabilization of centered fluxes have been proposed though [12]). Another important property is the preservation of the divergence relations in the absence of sources: the electric and magnetic fields should remain solenoidal throughout the computation. Much work has recently been done in that direction, including divergence-free basis functions [2, 6].

We present in this paper a discontinuous Galerkin method for the time domain Maxwell equations based on centered numerical fluxes and a leap-frog time-scheme. We show that this method yields discrete energy conservation and divergence preservation (in a weak sense). We dress the outline of the method in the general case in Section 2, then we prove the stability of the resulting time and space fully discretized scheme (with metallic and absorbing conditions) and its energy conservation properties in Section 3. The convergence of the method and the discrete divergence preservation property of the method are analyzed on tetrahedral meshes in Section 4. The convergence result is slightly weaker than available results for general triangulations and upwind (or monotone) fluxes [5, 9] and the same as one given recently [11] (this is also the case for the error bound on the divergence of the fields). In Section 5, we pay a particular attention to low order (\mathbb{P}_0 - \mathbb{P}_1 -DGTD) methods on tetrahedral meshes, for which numerical results are presented and compared with the exact solutions.

We consider in this paper the Maxwell equations in three space dimensions for heterogeneous anisotropic linear media with no source. The electric permittivity tensor $\bar{\epsilon}(x)$ and the magnetic permeability tensor $\bar{\mu}(x)$ are varying in space and both symmetric positive definite (for almost every $x \in \Omega$). We assume they are uniformly bounded with a strictly positive lower bound, *i.e.* there are constants $\underline{\lambda} > 0$ and $\bar{\lambda} > 0$ such that for almost every $x \in \Omega$,

$$\forall \xi \in \mathbb{R}^3, \quad \begin{cases} \underline{\lambda}|\xi|^2 \leq \xi \cdot \bar{\epsilon}(x)\xi \leq \bar{\lambda}|\xi|^2, \\ \underline{\lambda}|\xi|^2 \leq \xi \cdot \bar{\mu}(x)\xi \leq \bar{\lambda}|\xi|^2. \end{cases} \quad (1)$$

The electric field \vec{E} and the magnetic field \vec{H} verify

$$\bar{\epsilon} \frac{\partial \vec{E}}{\partial t} = \text{curl} \vec{H}, \quad \bar{\mu} \frac{\partial \vec{H}}{\partial t} = -\text{curl} \vec{E}. \quad (2)$$

These equations are set and solved on a bounded polyhedral domain Ω of \mathbb{R}^3 . Everywhere on the domain boundary $\partial\Omega$ (of unitary outwards normal \vec{n}), a boundary condition is set which is either metallic ($\vec{n} \times \vec{E} = \vec{0}$) on $\partial\Omega^m$ or absorbing ($\vec{n} \times \vec{E} = c\mu (\vec{n} \times \vec{H}) \times \vec{n}$), on $\partial\Omega^a$, where we assume the medium is isotropic near the

absorbing boundary, *i.e.* $\bar{\epsilon} = \epsilon \mathbb{I}_3$ $\bar{\mu} = \mu \mathbb{I}_3$ and the local light speed c is given by $\epsilon \mu c^2 = 1$. We recall the following existence and uniqueness result that has been proved in [18] for a regular domain and can be generalized for Lipschitz-polyhedra.

Theorem 0.1. *If $\bar{\epsilon}$ and $\bar{\mu}$ are uniformly bounded and verify (1), the problem (2) admits a unique solution $(\vec{\mathbf{E}}, \vec{\mathbf{H}}) \in C^1([0, T], (L^2(\Omega))^6) \cap C^0([0, T], \mathcal{H}(\text{curl}) \times \mathcal{H}(\text{curl}))$, for any initial data $(\vec{\mathbf{E}}_0, \vec{\mathbf{H}}_0) \in \mathcal{H}(\text{curl}) \times \mathcal{H}(\text{curl})$ satisfying the boundary conditions $\vec{n} \times \vec{\mathbf{E}} = \vec{0}$ on $\partial\Omega^m$ and $\vec{n} \times \vec{\mathbf{E}} = c\mu (\vec{n} \times \vec{\mathbf{H}}) \times \vec{n}$ on $\partial\Omega^a$.*

1. THE DISCONTINUOUS GALERKIN TIME-DOMAIN FRAMEWORK

1.1. Introduction

We assume we dispose of a partition of a polyhedral domain Ω_h (approximating the regular or Lipschitz-continuous domain of interest Ω) into a finite number of polyhedra (each one having a finite number of faces). For each polyhedral element \mathcal{T}_i , V_i denotes its volume, and $\bar{\epsilon}_i$ and $\bar{\mu}_i$ are respectively the local electric permittivity and magnetic permeability tensors of the medium, which could be varying inside the element \mathcal{T}_i . We call face between two finite elements their intersection, whenever it is a polyhedral surface. We denote by \mathcal{F}_h the union of faces and by $\mathcal{F}_h^{\text{int}} = \mathcal{F}_h / \partial\Omega_h$ the union of internal faces (common to two finite elements). For each internal face $a_{ik} = \mathcal{T}_i \cap \mathcal{T}_k$, we denote by S_{ik} the measure of a_{ik} and by \vec{n}_{ik} the unitary normal, oriented from \mathcal{T}_i towards \mathcal{T}_k . The same definitions are extended to boundary faces (in the intersection of the domain boundary $\partial\Omega_h = \partial\Omega_h^m \cup \partial\Omega_h^a$ with a finite element), the index k corresponding to a fictitious element outside the domain.

Finally, we denote by \mathcal{V}_i the set of indices of the neighboring elements of the \mathcal{T}_i (having a face in common). We also define the perimeter P_i of \mathcal{T}_i by $P_i = \sum_{k \in \mathcal{V}_i} S_{ik}$. We have the following geometrical property for all elements: $\sum_{k \in \mathcal{V}_i} S_{ik} \vec{n}_{ik} = 0$.

1.2. The time and space discretizations

Inside each finite element, the electric and magnetic fields are sought for as linear combinations $(\vec{\mathbf{E}}_i, \vec{\mathbf{H}}_i)$ of linearly independent vector fields $\vec{\varphi}_{ij}$, $1 \leq j \leq d_i$, where d_i denotes the local number of scalar degrees of freedom inside the finite element \mathcal{T}_i . We denote by $\mathcal{P}_i = \text{Span}(\vec{\varphi}_{ij}, 1 \leq j \leq d_i)$. The approximate fields $(\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h)$, defined by $(\forall i, \vec{\mathbf{E}}_h|_{\mathcal{T}_i} = \vec{\mathbf{E}}_i, \vec{\mathbf{H}}_h|_{\mathcal{T}_i} = \vec{\mathbf{H}}_i)$ are allowed to be discontinuous across element boundaries. For such a discontinuous field $\vec{\mathbf{U}}_h$, we define its average $\{\vec{\mathbf{U}}_h\}_{ik}$ through any internal face a_{ik} , as

$$\{\vec{\mathbf{U}}_h\}_{ik} = \frac{\vec{\mathbf{U}}_i|_{a_{ik}} + \vec{\mathbf{U}}_k|_{a_{ik}}}{2}.$$

Note that for any internal face a_{ik} , $\{\vec{\mathbf{U}}_h\}_{ki} = \{\vec{\mathbf{U}}_h\}_{ik}$. Dot-multiplying (2) by any given vector field $\vec{\varphi}$, integrating over \mathcal{T}_i and integrating by parts yields

$$\begin{cases} \int_{\mathcal{T}_i} \vec{\varphi} \cdot \bar{\epsilon}_i \frac{\partial \vec{\mathbf{E}}}{\partial t} = - \int_{\partial\mathcal{T}_i} \vec{\varphi} \cdot (\vec{\mathbf{H}} \times \vec{n}) + \int_{\mathcal{T}_i} \text{curl } \vec{\varphi} \cdot \vec{\mathbf{H}}, \\ \int_{\mathcal{T}_i} \vec{\varphi} \cdot \bar{\mu}_i \frac{\partial \vec{\mathbf{H}}}{\partial t} = \int_{\partial\mathcal{T}_i} \vec{\varphi} \cdot (\vec{\mathbf{E}} \times \vec{n}) - \int_{\mathcal{T}_i} \text{curl } \vec{\varphi} \cdot \vec{\mathbf{E}}. \end{cases} \tag{3}$$

In Equations (3), we now replace the exact fields $\vec{\mathbf{E}}$ and $\vec{\mathbf{H}}$ by the approximate fields $\vec{\mathbf{E}}_h$ and $\vec{\mathbf{H}}_h$ in order to evaluate volume integrals. For integrals over $\partial\mathcal{T}_i$, some additional approximations have to be done since the approximate fields are discontinuous through element faces. We choose to use completely centered fluxes, *i.e.*

$$\forall i, \forall k \in \mathcal{V}_i, \vec{\mathbf{E}}|_{a_{ik}} \simeq \{\vec{\mathbf{E}}_h\}_{ik}, \vec{\mathbf{H}}|_{a_{ik}} \simeq \{\vec{\mathbf{H}}_h\}_{ik}.$$

Concerning the time discretization, we propose to use a leap-frog advancing-in-time scheme. This kind of time scheme has both the advantages to be explicit and to be free of time-dissipation. In the sequel, superscripts refer to time stations and Δt is the fixed time-step. The unknowns related to the electric field are approximated at integer time-stations $t^n = n\Delta t$ and are denoted by $\vec{\mathbf{E}}_i^n$. The unknowns related to the magnetic field are approximated at half-integer time-stations $t^{n+\frac{1}{2}} = (n + \frac{1}{2})\Delta t$ and are denoted by $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$. Then, assuming we dispose of the fields $\vec{\mathbf{E}}_i^n$ and $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$, the fields $\vec{\mathbf{E}}_i^{n+1}$ and $\vec{\mathbf{H}}_i^{n+\frac{3}{2}}$ are sought for in \mathcal{P}_i such that, $\forall \varphi \in \mathcal{P}_i$,

$$\begin{cases} \int_{\mathcal{T}_i} \vec{\varphi} \cdot \vec{\varepsilon}_i \frac{\vec{\mathbf{E}}_i^{n+1} - \vec{\mathbf{E}}_i^n}{\Delta t} = - \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\varphi} \cdot \{\vec{\mathbf{H}}_h^{n+\frac{1}{2}}\}_{ik} \times \vec{n}_{ik} + \int_{\mathcal{T}_i} \text{curl} \vec{\varphi} \cdot \vec{\mathbf{H}}_i^{n+\frac{1}{2}}, \\ \int_{\mathcal{T}_i} \vec{\varphi} \cdot \vec{\mu}_i \frac{\vec{\mathbf{H}}_i^{n+\frac{3}{2}} - \vec{\mathbf{H}}_i^{n+\frac{1}{2}}}{\Delta t} = \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\varphi} \cdot \{\vec{\mathbf{E}}_h^{n+1}\}_{ik} \times \vec{n}_{ik} - \int_{\mathcal{T}_i} \text{curl} \vec{\varphi} \cdot \vec{\mathbf{E}}_i^{n+1}. \end{cases} \tag{4}$$

We can rewrite this formulation in terms of scalar unknowns. Inside each element, the fields $\vec{\mathbf{E}}_i^n$ and $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$ are decomposed according to

$$\forall x \in \mathcal{T}_i, \vec{\mathbf{E}}_i^n(x) = \sum_{1 \leq j \leq d_i} E_{ij}^n \vec{\varphi}_{ij}(x), \vec{\mathbf{H}}_i^{n+\frac{1}{2}}(x) = \sum_{1 \leq j \leq d_i} H_{ij}^{n+\frac{1}{2}} \vec{\varphi}_{ij}(x). \tag{5}$$

We denote by \mathbf{E}_i^n (resp. $\mathbf{H}_i^{n+\frac{1}{2}}$) the column $(E_{ij}^n)_{1 \leq l \leq d_i}$ (resp. $(H_{il}^{n+\frac{1}{2}})_{1 \leq l \leq d_i}$). The scheme (4) can be rewritten in the following semi-matrix form:

$$\begin{cases} \left[M_i^\varepsilon \frac{\mathbf{E}_i^{n+1} - \mathbf{E}_i^n}{\Delta t} \right]_j = - \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\varphi}_{ij} \cdot \{\vec{\mathbf{H}}_h^{n+\frac{1}{2}}\}_{ik} \times \vec{n}_{ik} + \int_{\mathcal{T}_i} \text{curl} \vec{\varphi}_{ij} \cdot \vec{\mathbf{H}}_i^{n+\frac{1}{2}}, \\ \left[M_i^\mu \frac{\mathbf{H}_i^{n+\frac{3}{2}} - \mathbf{H}_i^{n+\frac{1}{2}}}{\Delta t} \right]_j = \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\varphi}_{ij} \cdot \{\vec{\mathbf{E}}_h^{n+1}\}_{ik} \times \vec{n}_{ik} - \int_{\mathcal{T}_i} \text{curl} \vec{\varphi}_{ij} \cdot \vec{\mathbf{E}}_i^{n+1}, \end{cases} \tag{6}$$

where the j subscripts denote the j th component of vectors, the fields $\vec{\mathbf{E}}_i^n$, $\vec{\mathbf{E}}_i^{n+1}$, and $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$ are given in (5) in functions of scalar degrees of freedom, and M_i^ε and M_i^μ are square matrices of size d_i , given by

$$\begin{aligned} (M_i^\varepsilon)_{jl} &= \int_{\mathcal{T}_i} {}^t \vec{\varphi}_{ij} \vec{\varepsilon}_i \vec{\varphi}_{il}, \quad 1 \leq j, l \leq d_i, \\ (M_i^\mu)_{jl} &= \int_{\mathcal{T}_i} {}^t \vec{\varphi}_{ij} \vec{\mu}_i \vec{\varphi}_{il}, \quad 1 \leq j, l \leq d_i. \end{aligned} \tag{7}$$

It is clear that the matrices M_i^ε and M_i^μ are symmetric and definite positive, because the tensors $\vec{\varepsilon}_i$ and $\vec{\mu}_i$ are symmetric definite positive, and the basis vector fields $\vec{\varphi}_{ij}$ are assumed linearly independent. The semi-matrix form (6) makes clear that the actual implementation of the scheme (4) only requires the inversion of local mass matrices and the computation of numerical fluxes.

1.3. Weak treatment of boundary conditions

The metallic and absorbing conditions are dealt with in a weak sense by taking some values for the fields $\vec{\mathbf{E}}$ and $\vec{\mathbf{H}}$ at the boundary face on its outer side. In all cases, a boundary face is still denoted by a_{ik} , where i is the index of the boundary element \mathcal{T}_i and some fictitious neighbor \mathcal{T}_k is virtually created. The treatment of boundary conditions is *weak* in the sense that the traces on a_{ik} of fictitious fields $\vec{\mathbf{E}}_k^n$ and $\vec{\mathbf{H}}_k^{n+\frac{1}{2}}$ are used for the computation of numerical fluxes in (4) for the boundary element \mathcal{T}_i .

For a metallic boundary face a_{ik} , the boundary condition is $\vec{n}_{ik} \times \vec{\mathbf{E}}|_{a_{ik}} = 0$. We then choose to use

$$\text{Metallic boundary: } \begin{cases} \vec{\mathbf{E}}_k^n|_{a_{ik}} = -\vec{\mathbf{E}}_i^n|_{a_{ik}}, \\ \vec{\mathbf{H}}_k^{n+\frac{1}{2}}|_{a_{ik}} = \vec{\mathbf{H}}_i^{n+\frac{1}{2}}|_{a_{ik}}, \end{cases} \quad (8)$$

which leads to $\{\vec{\mathbf{E}}^n\}_{ik} = 0$ and $\{\vec{\mathbf{H}}^{n+\frac{1}{2}}\}_{ik} = \vec{\mathbf{H}}_i^{n+\frac{1}{2}}|_{a_{ik}}$.

For an absorbing boundary face a_{ik} , a first-order Silver–Müller absorbing condition is used. It takes the form: $\vec{n}_{ik} \times \vec{\mathbf{E}}|_{a_{ik}} = -c_i \mu_i \vec{n}_{ik} \times (\vec{n}_{ik} \times \vec{\mathbf{H}}|_{a_{ik}})$ and symmetrically $\vec{n}_{ik} \times \vec{\mathbf{H}}|_{a_{ik}} = c_i \epsilon_i \vec{n}_{ik} \times (\vec{n}_{ik} \times \vec{\mathbf{E}}|_{a_{ik}})$, where $c_i = 1/\sqrt{\mu_i \epsilon_i}$ is the local light speed (we recall we have assumed that the medium is isotropic near the absorbing boundary; the permeability and permittivity tensors $\vec{\epsilon}$ and $\vec{\mu}$ are scalars, respectively ϵ_i and μ_i). This boundary condition is exact for outgoing plane waves (with a wave vector collinear with \vec{n}_{ik}). It is a first-order approximation, asymptotically correct when the fictitious absorbing boundary is far enough and normal to the wave propagation.

In view of the absorbing boundary condition above, we propose the following fictitious fields $\vec{\mathbf{H}}_k^{n+\frac{1}{2}}$ and $\vec{\mathbf{E}}_k^{n+1}$ over an absorbing face a_{ik} .

$$\text{Absorbing boundary: } \begin{cases} \vec{\mathbf{H}}_k^{n+\frac{1}{2}}|_{a_{ik}} = c_i \epsilon_i \vec{n}_{ik} \times \frac{\vec{\mathbf{E}}_i^n|_{a_{ik}} + \vec{\mathbf{E}}_i^{n+1}|_{a_{ik}}}{2}, \\ \vec{\mathbf{E}}_k^{n+1}|_{a_{ik}} = -c_i \mu_i \vec{n}_{ik} \times \frac{\vec{\mathbf{H}}_i^{n+\frac{1}{2}}|_{a_{ik}} + \vec{\mathbf{H}}_i^{n+\frac{3}{2}}|_{a_{ik}}}{2}. \end{cases} \quad (9)$$

Some remarks can be made concerning these boundary fields:

- they are used in the computation of boundary fluxes in (4) for the element \mathcal{T}_i . Then they both are cross-multiplied by the local normal \vec{n}_{ik} , which yields back the original form of the Silver–Müller condition;
- the reader can check that the fields $\vec{\mathbf{E}}_k^n|_{a_{ik}}$ and $\vec{\mathbf{H}}_k^{n+\frac{1}{2}}|_{a_{ik}}$ are not actually available when needed for advancing in time. They lead to a locally implicit time scheme. The linear systems induced for each field are not singular since they correspond to the positive definite quadratic forms $\epsilon_i q_i$ and $\mu_i q_i$, where the positive definite quadratic form q_i is given by

$$q_i(\vec{\varphi}_{ij}, \vec{\varphi}_{il}) = \frac{1}{\Delta t} \int_{\mathcal{T}_i} \vec{\varphi}_{ij} \cdot \vec{\varphi}_{il} + \frac{c_i}{4} \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \subset \partial \Omega_h^a}} \int_{a_{ik}} (\vec{n}_{ik} \times \vec{\varphi}_{ij}) \cdot (\vec{n}_{ik} \times \vec{\varphi}_{il});$$

- one can also notice that the proposed formulae fluxes are time-consistent. They lead to a second-order time-accurate method;
- among many possible choices, the origin of these fluxes is not really obvious. In fact, these values correspond to upwind fluxes at the absorbing boundary, based on the hyperbolic nature of the global six-component Maxwell system.

2. STABILITY FOR PROBLEMS WITH METALLIC AND ABSORBING BOUNDARIES

We aim at giving and proving a sufficient condition for the L^2 -stability of the Discontinuous Galerkin scheme (4) or its equivalent matrix form (6) with only metallic and absorbing boundary conditions dealt with according to (8) and (9). We use the same kind of energy approach as in [19], where a quadratic form plays the role of a Lyapunov function of the whole set of numerical unknowns.

2.1. A discrete electromagnetic energy

Definition 2.1. We define the electromagnetic energy in any given connected group \mathcal{G} of finite elements by $\mathbb{E}_{\mathcal{G}}^n = \sum_{i \in \mathcal{G}} \mathbb{E}_i^n$, where the electromagnetic energy inside each element is given by

$$\mathbb{E}_i^n = \frac{1}{2} \int_{\mathcal{T}_i} \left({}^t \vec{\mathbf{E}}_i^n \bar{\epsilon}_i \vec{\mathbf{E}}_i^n + {}^t \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \bar{\mu}_i \vec{\mathbf{H}}_i^{n+\frac{1}{2}} \right) = \frac{1}{2} {}^t \mathbf{E}_i^n M_i^\epsilon \mathbf{E}_i^n + \frac{1}{2} {}^t \mathbf{H}_i^{n-\frac{1}{2}} M_i^\mu \mathbf{H}_i^{n+\frac{1}{2}}.$$

The reason why the discrete energy \mathbb{E}^n should be a positive definite quadratic form of all numerical unknowns is absolutely not obvious. We notice here that the situation is quite different from the proof of the L^2 -stability of the first-order upwind finite-volume scheme of [19], where the energy was obviously a positive definite quadratic form of all unknowns. At the same time, the energy proposed here depends explicitly on the numerical scheme, since it can be only written as a quadratic form of unknowns $(\mathbf{E}_i^n, \mathbf{H}_i^{n-\frac{1}{2}})$ through the use of the second part of the scheme (4) with boundary values given in (8) or (9).

The proof follows the same ideas as the one provided in [21] for the finite volume time domain method (*i.e.* the \mathbb{P}_0 -DGTD method), but we feel that the present generalization is worth writing it. In particular, energy conservation for centered schemes is not restricted to finite volumes, but would not be valid for systems where fluxes involve space varying coefficients (aeroacoustics for example). Also, the proof of the positiveness of the proposed energy is quite more complex in the general context of \mathbb{P}_k -DGTD methods with $k > 0$.

In the following, we shall prove that the proposed energy is not increasing through each time step and that it is a positive definite quadratic form of all unknowns under a CFL-like condition on the time-step Δt . This will yield the proof that the scheme (4) with metallic or absorbing boundary conditions is L^2 -stable under some sufficient condition on Δt .

2.2. Variation of the discrete electromagnetic energy through one time-step

We aim at building a discrete version of Poynting’s theorem, which states in the continuous case (with no current source) that the volumic electromagnetic energy $\mathcal{E} = \frac{1}{2} ({}^t \vec{\mathbf{E}} \bar{\epsilon} \vec{\mathbf{E}} + {}^t \vec{\mathbf{H}} \bar{\mu} \vec{\mathbf{H}})$ verifies the balance equation $\partial_t \mathcal{E} + \text{div}(\vec{\mathbf{E}} \times \vec{\mathbf{H}}) = 0$. We propose the following lemma.

Lemma 2.2. *Using the scheme (4) for an arbitrary connected group \mathcal{G} of finite elements, the variation during one time step of the discrete electromagnetic energy inside the group, defined in Definition 2.1 is given by*

$$\mathbb{E}_{\mathcal{G}}^{n+1} = \mathbb{E}_{\mathcal{G}}^n - \Delta t \sum_{a_{ik} \subset \partial \mathcal{G}} \int_{a_{ik}} \vec{n}_{ik} \cdot \frac{{}^{\vec{E}}_i^{[n+\frac{1}{2}]} \times {}^{\vec{H}}_k^{n+\frac{1}{2}} + {}^{\vec{E}}_k^{[n+\frac{1}{2}]} \times {}^{\vec{H}}_i^{n+\frac{1}{2}}}{2},$$

with the convention that $\vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} = (\vec{\mathbf{E}}_i^n + \vec{\mathbf{E}}_i^{n+1})/2$.

Proof. Because of the leap-frog structure of the energy in (2.1), and since the tensors $\bar{\epsilon}_i$ and $\bar{\mu}_i$ are symmetric, we have inside each finite element that

$$\mathbb{E}_i^{n+1} = \mathbb{E}_i^n + \int_{\mathcal{T}_i} \left({}^t \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \bar{\epsilon}_i (\vec{\mathbf{E}}_i^{n+1} - \vec{\mathbf{E}}_i^n) + \frac{1}{2} {}^t \vec{\mathbf{H}}_i^{n+\frac{1}{2}} \bar{\mu}_i (\vec{\mathbf{H}}_i^{n+\frac{3}{2}} - \vec{\mathbf{H}}_i^{n-\frac{1}{2}}) \right).$$

Since $\vec{\mathbf{E}}_i^{[n+\frac{1}{2}]}$ and $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$ are in \mathcal{P}_i , they can play the role of $\vec{\varphi}$ in (4). We obtain

$$\begin{aligned} \mathbb{E}_i^{n+1} = \mathbb{E}_i^n & - \Delta t \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \cdot \{ \vec{\mathbf{H}}^{n+\frac{1}{2}} \}_{ik} \times \vec{n}_{ik} + \Delta t \int_{\mathcal{T}_i} \vec{\mathbf{H}}_i^{n+\frac{1}{2}} \cdot \text{curl} \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \\ & + \Delta t \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} \vec{\mathbf{H}}_i^{n+\frac{1}{2}} \cdot \{ \vec{\mathbf{E}}^{[n+\frac{1}{2}]} \}_{ik} \times \vec{n}_{ik} - \Delta t \int_{\mathcal{T}_i} \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \cdot \text{curl} \vec{\mathbf{H}}_i^{n+\frac{1}{2}}. \end{aligned}$$

Reorganizing terms and integrating by parts yields

$$\mathbb{E}_i^{n+1} = \mathbb{E}_i^n - \frac{\Delta t}{2} \sum_{k \in \mathcal{V}_i} \int_{a_{ik}} (\vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \times \vec{\mathbf{H}}_k^{n+\frac{1}{2}} + \vec{\mathbf{E}}_k^{[n+\frac{1}{2}]} \times \vec{\mathbf{H}}_i^{n+\frac{1}{2}}) \cdot \vec{n}_{ik}.$$

Then, in the evaluation of $\mathbb{E}_{\mathcal{G}}^{n+1} - \mathbb{E}_{\mathcal{G}}^n$ derived from the expression above, all terms corresponding to the faces a_{ik} which are internal to the group \mathcal{G} vanish. Only boundary terms are conserved, and we simply obtain the result of the lemma. \square

Lemma 2.3. *Using the scheme (4) with metallic boundaries only, taken into account according to (8), the total discrete electromagnetic energy defined in Definition 2.1 is exactly conserved, i.e. $\mathbb{E}_{\Omega_h}^{n+1} = \mathbb{E}_{\Omega_h}^n$.*

Proof. This is a direct consequence of Lemma 2.2. Since all boundaries are metallic, we have on all boundaries of the whole polyhedral domain Ω_h : $\vec{\mathbf{E}}_k^{[n+\frac{1}{2}]}|_{a_{ik}} = -\vec{\mathbf{E}}_i^{[n+\frac{1}{2}]}|_{a_{ik}}$ and $\vec{\mathbf{H}}_k^{n+\frac{1}{2}}|_{a_{ik}} = \vec{\mathbf{H}}_i^{n+\frac{1}{2}}|_{a_{ik}}$. \square

Remark 2.4. The non-dissipative nature of the scheme proposed is already established by Lemma 2.3. It will be shown that the discrete electromagnetic energy of Definition 2.1 is a definite positive quadratic form of unknowns under some stability condition on the time-step Δt . Then, since the energy is conserved, no damped behaviour can be observed in a totally metallic cavity.

2.3. A corrected discrete electromagnetic energy for absorbing boundaries

We consider the case where some boundary faces of Ω_h are absorbing. The absorbing condition is dealt with according to (9). We can give the variation of the total electromagnetic energy given in Definition 2.1. This is the result of the following lemma.

Lemma 2.5. *Using the scheme (4)-(8)-(9), assuming the material is isotropic near absorbing boundaries, the variation $\Delta \mathbb{E}_{\Omega_h} \equiv \mathbb{E}_{\Omega_h}^{n+1} - \mathbb{E}_{\Omega_h}^n$ of the discrete electromagnetic energy (see Def. 2.1) through one time-step is given by*

$$\begin{aligned} \Delta \mathbb{E}_{\Omega_h} &= -\frac{\Delta t}{2} \sum_{a_{ik} \subset \partial \Omega_h^a} \int_{a_{ik}} \left[c_i \mu_i \left\| \vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n]} \right\|^2 + c_i \epsilon_i \left\| \vec{n}_{ik} \times \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \right\|^2 \right] \\ &\quad + \frac{\Delta t}{4} \sum_{a_{ik} \subset \partial \Omega_h^a} \int_{a_{ik}} c_i \mu_i \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \right) \cdot \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n]} \right) \\ &\quad - \frac{\Delta t}{4} \sum_{a_{ik} \subset \partial \Omega_h^a} \int_{a_{ik}} c_i \mu_i \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{n+\frac{1}{2}} \right) \cdot \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n+1]} \right), \end{aligned}$$

with the convention that $\vec{\mathbf{H}}_i^{[n]} = (\vec{\mathbf{H}}_i^{n-\frac{1}{2}} + \vec{\mathbf{H}}_i^{n+\frac{1}{2}})/2$.

Proof. This is a direct consequence of Lemma 2.2. Boundary terms corresponding to metallic boundaries vanish. For remaining terms, corresponding to absorbing faces, we just have to use the definitions of fictitious fields given in (9) and elementary recombinations lead to the result of the lemma. \square

We can point out here that the discrete energy is not anymore conserved (this is natural since we want waves to go out). It is probably non-increasing for very small time steps (since the first term in the energy variation is first-order in Δt and negative, and the two others compensate and lead to second-order terms). The form of the result of Lemma 2.5 suggests we introduce a corrected discrete electromagnetic energy.

Definition 2.6. We correct the electromagnetic energies inside each finite element and in any given connected group \mathcal{G} of finite elements \mathbb{E}_i^n and $\mathbb{E}_{\mathcal{G}}^n$ (given in Def. 2.1) by defining

$$\begin{aligned}
 \text{(i) } \forall i, \mathbb{F}_i^n &= \mathbb{E}_i^n + \frac{\Delta t}{4} \sum_{a_{ik} \subset \partial \Omega_h^n} \int_{a_{ik}} c_i \mu_i \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \right) \cdot \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n]} \right), \\
 \text{(ii) } \mathbb{F}_{\mathcal{G}}^n &= \mathbb{E}_{\mathcal{G}}^n + \frac{\Delta t}{4} \sum_{a_{ik} \subset \partial \Omega_h^n} \int_{a_{ik}} c_i \mu_i \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \right) \cdot \left(\vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n]} \right).
 \end{aligned}$$

The physical meaning of these corrected discrete energies is not clear. Correction terms are only related to absorbing boundaries (which means that $\mathbb{F}_{\Omega_h}^n = \mathbb{E}_{\Omega_h}^n$ if there are none). The additional terms probably find their origin in the time-asymmetry in the form of the electromagnetic energy. We can now prove that the discrete energy $\mathbb{F}_{\Omega_h}^n$ is non-increasing. Lemma 2.5 and Definition 2.6 lead simply to the following lemma.

Lemma 2.7. *Using the scheme (4)-(8)-(9), assuming the material is isotropic near absorbing boundaries, the corrected discrete energy $\mathbb{F}_{\Omega_h}^n$ of Definition 2.6 is non-increasing. More precisely, $\Delta \mathbb{F}_{\Omega_h} = \mathbb{F}_{\Omega_h}^{n+1} - \mathbb{F}_{\Omega_h}^n$ is given by*

$$\Delta \mathbb{F}_{\Omega_h} = -\frac{\Delta t}{2} \sum_{a_{ik} \subset \partial \Omega_h^n} \int_{a_{ik}} \left(c_i \mu_i \left\| \vec{n}_{ik} \times \vec{\mathbf{H}}_i^{[n]} \right\|^2 + c_i \epsilon_i \left\| \vec{n}_{ik} \times \vec{\mathbf{E}}_i^{[n+\frac{1}{2}]} \right\|^2 \right).$$

2.4. Definite positivity of the corrected discrete energy and stability

In order to prove that our scheme is stable, we finally show that the discrete energy $\mathbb{F}_{\Omega_h}^n$, under some stability condition on Δt , is a positive definite quadratic form of the numerical unknowns $\vec{\mathbf{H}}_i^{n-\frac{1}{2}}$ and $\vec{\mathbf{E}}_i^n$. Being non increasing, it is bounded, and the stability of the scheme is easy to prove. We assume that, inside each finite element \mathcal{T}_i , there exist two positive constants ϵ_i and μ_i such that

$$\forall \vec{\mathbf{X}}, \quad {}^t \vec{\mathbf{X}} \bar{\epsilon}_i \vec{\mathbf{X}} \geq \epsilon_i \|\vec{\mathbf{X}}\|^2, \quad {}^t \vec{\mathbf{X}} \bar{\mu}_i \vec{\mathbf{X}} \geq \mu_i \|\vec{\mathbf{X}}\|^2. \tag{10}$$

We denote by $c_i = 1/\sqrt{\epsilon_i \mu_i}$ an upper bound for the light speed in the finite element \mathcal{T}_i . We also assume that there exist dimensionless constants α_i and β_{ik} ($k \in \mathcal{V}_i$) such that

$$\forall \vec{\mathbf{X}} \in \mathcal{P}_i, \quad \begin{cases} \|\text{curl } \vec{\mathbf{X}}\|_{\mathcal{T}_i} \leq \frac{\alpha_i P_i}{V_i} \|\vec{\mathbf{X}}\|_{\mathcal{T}_i}, \\ \|\vec{\mathbf{X}}\|_{a_{ik}}^2 \leq \frac{\beta_{ik} S_{ik}}{V_i} \|\vec{\mathbf{X}}\|_{\mathcal{T}_i}^2, \end{cases} \tag{11}$$

where $\|\vec{\mathbf{X}}\|_{\mathcal{T}_i}$ and $\|\vec{\mathbf{X}}\|_{a_{ik}}$ denote the L^2 -norm of the vector field $\vec{\mathbf{X}}$ over \mathcal{T}_i and the face a_{ik} respectively. The constants α_i and β_{ik} do not depend on the size of the finite element \mathcal{T}_i (they are invariant by any homothetic transformation), but on its geometry and on the shapes of the basis fields $\vec{\varphi}_{ij}$.

Lemma 2.8. *Using the scheme (4)-(8)-(9), under assumptions (10) and (11), and assuming the material is isotropic near absorbing boundaries, the local discrete electromagnetic energy \mathbb{F}_i^n introduced in Definition 2.6 verifies*

$$\begin{aligned}
 \mathbb{F}_i^n &\geq \left(\frac{1}{2} - \frac{\alpha_i P_i c_i \Delta t}{4 V_i} \right) \left(\epsilon_i \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2 + \mu_i \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 \right) \\
 &\quad - \frac{\Delta t}{8} \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \cap \partial \Omega_h^n = \emptyset}} \left(\frac{\beta_{ik} S_{ik} \sqrt{\mu_i}}{V_i \sqrt{\epsilon_k}} \|\vec{\mathbf{H}}_i^{n-\frac{1}{2}}\|_{\mathcal{T}_i}^2 + \frac{\beta_{ki} S_{ik} \sqrt{\epsilon_k}}{V_k \sqrt{\mu_i}} \|\vec{\mathbf{E}}_k^n\|_{\mathcal{T}_k}^2 \right),
 \end{aligned}$$

where we set by convention for any metallic boundary face a_{ik} : $\|\vec{\mathbf{E}}_k^n\|_{\mathcal{T}_k} \equiv \|\vec{\mathbf{E}}_i^n\|_{\mathcal{T}_i}$, $\|\vec{\mathbf{H}}_k^{n-\frac{1}{2}}\|_{\mathcal{T}_k} \equiv \|\vec{\mathbf{H}}_i^{n-\frac{1}{2}}\|_{\mathcal{T}_i}$, $\beta_{ki} \equiv \beta_{ik}$, $V_k \equiv V_i$, $\epsilon_k \equiv \epsilon_i$, and $\mu_k \equiv \mu_i$.

Proof. Using the scheme (4) to replace the occurrences of $\vec{\mathbf{H}}_i^{n+\frac{1}{2}}$ in Definition 2.6, and using the boundary fluxes given in (8) and (9), we get

$$\begin{aligned} \mathbb{F}_i^n &= \frac{1}{2} \int_{\mathcal{T}_i} \left({}^t \vec{\mathbf{E}}_i^n \bar{\epsilon}_i \vec{\mathbf{E}}_i^n + {}^t \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \bar{\mu}_i \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \right) \\ &\quad - \frac{\Delta t}{4} \int_{\mathcal{T}_i} \left(\vec{\text{curl}} \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \cdot \vec{\mathbf{E}}_i^n + \vec{\text{curl}} \vec{\mathbf{E}}_i^n \cdot \vec{\mathbf{H}}_i^{n-\frac{1}{2}} \right) \\ &\quad + \frac{\Delta t}{4} \sum_{a_{ik} \subset \mathcal{F}_h^{\text{int}}} \int_{a_{ik}} \left(\vec{\mathbf{H}}_i^{n-\frac{1}{2}} \times \vec{\mathbf{E}}_k^n \right) \cdot \vec{n}_{ik} - \frac{\Delta t}{4} \sum_{a_{ik} \subset \partial \Omega_h^n} \int_{a_{ik}} \left(\vec{\mathbf{H}}_i^{n-\frac{1}{2}} \times \vec{\mathbf{E}}_i^n \right) \cdot \vec{n}_{ik}. \end{aligned}$$

In the remainder of this proof, we omit the superscripts n and $n-1/2$ respectively in electric and magnetic variables. We have the following inequalities:

$$\begin{aligned} \frac{1}{2} \int_{\mathcal{T}_i} \left({}^t \vec{\mathbf{E}}_i^n \bar{\epsilon}_i \vec{\mathbf{E}}_i^n + {}^t \vec{\mathbf{H}}_i^n \bar{\mu}_i \vec{\mathbf{H}}_i^n \right) &\geq \frac{\epsilon_i}{2} \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2 + \frac{\mu_i}{2} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2, \\ \frac{\Delta t}{4} \left| \int_{\mathcal{T}_i} \left(\vec{\text{curl}} \vec{\mathbf{H}}_i \cdot \vec{\mathbf{E}}_i + \vec{\text{curl}} \vec{\mathbf{E}}_i \cdot \vec{\mathbf{H}}_i \right) \right| &\leq \frac{\alpha_i P_i \Delta t}{2V_i} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i} \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}, \\ \int_{a_{ik}} \left(c_i \mu_i \|\vec{n}_{ik} \times \vec{\mathbf{H}}_i\|^2 + c_i \epsilon_i \|\vec{n}_{ik} \times \vec{\mathbf{E}}_i\|^2 \right) &\leq \frac{c_i \beta_{ik} S_{ik}}{V_i} \left(\mu_i \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 + \epsilon_i \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2 \right). \end{aligned}$$

We also have, for any metallic or internal face a_{ik} :

$$\begin{aligned} \left| \int_{a_{ik}} \left(\vec{\mathbf{H}}_i^{n-\frac{1}{2}} \times \vec{\mathbf{E}}_k^n \right) \cdot \vec{n}_{ik} \right| &\leq \frac{1}{\sqrt{\mu_i \epsilon_k}} \int_{a_{ik}} \|\sqrt{\mu_i} \vec{\mathbf{H}}_i\| \|\sqrt{\epsilon_k} \vec{\mathbf{E}}_k\| \\ &\leq \frac{1}{2} \sqrt{\frac{\mu_i}{\epsilon_k}} \|\vec{\mathbf{H}}_i\|_{a_{ik}}^2 + \frac{1}{2} \sqrt{\frac{\epsilon_k}{\mu_i}} \|\vec{\mathbf{E}}_k\|_{a_{ik}}^2 \\ &\leq \frac{1}{2} \sqrt{\frac{\mu_i}{\epsilon_k}} \frac{\beta_{ik} S_{ik}}{V_i} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 + \frac{1}{2} \sqrt{\frac{\epsilon_k}{\mu_i}} \frac{\beta_{ki} S_{ik}}{V_k} \|\vec{\mathbf{E}}_k\|_{\mathcal{T}_i}^2. \end{aligned}$$

Noticing that $\|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i} \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i} = c_i \sqrt{\mu_i} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i} \sqrt{\epsilon_i} \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i} \leq \frac{c_i}{2} (\mu_i \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 + \epsilon_i \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2)$ and gathering all lower bounds for terms in the expression of \mathbb{F}_i^n leads to

$$\begin{aligned} \mathbb{F}_i^n &\geq \left(\frac{1}{2} - \frac{\alpha_i P_i c_i \Delta t}{4V_i} \right) \left(\epsilon_i \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2 + \mu_i \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 \right) \\ &\quad - \frac{\Delta t}{8} \sum_{a_{ik} \subset \partial \Omega_h^n} \left(\sqrt{\frac{\mu_i}{\epsilon_k}} \frac{\beta_{ik} S_{ik}}{V_i} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 + \sqrt{\frac{\epsilon_k}{\mu_i}} \frac{\beta_{ik} S_{ik}}{V_i} \|\vec{\mathbf{E}}_i\|_{\mathcal{T}_i}^2 \right) \\ &\quad - \frac{\Delta t}{8} \sum_{a_{ik} \subset \mathcal{F}_h^{\text{int}}} \left(\sqrt{\frac{\mu_i}{\epsilon_k}} \frac{\beta_{ik} S_{ik}}{V_i} \|\vec{\mathbf{H}}_i\|_{\mathcal{T}_i}^2 + \sqrt{\frac{\epsilon_k}{\mu_i}} \frac{\beta_{ki} S_{ik}}{V_k} \|\vec{\mathbf{E}}_k\|_{\mathcal{T}_i}^2 \right), \end{aligned}$$

which is a less compact form of the result of the lemma (with the proposed conventions). \square

Lemma 2.9. *Using the scheme (4)-(8)-(9), under assumptions (10) and (11), and assuming the material is isotropic near absorbing boundaries, the corrected total discrete electromagnetic energy $\mathbb{F}_{\Omega_h}^n$ introduced in Definition 2.6 is a positive definite quadratic form of all unknowns (i.e. the $\vec{\mathbf{E}}_i^n$ and $\vec{\mathbf{H}}_i^{n-\frac{1}{2}}$) if*

$$\forall i, \forall k \in \mathcal{V}_i, \quad c_i \Delta t \left[2\alpha_i + \beta_{ik} \max \left(\sqrt{\frac{\mu_i}{\mu_k}}, \sqrt{\frac{\epsilon_i}{\epsilon_k}} \right) \right] < \frac{4V_i}{P_i}.$$

Proof. Following the result of the previous lemma, we first split the lower bound for \mathbb{F}_i^n using $P_i = \sum_{k \in \mathcal{V}_i} S_{ik}$. We have

$$\mathbb{F}_i^n \geq \sum_{k \in \mathcal{V}_i} \frac{S_{ik}}{P_i} \left(\frac{1}{2} - \frac{\alpha_i P_i c_i \Delta t}{4V_i} \right) \left(\epsilon_i \|\vec{\mathbf{E}}_i\|_{T_i}^2 + \mu_i \|\vec{\mathbf{H}}_i\|_{T_i}^2 \right) - \frac{\Delta t}{8} \sum_{k \in \mathcal{V}_i} \left(\frac{\beta_{ik} S_{ik} \sqrt{\mu_i}}{V_i \sqrt{\epsilon_k}} \|\vec{\mathbf{H}}_i^{n-\frac{1}{2}}\|_{T_i}^2 + \frac{\beta_{ki} S_{ik} \sqrt{\epsilon_k}}{V_k \sqrt{\mu_i}} \|\vec{\mathbf{E}}_k^n\|_{T_k}^2 \right).$$

Then, summing up these inequalities in order to obtain a lower bound for $\mathbb{F}_{\Omega_h}^n$ leads to an expression we reorganize as sum over faces. We omit again the superscripts n and $n-1/2$ in electric and magnetic variables). We find that $\mathbb{F}_{\Omega_h} \geq \sum_{a_{ik}} S_{ik} T_{ik}$ with

$$T_{ik} = \left(\frac{1}{2P_i} - \frac{\alpha_i c_i \Delta t}{4V_i} \right) \left(\epsilon_i \|\vec{\mathbf{E}}_i\|_{T_i}^2 + \mu_i \|\vec{\mathbf{H}}_i\|_{T_i}^2 \right) + \left(\frac{1}{2P_k} - \frac{\alpha_k c_k \Delta t}{4V_k} \right) \left(\epsilon_k \|\vec{\mathbf{E}}_k\|_{T_k}^2 + \mu_k \|\vec{\mathbf{H}}_k\|_{T_k}^2 \right) - \frac{\Delta t}{8} \left(\frac{\beta_{ik} \sqrt{\mu_i}}{V_i \sqrt{\epsilon_k}} \|\vec{\mathbf{H}}_i\|_{T_i}^2 + \frac{\beta_{ki} \sqrt{\epsilon_k}}{V_k \sqrt{\mu_i}} \|\vec{\mathbf{E}}_k\|_{T_k}^2 \right) - \frac{\Delta t}{8} \left(\frac{\beta_{ki} \sqrt{\mu_k}}{V_k \sqrt{\epsilon_i}} \|\vec{\mathbf{H}}_k\|_{T_k}^2 + \frac{\beta_{ik} \sqrt{\epsilon_i}}{V_i \sqrt{\mu_k}} \|\vec{\mathbf{E}}_i\|_{T_i}^2 \right).$$

This can be rewritten as

$$T_{ik} = \epsilon_i \|\vec{\mathbf{E}}_i\|_{T_i}^2 \left(\frac{1}{2P_i} - \frac{\alpha_i c_i \Delta t}{4V_i} - \frac{\beta_{ik} \Delta t}{8V_i \sqrt{\epsilon_i \mu_k}} \right) + \mu_i \|\vec{\mathbf{H}}_i\|_{T_i}^2 \left(\frac{1}{2P_i} - \frac{\alpha_i c_i \Delta t}{4V_i} - \frac{\beta_{ik} \Delta t}{8V_i \sqrt{\epsilon_k \mu_i}} \right) + \epsilon_k \|\vec{\mathbf{E}}_k\|_{T_k}^2 \left(\frac{1}{2P_k} - \frac{\alpha_k c_k \Delta t}{4V_k} - \frac{\beta_{ki} \Delta t}{8V_k \sqrt{\epsilon_k \mu_i}} \right) + \mu_k \|\vec{\mathbf{H}}_k\|_{T_k}^2 \left(\frac{1}{2P_k} - \frac{\alpha_k c_k \Delta t}{4V_k} - \frac{\beta_{ki} \Delta t}{8V_k \sqrt{\epsilon_i \mu_k}} \right).$$

Under the conditions proposed in the lemma, the T_{ik} are positive definite quadratic forms and $\mathbb{F}_{\Omega_h}^n$ is a positive definite quadratic form of all unknowns (the careful reader can check that the stability conditions also include the boundary faces). This concludes the proof. \square

Under the hypotheses of Lemma 2.9, the corrected energy $\mathbb{F}_{\Omega_h}^n$ is non-increasing, then it is bounded. Being a positive definite quadratic form of all unknowns, we get the following stability result.

Theorem 2.10. *Using the scheme (4)-(8)-(9) on arbitrary finite elements as described in this section, under assumptions (10) and (11), and assuming the material is isotropic near absorbing boundaries, the corrected total discrete electromagnetic energy $\mathbb{F}_{\Omega_h}^n$ introduced in Definition 2.6 is non-increasing through iterations. It is also a positive definite quadratic form of all unknowns ($\vec{\mathbf{E}}_i^n$ and $\vec{\mathbf{H}}_i^{n-\frac{1}{2}}$), and therefore the scheme is L^2 -stable, if the time step Δt is such that*

$$\forall i, \forall k \in \mathcal{V}_i, c_i \Delta t \left[2\alpha_i + \beta_{ik} \max \left(\sqrt{\frac{\mu_i}{\mu_k}}, \sqrt{\frac{\epsilon_i}{\epsilon_k}} \right) \right] < \frac{4V_i}{P_i} \tag{12}$$

(with the convention that $\mu_k \equiv \mu_i$ and $\epsilon_k \equiv \epsilon_i$ for boundary faces a_{ik}).

The stability condition (12) on Δt is a sufficient condition. It might be suboptimal. It has the form of a CFL stability condition, since the fraction V_i/P_i has the dimension of a length and gives an approximation of the diameter h_i of the finite element, the parameters α_i and β_{ik} are dimensionless and independent of h_i , and finally

c_i has the dimension of a wave speed (it is indeed an upper bound for the local wave speed in the heterogeneous anisotropic medium). The fact that the scheme is conditionally stable is not surprising, since it is explicit.

3. CONVERGENCE ANALYSIS AND DISCRETE DIVERGENCE PRESERVATION ON TETRAHEDRAL MESHES

In this section we shall prove the convergence of the DGTD-method in the case where all finite elements are tetrahedra and the discretization space V_h is given by discontinuous piecewise polynomials of degree at most k inside each tetrahedron:

$$V_h = \left\{ \vec{V}_h \in L^2(\Omega)^3 \mid \forall i, \vec{V}_h|_{T_i} \in \mathbb{P}_k(T_i)^3 \right\}. \tag{13}$$

We call this method the \mathbb{P}_k -DGTD. The proof of the convergence should be adapted for other spaces V_h (but the stability result of the previous section would remain valid). We shall also prove a discrete divergence preservation property which yields an upper bound for the divergence error.

We recall here that we use totally centered numerical fluxes and, because of this particular choice, the proof given here is slightly different from the one given for upwind fluxes [11] (regularity hypotheses on the solution are slightly weaker here). Other weak convergence results are also available with upwind fluxes in the general context of nonlinear systems [16].

Let us make precise some notations that will be useful in the sequel. We consider a family of unstructured grids \mathcal{T}_h , h being the mesh parameter of each unstructured grid, defined by $h = \max_i h_i$. We assume that the unstructured grids \mathcal{T}_h are uniformly *shape regular* in the sense that there is a constant $\sigma > 0$ such that

$$\forall h, \forall T_i \in \mathcal{T}_h, \frac{h_i}{\rho_i} \leq \sigma, \tag{14}$$

where ρ_i is the diameter of the biggest ball included in T_i . We shall further make the following *inverse assumption*: there is a positive constant η (independent of h) such that

$$\forall h, \forall T_i \in \mathcal{T}_h, \forall k \in \mathcal{V}_i, \frac{h_i}{h_k} \leq \eta. \tag{15}$$

We also assume that the meshes \mathcal{T}_h are compatible with the domain boundary $\partial\Omega$, *i.e.* the discretized volume $\Omega_h \equiv \bigcup_{T_i \in \mathcal{T}_h} T_i$ is equal to Ω . We finally assume that the electromagnetic coefficients $\bar{\epsilon}$ and $\bar{\mu}$ are piecewise constant (this assumption is not restrictive and the proofs could be adapted if the electromagnetic tensors were piecewise regular) and that the induced partition is also compatible in the sense that $\forall h, \forall T_i \in \mathcal{T}_h, \forall x \in T_i, \bar{\epsilon}(x) = \bar{\epsilon}_i$ and $\bar{\mu}(x) = \bar{\mu}_i$ with $\bar{\epsilon}_i, \bar{\mu}_i \in \mathbb{R}^{3 \times 3}$. We denote by Ω^j the subdomains of Ω where the electromagnetic tensors are constant. We then also have $\Omega_h^j \equiv \Omega^j$.

We next introduce the broken Sobolev spaces $PH^s(\Omega) = \{v : \forall j, v|_{\Omega^j} \in H^s(\Omega^j)\}$ equipped with the norm $\|v\|_{PH^s(\Omega)} = (\sum_j \|v|_{\Omega^j}\|_{s,\Omega^j}^2)^{1/2}$, where $\|\cdot\|_{s,\Omega^j}$ denotes the standard H^s -norm on Ω^j . Note that the element-wise traces of functions in $PH^s(\Omega)$ on the set of all faces \mathcal{F} are well defined as elements of $\text{tr}(\mathcal{F}) = \prod_{T \in \mathcal{T}_h} L^2(\partial T)$ for any $s > 1/2$. In order to avoid problems for the definition of traces in general, we make the following *regularity assumption* for the exact solution (\vec{E}, \vec{H}) of (2) in the sense of Theorem 0.1:

$$\exists s > 1/2 \text{ such that } \forall h, \forall t \in [0, T] : (\vec{E}(t), \vec{H}(t)) \in PH^s(\Omega)^6. \tag{16}$$

This assumption guarantees that $\{\vec{E}\} \times \vec{n}$ and $\{\vec{H}\} \times \vec{n}$ are well defined as elements in $s\text{ftr}(\mathcal{F})$. Moreover, we can state that $\{\vec{E}\} \times \vec{n} = \vec{E} \times \vec{n}$ and $\{\vec{H}\} \times \vec{n} = \vec{H} \times \vec{n}$ on $[0, T] \times a$ for any internal face $a \in \mathcal{F}_h^{\text{int}}$, since $\vec{E}(t)$ and $\vec{H}(t)$ belong to $\mathcal{H}(\text{curl})$ for any $t \in [0, T]$.

3.1. Properties of the semi-discretized scheme

In a first step we shall discuss consistency and stability of the spatially semi-discretized method. In the following, we denote by $\vec{\mathbf{U}}_i$ the restriction of $\vec{\mathbf{U}} \in H^s(\mathcal{T}_h)^3$ with $s > 1/2$ to the element \mathcal{T}_i . According to (8) and (9), boundary conditions for the semi-discretized scheme are taken into account as follows: for any metallic boundary face $a_{ik} \subset \partial\Omega_h^m$, we take $\vec{\mathbf{E}}_{k|a_{ik}}(t) = -\vec{\mathbf{E}}_{i|a_{ik}}(t)$ and $\vec{\mathbf{H}}_{k|a_{ik}}(t) = \vec{\mathbf{H}}_{i|a_{ik}}(t)$, while for any absorbing boundary face $a_{ik} \subset \partial\Omega_h^a$, we take $\vec{\mathbf{E}}_{k|a_{ik}}(t) = -c_i\mu_i \vec{n}_{ik} \times \vec{\mathbf{H}}_{i|a_{ik}}(t)$ and $\vec{\mathbf{H}}_{k|a_{ik}}(t) = c_i\epsilon_i \vec{n}_{ik} \times \vec{\mathbf{E}}_{i|a_{ik}}(t)$. The semi-discrete solution $(\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h)$ is thus the solution in $C^1([0, T]; V_h^2)$ of the following weak formulation: for any test field $(\vec{\mathbf{U}}_h, \vec{\mathbf{V}}_h) \in V_h^2$, for $0 < t \leq T$ and all i ,

$$\left\{ \begin{aligned} & \int_{T_i} \vec{\mathbf{V}}_i \cdot \bar{\mu}_i \frac{\partial \vec{\mathbf{H}}_i}{\partial t} + \int_{T_i} \vec{\mathbf{E}}_i \cdot \text{curl} \vec{\mathbf{V}}_i - \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \in \mathcal{F}_h^{\text{int}}}} \int_{a_{ik}} \vec{\mathbf{V}}_i \cdot (\{\vec{\mathbf{E}}_h\}_{ik} \times \vec{n}_{ik}) \\ & - \frac{1}{2} \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \subset \partial\Omega_h^a}} \int_{a_{ik}} \vec{\mathbf{V}}_i \cdot (\vec{\mathbf{E}}_i - c_i\mu_i \vec{n}_{ik} \times \vec{\mathbf{H}}_i) \times \vec{n}_{ik} = 0, \\ & \int_{T_i} \vec{\mathbf{U}}_i \cdot \bar{\epsilon}_i \frac{\partial \vec{\mathbf{E}}_i}{\partial t} - \int_{T_i} \vec{\mathbf{H}}_i \cdot \text{curl} \vec{\mathbf{U}}_i + \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \in \mathcal{F}_h^{\text{int}}}} \int_{a_{ik}} \vec{\mathbf{U}}_i \cdot (\{\vec{\mathbf{H}}_h\}_{ik} \times \vec{n}_{ik}) \\ & + \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \subset \partial\Omega_h^m}} \int_{a_{ik}} \vec{\mathbf{U}}_i \cdot (\vec{\mathbf{H}}_i \times \vec{n}_{ik}) + \frac{1}{2} \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \subset \partial\Omega_h^a}} \int_{a_{ik}} \vec{\mathbf{U}}_i \cdot (\vec{\mathbf{H}}_i + c_i\epsilon_i \vec{n}_{ik} \times \vec{\mathbf{E}}_i) \times \vec{n}_{ik} = 0, \end{aligned} \right. \tag{17}$$

together with the initial conditions

$$\vec{\mathbf{E}}_h(0) = P_h(\vec{\mathbf{E}}_0) \quad \text{and} \quad \vec{\mathbf{H}}_h(0) = P_h(\vec{\mathbf{H}}_0), \tag{18}$$

where $P_h : L^2(\Omega)^3 \rightarrow V_h$ denotes the orthogonal projection onto V_h . We introduce the following bilinear forms that are well defined on V_h . For $\vec{\mathbf{Q}} = (\vec{\mathbf{U}}, \vec{\mathbf{V}})$ and $\vec{\mathbf{Q}}' = (\vec{\mathbf{U}}', \vec{\mathbf{V}}')$, we set

$$m(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') = \int_{\Omega} (\bar{\epsilon} \vec{\mathbf{U}} \cdot \vec{\mathbf{U}}' + \bar{\mu} \vec{\mathbf{V}} \cdot \vec{\mathbf{V}}'), \tag{19}$$

$$a(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') = \int_{\Omega} (\vec{\mathbf{U}} \cdot \text{curl}_h \vec{\mathbf{V}}' - \vec{\mathbf{V}} \cdot \text{curl}_h \vec{\mathbf{U}}'), \tag{20}$$

$$\begin{aligned} b(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') &= \int_{\mathcal{F}_h^{\text{int}}} \{\vec{\mathbf{V}}\} \cdot [\vec{\mathbf{U}}'] - \int_{\mathcal{F}_h^{\text{int}}} \{\vec{\mathbf{U}}\} \cdot [\vec{\mathbf{V}}'] + \int_{\partial\Omega_h^m} (\vec{\mathbf{V}} \times \vec{n}) \cdot \vec{\mathbf{U}}' \\ &+ \frac{1}{2} \int_{\partial\Omega_h^a} (\vec{\mathbf{V}} \times \vec{n}) \cdot \vec{\mathbf{U}}' - (\vec{\mathbf{U}} \times \vec{n}) \cdot \vec{\mathbf{V}}' \\ &+ \frac{1}{2} \int_{\partial\Omega_h^a} (c\epsilon(\vec{n} \times \vec{\mathbf{U}}) \times \vec{n}) \cdot \vec{\mathbf{U}}' + (c\mu(\vec{n} \times \vec{\mathbf{V}}) \times \vec{n}) \cdot \vec{\mathbf{V}}' \end{aligned} \tag{21}$$

where curl_h is the piecewise curl-operator given by $\forall i, (\text{curl}_h \vec{\mathbf{U}})|_{\mathcal{T}_i} = \text{curl}(\vec{\mathbf{U}}|_{\mathcal{T}_i})$ and $[\vec{\mathbf{U}}_h]_{ik}$ denotes the tangential jump of the discontinuous field $\vec{\mathbf{U}}_h$ through any internal face a_{ik} , defined as follows:

$$[\vec{\mathbf{U}}_h]_{ik} = (\vec{\mathbf{U}}_{k|a_{ik}} - \vec{\mathbf{U}}_{i|a_{ik}}) \times \vec{n}_{ik}.$$

Note that for any internal face a_{ik} , $[[\vec{\mathbf{U}}_h]]_{ki} = [[\vec{\mathbf{U}}_h]]_{ik}$. Summing up the identities in (17) with respect to i , we see that the discrete solution $\vec{\mathbf{Q}}_h = (\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h)$ satisfies

$$m(\partial_t \vec{\mathbf{Q}}_h, \vec{\mathbf{Q}}') + a(\vec{\mathbf{Q}}_h, \vec{\mathbf{Q}}') + b(\vec{\mathbf{Q}}_h, \vec{\mathbf{Q}}') = 0 \quad \forall \vec{\mathbf{Q}}' \in V_h \times V_h. \tag{22}$$

We now state that the method is consistent with (2) in the following sense.

Proposition 3.1. *Let $\vec{\mathbf{Q}} = (\vec{\mathbf{E}}, \vec{\mathbf{H}})$ be the exact solution of (2). We assume that the regularity assumption (16) holds. Then*

$$m(\partial_t \vec{\mathbf{Q}}, \vec{\mathbf{Q}}') + a(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') + b(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') = 0, \quad \forall \vec{\mathbf{Q}}' \in V_h \times V_h. \tag{23}$$

Proof. We first notice that the bilinear forms are well defined for $\vec{\mathbf{Q}}$ under the regularity assumption. Next, we know from Theorem 0.1 that $\vec{\mathbf{Q}}(t) \in \mathcal{H}(\text{curl}) \times \mathcal{H}(\text{curl})$ for $t \in [0, T]$ and thus $\forall i, \forall k \in \mathcal{V}_i, \{\vec{\mathbf{H}}\}_{ik} \cdot [[\vec{\mathbf{U}}]]_{ik} = (\vec{\mathbf{H}}_{a_{ik}} \times \vec{n}_{ik}) \cdot (\vec{\mathbf{U}}_{i|a_{ik}} - \vec{\mathbf{U}}_{k|a_{ik}})$. Taking into account that $ce(\vec{n} \times \vec{\mathbf{E}}) \times \vec{n} = \vec{\mathbf{H}} \times \vec{n}$ and $c\mu(\vec{n} \times \vec{\mathbf{H}}) \times \vec{n} = -\vec{\mathbf{E}} \times \vec{n}$ on absorbing boundary faces $a_{ik} \subset \Omega_h^a$ as well as $\vec{\mathbf{E}} \times \vec{n} = 0$ on metallic boundary faces yields

$$\begin{aligned} b(\vec{\mathbf{Q}}, \vec{\mathbf{Q}}') &= \sum_{a_{ik} \subset \mathcal{F}_h^{\text{int}}} \int_{a_{ik}} (\vec{\mathbf{H}}_{|a_{ik}} \times \vec{n}_{ik}) \cdot (\vec{\mathbf{U}}_i - \vec{\mathbf{U}}_k) - (\vec{\mathbf{E}}_{|a_{ik}} \times \vec{n}_{ik}) \cdot (\vec{\mathbf{V}}_i - \vec{\mathbf{V}}_k) \\ &\quad + \sum_{a_{ik} \subset \partial\Omega_h} \int_{a_{ik}} (\vec{\mathbf{H}}_{|a_{ik}} \times \vec{n}_{ik}) \cdot \vec{\mathbf{U}}_i - (\vec{\mathbf{E}}_{|a_{ik}} \times \vec{n}_{ik}) \cdot \vec{\mathbf{V}}_i \\ &= \sum_{\mathcal{T}_i} \int_{\partial\mathcal{T}_i} (\vec{\mathbf{H}} \times \vec{n}) \cdot \vec{\mathbf{U}}_i - (\vec{\mathbf{E}} \times \vec{n}) \cdot \vec{\mathbf{V}}_i, \end{aligned}$$

and (23) then follows by integration by parts. □

In the sequel, the following stability lemma will be useful, which can be obtained by simple calculus.

Lemma 3.2. *For any $\vec{\mathbf{Q}}_h$ in $V_h \times V_h$, we have*

$$a(\vec{\mathbf{Q}}_h, \vec{\mathbf{Q}}_h) + b(\vec{\mathbf{Q}}_h, \vec{\mathbf{Q}}_h) = \frac{1}{2} \int_{\partial\Omega_h^a} ce|\vec{n} \times \vec{\mathbf{U}}_h|^2 + c\mu|\vec{n} \times \vec{\mathbf{V}}_h|^2. \tag{24}$$

The semi-discrete energy $\mathcal{E}_h(t) = \frac{1}{2}m(\vec{\mathbf{Q}}_h(t), \vec{\mathbf{Q}}_h(t))$, where $\vec{\mathbf{Q}}_h = (\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h)$ is the solution of (17)-(18), is thus non-increasing on $[0, T]$ and we have

$$\mathcal{E}_h(t) \leq \mathcal{E}_h(0) \leq \mathcal{E}(0) \quad \forall t \in [0, T].$$

3.2. Convergence of the semi-discretized problem

We recall some standard approximation results known in the context of finite element methods:

Lemma 3.3. (cf. [3]) *Let $T \in \mathcal{T}_h$ and assume that u belongs to the space $H^{s+1}(T)$ for $s \geq 0$. Let Π be a linear continuous operator from $H^{s+1}(T)$ onto $\mathbb{P}_k(T)$ such that $\Pi(u) = u$ for all $u \in \mathbb{P}_k(T)$. Then we have*

$$|u - \Pi(u)|_{m,T} \leq Ch_T^{\min\{s,k\}+1-m} \|u\|_{s+1,T}, \quad m = 0, 1 \tag{25}$$

$$\|u - \Pi(u)\|_{0,\partial T} \leq Ch_T^{\min\{s,k\}+1/2} \|u\|_{s+1,T} \tag{26}$$

where C is a positive constant depending only on k, s and the regularity parameter σ of the mesh.

Lemma 3.4. (cf. [3]) For all $p \in \mathbb{P}_k(T)$ we have

$$\|p\|_{0,\partial T} \leq C_{\text{inv}} h_T^{-1/2} \|p\|_{0,T} \tag{27}$$

$$\|p\|_{1,T} \leq C_{\text{inv}} h_T^{-1} \|p\|_{0,T} \tag{28}$$

where C_{inv} is a positive constant depending only on k and the regularity parameter σ of the mesh.

Theorem 3.5. Let $(\mathcal{T}_h)_h$ be a family of unstructured meshes satisfying (14) and (15). Let $\bar{\epsilon}$ and $\bar{\mu}$ satisfy (1). V_h is given by (13).

Let $(\vec{\mathbf{E}}, \vec{\mathbf{H}})$ be the exact solution of (2) and let $(\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h) \in C^1([0, T], V_h^2)$ be the semi-discrete solution satisfying (17)-(18). If in addition $(\vec{\mathbf{E}}, \vec{\mathbf{H}})$ belongs to $C^0([0, T]; \text{PH}^{s+1}(\Omega)^6)$ for $s \geq 0$, then there is a constant $C > 0$ independent of h such that

$$\max_{t \in [0, T]} \left(\|P_h(\vec{\mathbf{E}}) - \vec{\mathbf{E}}_h\|_{0,\Omega}^2 + \|P_h(\vec{\mathbf{H}}) - \vec{\mathbf{H}}_h\|_{0,\Omega}^2 \right)^{1/2} \leq C T h^{\min(s,k)} \|(\vec{\mathbf{E}}, \vec{\mathbf{H}})\|_{C^0([0, T], \text{PH}^{s+1}(\Omega))}. \tag{29}$$

Proof. To prove (29), we proceed in two steps. First, we will show that the left hand side in (29) actually may be majored by a term involving only surface integrals. In a second step, we then will provide estimates with the help of the standard error estimates (26) and (27).

Let $\vec{\mathbf{Q}} = (\vec{\mathbf{E}}, \vec{\mathbf{H}})$, $\vec{\mathbf{Q}}_h = (\vec{\mathbf{E}}_h, \vec{\mathbf{H}}_h)$ and $\vec{\mathbf{q}} = \vec{\mathbf{Q}} - \vec{\mathbf{Q}}_h = (\vec{\mathbf{e}}, \vec{\mathbf{h}})$ (thus $\vec{\mathbf{e}}$ and $\vec{\mathbf{h}}$ denote the error of, respectively, the electric field and the magnetic field). Let us introduce $\epsilon(t) = \frac{1}{2} m(P_h(\vec{\mathbf{q}})(t), P_h(\vec{\mathbf{q}})(t))$. Using the lower bound $\lambda > 0$ of $\bar{\epsilon}$ and $\bar{\mu}$, we first get, for $0 < t \leq T$,

$$\frac{\lambda}{2} \left(\|P_h(\vec{\mathbf{e}})(t)\|_{0,\Omega}^2 + \|P_h(\vec{\mathbf{h}})(t)\|_{0,\Omega}^2 \right) \leq \epsilon(t).$$

Due to the definition of the discrete initial conditions, we have $\epsilon(0) = 0$ and then, for $0 < t \leq T$,

$$\epsilon(t) = \frac{1}{2} \int_0^t \frac{d}{ds} m(P_h(\vec{\mathbf{q}})(s), P_h(\vec{\mathbf{q}})(s)) ds = \int_0^t m(\partial_s P_h(\vec{\mathbf{q}})(s), P_h(\vec{\mathbf{q}})(s)) ds.$$

Applying the first result of Lemma 3.2 to $P_h(\vec{\mathbf{q}}) \in C^1([0, T]; V_h \times V_h)$, we get

$$\epsilon(t) \leq \int_0^t \left(m(\partial_s P_h(\vec{\mathbf{q}})(s), P_h(\vec{\mathbf{q}})(s)) + a(P_h(\vec{\mathbf{q}})(s), P_h(\vec{\mathbf{q}})(s)) + b(P_h(\vec{\mathbf{q}})(s), P_h(\vec{\mathbf{q}})(s)) \right) ds. \tag{30}$$

Subtracting (22) from the consistency result (23) with $\vec{\mathbf{Q}}'_h = P_h(\vec{\mathbf{q}})(s)$ yields

$$m(\partial_s \vec{\mathbf{q}}(s), P_h(\vec{\mathbf{q}})(s)) + a(\vec{\mathbf{q}}(s), P_h(\vec{\mathbf{q}})(s)) + b(\vec{\mathbf{q}}(s), P_h(\vec{\mathbf{q}})(s)) = 0.$$

Since $P_h(\vec{\mathbf{Q}}_h) = \vec{\mathbf{Q}}_h$, subtracting the above equality from (30) leads to

$$\begin{aligned} \epsilon(t) \leq & \int_0^t \left(m([P_h(\partial_s \vec{\mathbf{Q}}) - \partial_s \vec{\mathbf{Q}}](s), P_h(\vec{\mathbf{q}})(s)) \right. \\ & \left. + a([P_h(\vec{\mathbf{Q}}) - \vec{\mathbf{Q}}](s), P_h(\vec{\mathbf{q}})(s)) + b([P_h(\vec{\mathbf{Q}}) - \vec{\mathbf{Q}}](s), P_h(\vec{\mathbf{q}})(s)) \right) ds. \end{aligned}$$

But P_h is a projector on V_h and $P_h(\vec{\mathbf{q}})$ belongs to $V_h \times V_h$. Thus, $m(P_h(\partial_s \vec{\mathbf{Q}}) - \partial_s \vec{\mathbf{Q}}, P_h(\vec{\mathbf{q}})) = 0$. In the same way, it follows that $a([P_h(\vec{\mathbf{Q}}) - \vec{\mathbf{Q}}], P_h(\vec{\mathbf{q}})) = 0$ since $\text{curl}_h(P_h(\vec{\mathbf{q}}))(s) \in V_h \times V_h$ for all $0 < s \leq t$. We thus get

$$\frac{\lambda}{2} \left(\|P_h(\vec{\mathbf{e}})(t)\|_{0,\Omega}^2 + \|P_h(\vec{\mathbf{h}})(t)\|_{0,\Omega}^2 \right) \leq \int_0^t b([P_h(\vec{\mathbf{Q}}) - \vec{\mathbf{Q}}](s), P_h(\vec{\mathbf{q}})(s)) ds \tag{31}$$

which proves the first step. Now, we estimate the surface integrals in the preceding expression which derive from the definition of $b(\cdot, \cdot)$. Let $a_{ik} \in \mathcal{F}_h^{\text{int}}$ be an internal face shared by the tetrahedra \mathcal{T}_i and \mathcal{T}_k . Due to Lemmas 3.3 and 3.4, we get, according to the inverse assumption (15) on the mesh:

$$\int_{a_{ik}} \{P_h(\vec{\mathbf{H}}) - \vec{\mathbf{H}}\}_{ik} \cdot \llbracket P_h \vec{\mathbf{e}} \rrbracket_{ik} d\sigma \leq C\eta h^{\min(s,k)} \left(\|\vec{\mathbf{H}}\|_{s+1, \mathcal{T}_i}^2 + \|\vec{\mathbf{H}}\|_{s+1, \mathcal{T}_k}^2 \right)^{\frac{1}{2}} \left(\|P_h(\vec{\mathbf{e}})\|_{0, \mathcal{T}_i}^2 + \|P_h(\vec{\mathbf{e}})\|_{0, \mathcal{T}_k}^2 \right)^{\frac{1}{2}},$$

with a generic constant $C > 0$ independent from h and $\vec{\mathbf{Q}}$. We get the corresponding upper bound for the term $\int_{a_{ik}} \{P_h(\vec{\mathbf{H}}) - \vec{\mathbf{H}}\}_{ik} \cdot \llbracket P_h \vec{\mathbf{e}} \rrbracket_{ik} d\sigma$. Next, let a_{ik} be a boundary face of $\partial\Omega_h^m$ such that $a_{ik} \subset \partial\mathcal{T}_i$. We have

$$\int_a \left[\left(P_h(\vec{\mathbf{H}}) - \vec{\mathbf{H}} \right) \times \vec{n} \right] \cdot P_h \vec{\mathbf{e}} d\sigma \leq Ch^{\min(s,k)} \|\vec{\mathbf{H}}\|_{s+1, \mathcal{T}_i} \|P_h(\vec{\mathbf{e}})\|_{0, \mathcal{T}_i}.$$

In the same way we deal with the other terms involving boundary faces. Summing up with respect to all interfaces together with Cauchy-Schwartz' inequality (and taking into account that each element \mathcal{T}_i of the mesh has at most four neighbors) then yields

$$b([P_h(\vec{\mathbf{Q}}) - \vec{\mathbf{Q}}](s), P_h(\vec{\mathbf{q}})(s)) \leq C\eta h^{\min(s,k)} \left(\|P_h(\vec{\mathbf{e}})(s)\|_{0, \Omega}^2 + \|P_h(\vec{\mathbf{h}})(s)\|_{0, \Omega}^2 \right)^{1/2} \|(\vec{\mathbf{E}}(s), \vec{\mathbf{H}}(s))\|_{\mathbf{PH}^{s+1}(\Omega)},$$

Integration in time, together with (31), finally leads to the result. □

In [11], the same convergence order is obtained for upwind fluxes and problems with periodic or perfect conducting boundary conditions. Notice that the regularity assumptions on the exact solution are slightly stronger than those in Theorem 3.5.

An estimation for the error $\vec{\mathbf{q}}$ can be easily deduced from Theorem 3.5 since $\vec{\mathbf{q}} = \vec{\mathbf{Q}} - \vec{\mathbf{Q}}_h = \vec{\mathbf{Q}} - P_h(\vec{\mathbf{Q}}) + P_h(\vec{\mathbf{q}})$ and the approximation error $\|\vec{\mathbf{Q}} - P_h(\vec{\mathbf{Q}})\|_{0, \Omega}$ is of order $\mathcal{O}(h^{\min(s,k)+1})$.

Corollary 3.6. *Under the assumptions of Theorem 3.5, there is a constant $C > 0$ independent from the mesh parameter h such that the error $\vec{\mathbf{q}} = (\vec{\mathbf{E}} - \vec{\mathbf{E}}_h, \vec{\mathbf{H}} - \vec{\mathbf{H}}_h)$ satisfies the estimate*

$$\|\vec{\mathbf{q}}\|_{C^0((0,T), L^2(\Omega))} \leq CT h^{\min(s,k)} \|(\vec{\mathbf{E}}, \vec{\mathbf{H}})\|_{C^0((0,T), \mathbf{PH}^{s+1}(\Omega))}. \tag{32}$$

3.3. Convergence of the totally discretized problem

The completely discretized scheme (6) may be seen as the discretization in time of a system of ordinary differential equations. This time discretization is everywhere second-order accurate. Estimates of the consistency error then follow directly from Taylor expansions. Indeed, if $(\vec{\mathbf{E}}_h^{n+1}, \vec{\mathbf{H}}_h^{n+\frac{3}{2}})$ have been computed from $\vec{\mathbf{E}}_h^n = \vec{\mathbf{E}}_h(t_n)$ and $\vec{\mathbf{H}}_h^{n+\frac{1}{2}} = \vec{\mathbf{H}}_h(t_{n+\frac{1}{2}})$ by (6) where $(\vec{\mathbf{E}}_h(\cdot), \vec{\mathbf{H}}_h(\cdot))$ denotes the semi-discrete solution of (17), we simply get that for some constant C independent of Δt and h ,

$$\|\vec{\mathbf{E}}_h(t_{n+1}) - \vec{\mathbf{E}}_h^{n+1}\|_{0, \Omega} + \|\vec{\mathbf{H}}_h(t_{n+3/2}) - \vec{\mathbf{H}}_h^{n+3/2}\|_{0, \Omega} \leq C \frac{\bar{\lambda}}{\Delta} \Delta t^3 \left(\|\partial_t^3 \vec{\mathbf{E}}_h\|_{C^0(0,T; L^2(\Omega))} + \|\partial_t^3 \vec{\mathbf{H}}_h\|_{C^0(0,T; L^2(\Omega))} \right).$$

The main ideas which should easily lead to the result are based on the following assumptions and remarks:

- the exact solution of the Maxwell system (2) is regular enough in time and space;
- second and first order accurate estimates for the initial values of first and second time derivatives of $(\vec{\mathbf{E}}, \vec{\mathbf{H}})$ are available;

- since we have noticed that $(\partial_t^m \vec{\mathbf{E}}, \partial_t^m \vec{\mathbf{H}})$ are also solutions of Maxwell equations (time derivative of fields are also solutions of Maxwell equations, with the same boundary conditions), $\|\partial_t^m \vec{\mathbf{E}}_h\|_{C^0(0,T;L^2(\Omega))}$ and $\|\partial_t^m \vec{\mathbf{H}}_h\|_{C^0(0,T;L^2(\Omega))}$ can be bounded independently from h ;
- $(\partial_t^m \vec{\mathbf{E}}_h, \partial_t^m \vec{\mathbf{H}}_h)$ are some discrete approximation of $(\partial_t^m \vec{\mathbf{E}}, \partial_t^m \vec{\mathbf{H}})$.

In that case, the consistency error is altogether of $\mathcal{O}(\Delta t^2)$ (the leapfrog time-scheme is second-order accurate and time-derivatives of fields are also solution of Maxwell equations; then the total error is the sum of terms in $\mathcal{O}(\Delta t^2)$ for the fields and derivatives, summed up with Δt times the initial error on the time-derivatives, which was of order Δt , and Δt^2 times the initial error on the time-derivatives, which was of order 1). Together with the stability result of Theorem 2.10 we thus get an error of order $\mathcal{O}(Th^{\min(s,k)}) + \mathcal{O}(\Delta t^2)$ for the total error in $C^0([0, T]; L^2(\Omega))$.

3.4. A discrete divergence preservation property

In the Maxwell equations (2), the divergence relations $\text{div}(\vec{\varepsilon} \vec{\mathbf{E}}) = 0$ and $\text{div}(\vec{\mu} \vec{\mathbf{H}}) = 0$ have not been considered because they are redundant in the absence of sources provided they are verified initially. This is true for the exact solutions but might lead to spurious effects (equivalent to artificial creation of electric charges) in the approximate solution. We prove here that the \mathbb{P}_k -DGTD method (tetrahedral mesh, piecewise polynomials) preserves the divergence relations in the following discrete weak sense.

Proposition 3.7. *Let $X_h \subset H_0^1(\Omega)$ be the space of continuous, element-wise polynomial functions defined as $X_h = \{v \in C^0(\bar{\Omega}) \mid \forall i, v|_{T_i} \in \mathbb{P}_{k+1}, v|_{\Gamma} = 0\}$. The electromagnetic fields obtained with the scheme (4) verify*

$$\forall \psi \in X_h, \begin{cases} \langle \text{div} \vec{\varepsilon} \vec{\mathbf{E}}_h^{n+1}, \psi \rangle_{-1} = \langle \text{div} \vec{\varepsilon} \vec{\mathbf{E}}_h^n, \psi \rangle_{-1} \\ \langle \text{div} \vec{\mu} \vec{\mathbf{H}}_h^{n+\frac{1}{2}}, \psi \rangle_{-1} = \langle \text{div} \vec{\mu} \vec{\mathbf{H}}_h^{n-\frac{1}{2}}, \psi \rangle_{-1}, \end{cases} \tag{33}$$

where $\langle \cdot, \cdot \rangle_{-1}$ denotes the duality product between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. In addition, if the initial data $\vec{\varepsilon} \vec{\mathbf{E}}_0$ and $\vec{\mu} \vec{\mathbf{H}}_0$ are divergence-free, the approximate solution is discrete divergence-free, i.e.

$$\forall \psi \in X_h, \begin{cases} \langle \text{div} \vec{\varepsilon} \vec{\mathbf{E}}_h^{n+1}, \psi \rangle_{-1} = 0 \\ \langle \text{div} \vec{\mu} \vec{\mathbf{H}}_h^{n+\frac{1}{2}}, \psi \rangle_{-1} = 0. \end{cases}$$

Proof. The proof is quite simple. If ψ belongs to $X_h \subset H_0^1(\Omega)$, we have

$$\begin{aligned} \langle \text{div} (\vec{\varepsilon} (\vec{\mathbf{E}}_h^{n+1} - \vec{\mathbf{E}}_h^n)), \psi \rangle_{-1} &= - \int_{\Omega} \nabla \psi \cdot \vec{\varepsilon} (\vec{\mathbf{E}}_h^{n+1} - \vec{\mathbf{E}}_h^n) \\ &= - \sum_i \int_{T_i} \vec{\nabla} \psi \cdot \vec{\varepsilon}_i (\vec{\mathbf{E}}_i^{n+1} - \vec{\mathbf{E}}_i^n). \end{aligned}$$

Since $\vec{\nabla} \psi$ is defined inside all tetrahedra and is a discontinuous polynomial vector field of degree at most k , it belongs to V_h . Applying (4) on any tetrahedron to $\vec{\varphi} = \vec{\nabla} \psi$ and noticing that $\text{curl} \vec{\nabla} \psi = 0$, we get:

$$\begin{aligned} & - \frac{1}{\Delta t} \int_{T_i} \vec{\nabla} \psi \cdot \vec{\varepsilon}_i (\vec{\mathbf{E}}_i^{n+1} - \vec{\mathbf{E}}_i^n) \\ &= \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \in \mathcal{F}_h^{\text{int}}}} \int_{a_{ik}} \vec{\nabla} \psi \cdot (\{\vec{\mathbf{H}}_h^{n+\frac{1}{2}}\}_{ik} \times \vec{n}_{ik}) + \sum_{\substack{k \in \mathcal{V}_i \\ a_{ik} \subset \partial \Omega}} \int_{a_{ik}} \vec{\nabla} \psi \cdot (\vec{\mathbf{H}}_i^{n+\frac{1}{2}} \times \vec{n}_{ik}). \end{aligned}$$

Summing over all tetrahedra yields

$$\frac{1}{\Delta t} \langle \operatorname{div} \left(\bar{\varepsilon}(\vec{\mathbf{E}}_h^{n+1} - \vec{\mathbf{E}}_h^n) \right), \psi \rangle_{-1} = \int_{\mathcal{F}_h^{\text{int}}} \{ \vec{\mathbf{H}}_h^{n+\frac{1}{2}} \} \cdot \llbracket \vec{\nabla} \psi \rrbracket + \int_{\partial\Omega} \vec{\mathbf{H}}_h^{n+\frac{1}{2}} \cdot (\vec{n} \times \vec{\nabla} \psi).$$

Since ψ is continuous, the tangential jump $\llbracket \vec{\nabla} \psi \rrbracket \equiv 0$ obviously vanishes on internal faces. Similarly, since $\psi = 0$ on $\partial\Omega$, $\vec{\nabla} \psi \times \vec{n} = 0$ on $\partial\Omega$. This leads to the result for the electric field. The proof follows the same line for the magnetic field.

Now, assume that $\operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_0 = 0$ and $\operatorname{div} \bar{\mu} \vec{\mathbf{H}}_0 = 0$. The discrete initial data are given by $\vec{\mathbf{E}}_h^0 = P_h(\vec{\mathbf{E}}_0)$ and $\vec{\mathbf{H}}_h^{1/2} = P_h(\vec{\mathbf{H}}_0)$. Hence,

$$\langle \operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h^0, \psi \rangle_{-1} = - \int_{\Omega} P_h(\vec{\mathbf{E}}_0) \cdot \bar{\varepsilon} \nabla \psi = - \int_{\Omega} \bar{\varepsilon} \vec{\mathbf{E}}_0 \cdot \nabla \psi = 0$$

since $\bar{\varepsilon} \nabla \psi$ belongs to V_h for piecewise constant $\bar{\varepsilon}$ and P_h is a projector. □

In the same way, we prove that the semi-discrete approximation $(\vec{\mathbf{E}}_h(t), \vec{\mathbf{H}}_h(t))$ is discrete divergence-free:

Proposition 3.8. *Let X_h be the vector space defined in Proposition 3.7. The solution $(\vec{\mathbf{E}}_h(t), \vec{\mathbf{H}}_h(t))$ of the semi-discrete scheme (17)-(18) satisfies*

$$\forall \psi \in X_h, \begin{cases} \langle \operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h(t), \psi \rangle_{-1} = 0 \\ \langle \operatorname{div} \bar{\mu} \vec{\mathbf{H}}_h(t), \psi \rangle_{-1} = 0 \end{cases} \tag{34}$$

for all $t \in [0, T]$ provided the initial data $\vec{\mathbf{E}}_0$ and $\vec{\mathbf{H}}_0$ are divergence-free.

Proposition 3.9. *Let the initial data $\vec{\mathbf{E}}_0$ and $\vec{\mathbf{H}}_0$ be divergence-free. Under the assumptions of Theorem 3.5, we have the following estimate for the divergence of the semi-discrete solution*

$$\max_{t \in [0, T]} \left(\|\operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h(t)\|_{-1, \Omega} + \|\operatorname{div} \bar{\mu} \vec{\mathbf{H}}_h(t)\|_{-1, \Omega} \right) \leq CT h^{\min(s, k)} \|(\vec{\mathbf{E}}, \vec{\mathbf{H}})\|_{C^0((0, T), \mathcal{PH}^{s+1}(\Omega))}. \tag{35}$$

Proof. Let $\psi \in C_0^\infty(\Omega)$ and denote by $\Pi_h \psi$ its interpolate in X_h in the sense of Lagrange finite elements of type P^{k+1} . Due to Proposition 3.8 we have

$$\langle \operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h(t), \psi \rangle_{-1} = - \int_{\Omega} \nabla(\psi - \Pi_h \psi) \cdot \bar{\varepsilon} \vec{\mathbf{E}}_h(t).$$

Next, taking into account that $\psi - \Pi_h \psi \in H_0^1(\Omega)$, we get by integration by parts on each tetrahedron

$$\begin{aligned} - \int_{\Omega} \nabla(\psi - \Pi_h \psi) \cdot \bar{\varepsilon} \vec{\mathbf{E}}_h(t) &\leq \sum_{\mathcal{T}_i} \|\operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h(t)\|_{0, \mathcal{T}_i} \|\psi - \Pi_h \psi\|_{0, \mathcal{T}_i} \\ &+ \sum_{a_{ik} \subset \mathcal{F}_h^{\text{int}}} \|\llbracket \bar{\varepsilon} \vec{\mathbf{E}}_h(t) \cdot \vec{n} \rrbracket\|_{0, a_{ik}} \|\psi - \Pi_h \psi\|_{0, a_{ik}} \\ &\leq Ch \left(\sum_{\mathcal{T}_i} \|\operatorname{div} \bar{\varepsilon} \vec{\mathbf{E}}_h(t)\|_{0, \mathcal{T}_i}^2 \right)^{1/2} |\psi|_{1, \Omega} \\ &+ Ch^{1/2} \left(\sum_{a_{ik} \subset \mathcal{F}_h^{\text{int}}} \|\llbracket \bar{\varepsilon} \vec{\mathbf{E}}_h(t) \cdot \vec{n} \rrbracket\|_{0, a_{ik}}^2 \right)^{1/2} |\psi|_{1, \Omega}, \end{aligned}$$

where the last estimate follows from the approximation properties of Π_h (see Lemma 3.3). Applying the inverse estimate (28) and the convergence Theorem 3.5 yields, after division by $|\psi|_{1,\Omega}$, the result of the proposition (and to the one of [11] as well). \square

4. NUMERICAL EXAMPLES FOR LOW-ORDER METHODS

We have implemented the low order \mathbb{P}_0 -DGTD and \mathbb{P}_1 -DGTD methods on tetrahedral meshes. The corresponding solvers were parallelized according to a widely used SPMD (single program multiple data) strategy that combines a partitioning of the underlying mesh (element-wise decomposition and minimum overlap of sub-domain artificial interfaces) and a message passing programming model (MPICH implementation of the message passing interface).

In this section, we first compare the sufficient stability conditions obtained for both methods as well as memory requirements. We then present numerical results in order to validate and compare the two methods.

4.1. A comparison of the \mathbb{P}_0 - and \mathbb{P}_1 -DGTD methods on tetrahedral meshes

For tetrahedral discretizations, the \mathbb{P}_0 -DGTD method is exactly the classical finite volume approach proposed in [21], for which assumptions (11) are verified with $\forall i \in \mathcal{G}, \alpha_i = 0$ and $\forall i \in \mathcal{G}, \forall k \in \mathcal{V}_i, \beta_{ik} = 1$. In that case, Theorem 2.10 states that the finite volume method is stable, if

$$\forall i, \forall k \in \mathcal{V}_i, c_i \Delta t \left[\max \left(\sqrt{\frac{\mu_i}{\mu_k}}, \sqrt{\frac{\epsilon_i}{\epsilon_k}} \right) \right] < \frac{4V_i}{P_i}.$$

This condition is slightly more restrictive than the condition obtained in the less general context of [21].

Similarly, the basis vector fields considered in the \mathbb{P}_1 -DGTD method are simply \mathbb{P}_1 fields inside the tetrahedron. This leads to twelve degrees of freedom for each field inside each tetrahedron (three components times four \mathbb{P}_1 scalar basis functions). Therefore, we have $\forall i \in \mathcal{G}, d_i = 12$. We have chosen to limit our first implementation to isotropic materials which are also homogeneous inside each tetrahedron. Because of this simple choice, exact integrations were performed for volume integrals over tetrahedra and surface integrals over faces.

Lemma 4.1. *For this \mathbb{P}_1 -DGTD method, assumptions (11) are verified with $\forall i, \alpha_i^2 = \frac{20}{9} \max_{k \in \mathcal{V}_i} (S_{ik}/P_i)$ and $\forall i, \forall k \in \mathcal{V}_i, \beta_{ik} = 8/3$.*

Proof. Let us consider the standard \mathbb{P}_1 scalar basis functions φ_{ij} (equal to 0 on a_{ij} and 1 on the other vertex; then degrees of freedom can be grouped into vectors, related to a given vertex, or equivalently to a given opposite neighboring tetrahedron). We have:

$$\int_{\mathcal{T}_i} \varphi_{ij} \varphi_{ij'} = (1 + \delta_{jj'}) V_i / 20, \quad \int_{a_{ik}} \varphi_{ij} \varphi_{ij'} = (1 - \delta_{kj})(1 - \delta_{kj'})(1 + \delta_{jj'}) S_{ik} / 12.$$

Then, for any \mathbb{P}_1 field $\vec{\mathbf{X}} = \sum_{k \in \mathcal{V}_i} \vec{X}_{ik} \varphi_{ik}$, $\|\vec{\mathbf{X}}\|_{\mathcal{T}_i}^2 \geq (V_i/20) \sum_{k \in \mathcal{V}_i} \|\vec{X}_{ik}\|^2$. Also,

$$\begin{aligned} \|\text{curl } \vec{\mathbf{X}}\|_{\mathcal{T}_i}^2 &= \frac{1}{9V_i} \left(\sum_{k \in \mathcal{V}_i} S_{ik} \vec{n}_{ik} \times \vec{X}_{ik} \right)^2 = \frac{P_i^2}{9V_i} \left(\sum_{k \in \mathcal{V}_i} \frac{S_{ik}}{P_i} \vec{n}_{ik} \times \vec{X}_{ik} \right)^2 \\ &\leq \frac{P_i^2}{9V_i} \sum_{k \in \mathcal{V}_i} \frac{S_{ik}}{P_i} \left(\vec{n}_{ik} \times \vec{X}_{ik} \right)^2 \leq \frac{P_i}{9V_i} \max_{k \in \mathcal{V}_i} (S_{ik}) \sum_{k \in \mathcal{V}_i} \|\vec{X}_{ik}\|^2, \end{aligned}$$

which leads to the given value for α_i of Definition 11 in the lemma (the value is optimal). Concerning the β_{ik} , they are all identical and can be obtained explicitly by a simple minimization problem. \square

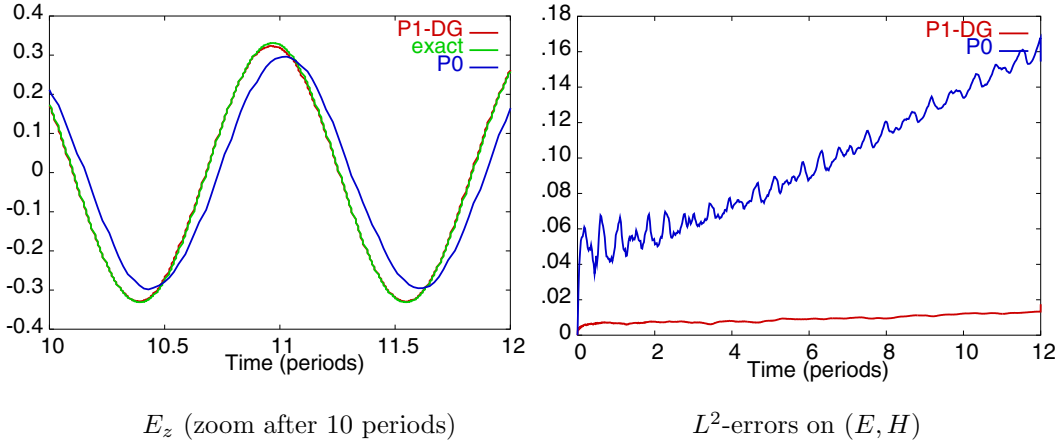


FIGURE 1. Cubic cavity: \mathbb{P}_1 -DGTD, \mathbb{P}_0 , and exact solutions.

These results lead to the following sufficient stability condition for our \mathbb{P}_1 -DGTD method on unstructured tetrahedral meshes:

$$\forall i, \forall k \in \mathcal{V}_i, c_i \Delta t \left[\frac{4\sqrt{5}}{3} \sqrt{\frac{\max_{k \in \mathcal{V}_i}(S_{ik})}{P_i}} + \frac{8}{3} \max \left(\sqrt{\frac{\mu_i}{\mu_k}}, \sqrt{\frac{\epsilon_i}{\epsilon_k}} \right) \right] < \frac{4V_i}{P_i}.$$

Noticing that $\max_{k \in \mathcal{V}_i}(S_{ik}) \leq P_i/2$, this expression can be simplified, for example in the case of an homogeneous medium, into the following slightly more restrictive sufficient condition:

$$\forall i, \forall j \in \mathcal{V}_i, c_i \Delta t \frac{8 + \sqrt{40}}{3} < \frac{4V_i}{P_i}.$$

This means that, theoretically, the limit possible time step for the \mathbb{P}_1 -DGTD method on unstructured tetrahedral meshes should be roughly 4.7 times more restrictive than the FVTD method.

4.2. Resonance in metallic cavities

We consider here resonance inside close metallic cavities, in order to verify energy conservation and to compare with analytical solution. At the same time, these computations where numerical errors are accumulated are good benchmark problems for numerical methods.

We first compute the (1, 1, 1) mode which is a standing wave of 0.260 GHz frequency in a 1m-side cubic cavity. We use an unstructured grid of 16 464 tetrahedra and 3375 nodes which gives 13 points per wavelength. We plot in Figure 1 (left) the time evolution of E_z after ten periods at a fixed point in the cavity. One can see that the \mathbb{P}_1 -DG solution compares very well with the exact one, with a much smaller dispersion error compared to the \mathbb{P}_0 (finite volume) solution (we recall that the finite volume scheme has the same order of dispersion error as the Yee scheme [22]). The overall L^2 -error on the electromagnetic field (E, H) of the \mathbb{P}_1 -DGTD and \mathbb{P}_0 approximate solutions are plotted in Figure 1 (right). The errors are increasing in time because of the dispersion and the level of dispersion is a lot smaller for the \mathbb{P}_1 -DGTD method. Figure 2 shows some contours of the electric and magnetic fields respectively, for the exact and \mathbb{P}_1 -DGTD solutions in the cut plane $x + y + z = 1.5$. A similar experiment was done for the lowest (0, 1, 1) TE mode in a spherical cavity (1m radius). The resonant frequency is 0.21 GHz and the 81 920-tetrahedron 14 993-vertex mesh with an average edge length of 56 mm (which corresponds to an average of 10 points per wavelength for the frequency 0.53 GHz). We compare in Figure 3 (left) the time evolution of the H_z component of the exact and computed magnetic field during seven

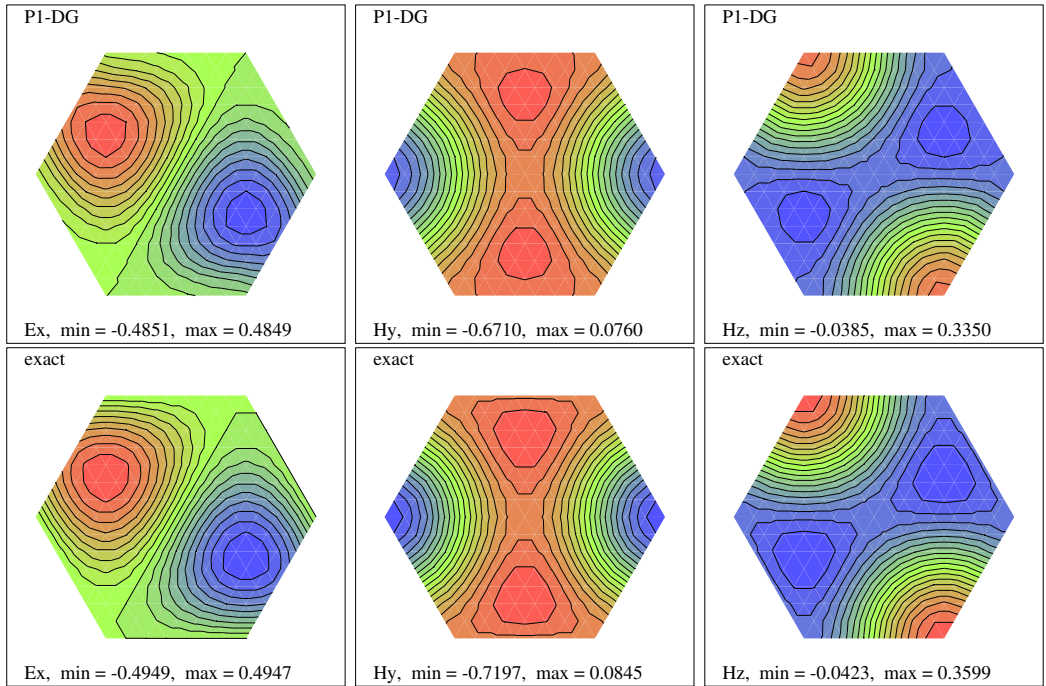


FIGURE 2. Cubic cavity: \mathbb{P}_1 -DGTD and exact solutions (plane $x + y + z = 1.5$).

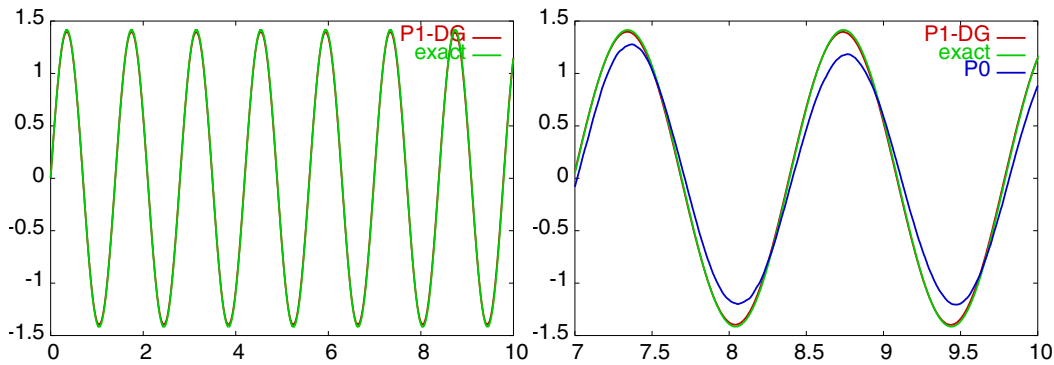


FIGURE 3. Spherical cavity: H_z in \mathbb{P}_1 -DGTD and exact solutions - first 7 periods (left) and zoom after 5 periods (right).

periods. One may see again that the two solutions compare very well. These solutions are compared to the \mathbb{P}_0 approximate solution in Figure 3 (right) and the gain in accuracy in favor of the \mathbb{P}_1 -DGTD solution is obvious. Figure 4 shows contours of the computed and exact magnetic field respectively in the plane $z = 0$.

We have also considered the transient simulation of waves inside the cubic and spherical cavities for an initial Gaussian pulse, in order to evaluate the ability of both methods to recover the resonating eigenfrequencies. The normalized discrete Fourier transforms of the signal obtained with the \mathbb{P}_1 -DGTD method (vertical electric field at some observation point) are shown in Figure 5 for the cube and Figure 6 for the sphere. The signals obtained with the \mathbb{P}_0 -DGTD were too noisy to allow a good resolution for the frequencies. Results shown for the cube are quite accurate and eigenfrequencies are well captured up to 700 MHz, which correspond for this mesh (made of cubes divided into tetrahedra) to six points per wavelength at this frequency. Results shown for

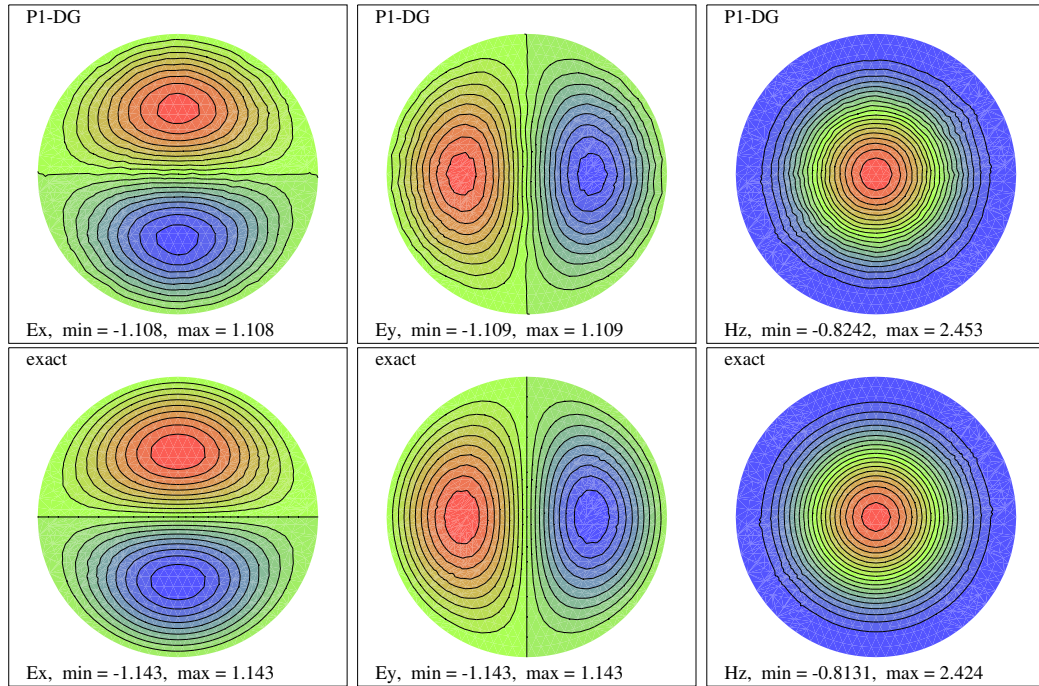


FIGURE 4. Spherical cavity (plane $z = 0$): E_x , E_y , and H_z for \mathbb{P}_1 -DGTD and exact solutions.

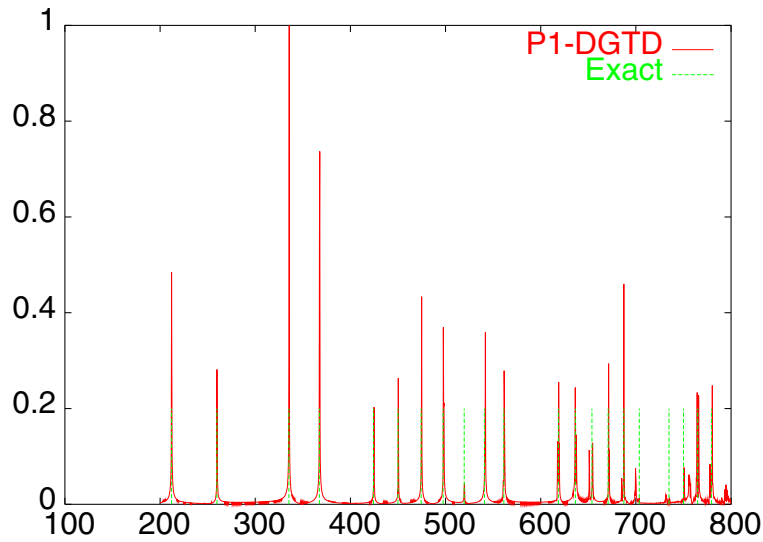


FIGURE 5. Cubic cavity: Fourier transforms for the \mathbb{P}_1 -DGTD approximate solution (amplitude in function of the frequency in MHz).

the sphere correspond to the same mesh (14 993 vertices, 81 920 tetrahedra). All frequencies are well captured up to somewhere near 0.5 GHz, which corresponds to ten points per wavelength.

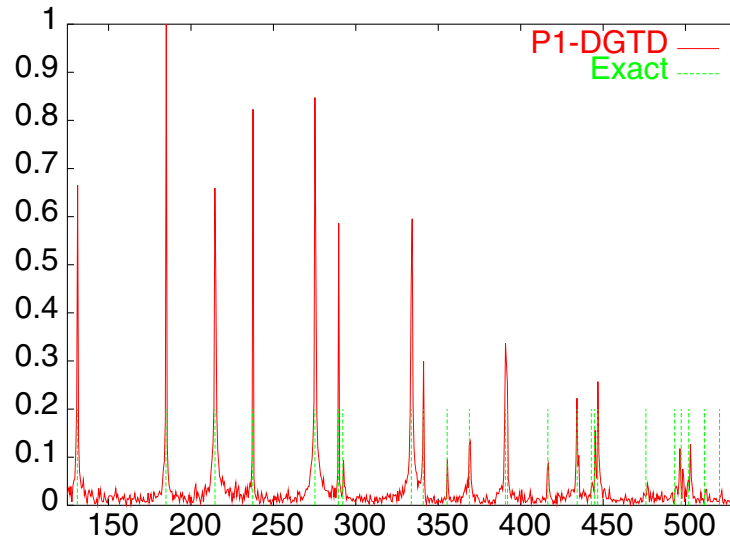


FIGURE 6. Spherical cavity: Fourier transforms for the \mathbb{P}_1 -DGTD approximate solution (amplitude in function of the frequency in MHz).

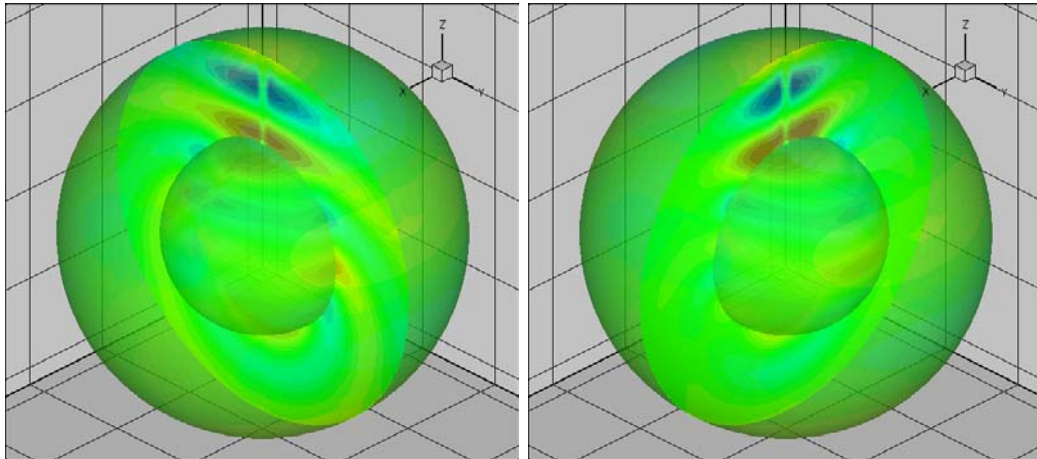


FIGURE 7. Component E_x of the electric field (cuts in the $x = 0$ and $y = 0$ planes).

4.3. Scattering by a metallic sphere

We use a 43 932 node tetrahedral grid with 244 608 tetrahedra to mesh the exterior of a one meter radius perfectly conducting sphere. The mesh is quite uniform, with 13 points per wavelength in the radial direction. In Figures 7 and 8, we have respectively shown contours of the E_x and H_y components scattered by a Ox polarized incident plane wave propagating along the Oz axis at a frequency of 0.3 GHz (which corresponds to $ka = 2\pi$). Let us note that the absorbing boundary is set at only one wavelength away from the sphere. In Figure 9, we have shown both analytic and computed radar cross sections (RCS) for the metallic sphere under the selected incident wave. One can see that both RCS compare very well.

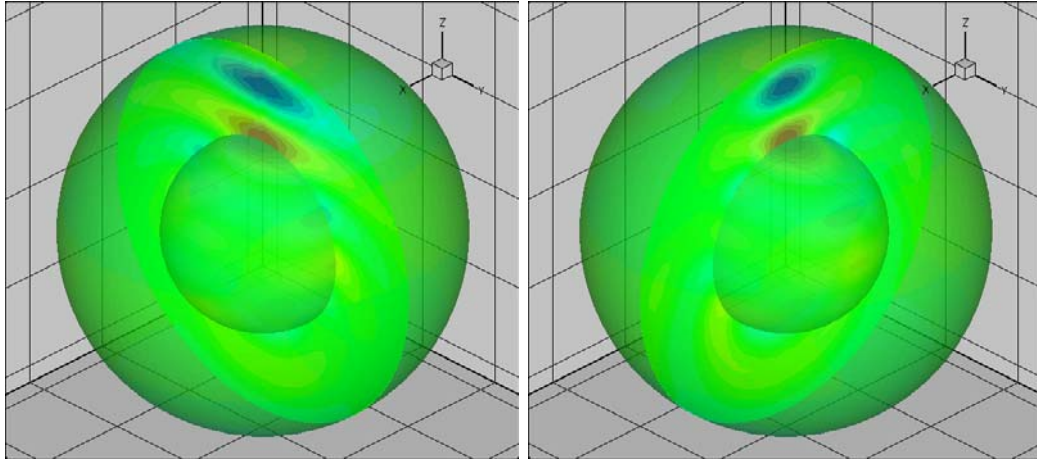


FIGURE 8. Component H_y of the magnetic field (cuts in the $x = 0$ and $y = 0$ planes).

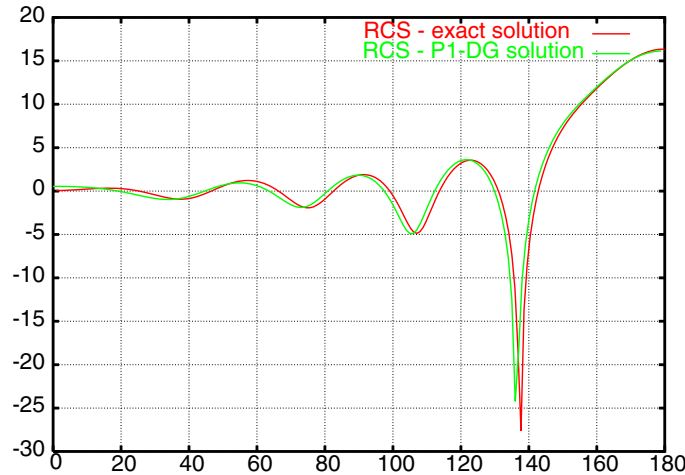


FIGURE 9. RCS of the exact and approximate solutions ($\theta = 0, 0 \leq \phi \leq 180$).

4.4. Scattering by a dielectric sphere

The problem considered here deals with the propagation of a 1 GHz Gaussian pulse past a dielectric ball. The computational domain is setup has a dielectric ball (radius 0.138 m, $\epsilon_r = 4.431$) surrounded by air and embedded in a outer sphere of radius 0.438 m delimiting the computational domain and thus supporting the first-order absorbing boundary condition. The Gaussian pulse is defined as $G(x, t) = G_0 \exp(-[t - t_r - (x - x_r)/c_0]^2/\alpha_r^2)$ with $G_0 = 10^2$ V/m, $\alpha_r = 0.7$ ns, $t_r = 1.5$ ns, and $x_r = -0.3$ m. The simulation has been carried out for a physical duration of 15 ns. Two unstructured (yet regular) tetrahedral meshes have been constructed: a coarse one named MS1, with 39 091 vertices and 229 873 tetrahedra, and a finer one named MS2, with 163 132 vertices and 968 027 tetrahedra. The coarser grid (resp. the finer grid) is meshed with around 11 points per wavelength (resp. 17) in the dielectric sphere and 12 points per wavelength (resp. 20) outside. The surfacic triangular mesh of the inner sphere for the coarser mesh is shown in Figure 10, where we also give contour lines of the maximum of the electric field amplitude: $\mathbf{E}_{\max}(\vec{x}_i) = \max_{t \in [t_0, t_e]} |\vec{\mathbf{E}}(\vec{x}_i, t)|$ on the surface of the inner sphere. We note that the \mathbb{P}_1 -DGTD method produces a smoother solution even on a coarser mesh while the mesh refinement

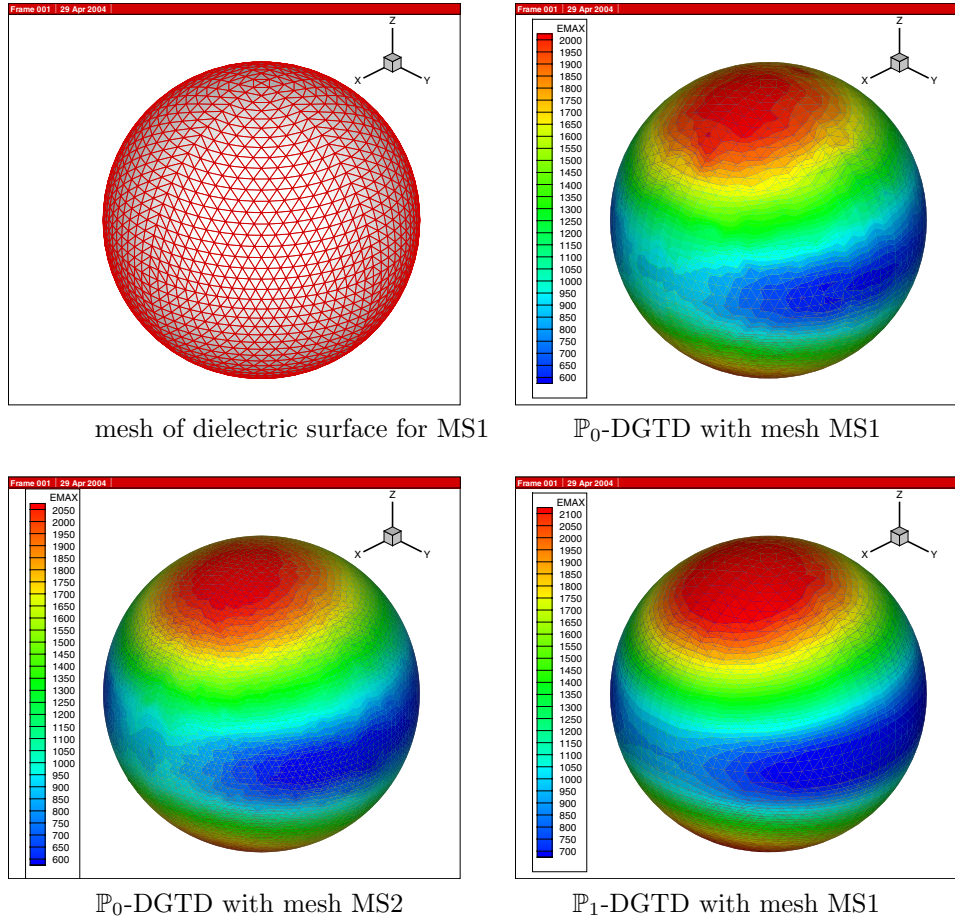


FIGURE 10. Gaussian pulse past a dielectric sphere.

considered here is not sufficient to obtain an acceptable convergence with the \mathbb{P}_0 -DGTD method. For the sake of completeness, we compare in Figure 11 the time evolutions of E_z at a specific point location in the free space (ahead of the dielectric sphere), for numerical simulations performed with a FDTD method and with the two DGTD methods considered here. With such comparisons, the advantage of the \mathbb{P}_1 -DGTD method over the \mathbb{P}_0 -DGTD method is less remarkable.

Finally, we give some informations on the simulation times required by both methods. On a cluster of PC computers (Intel Pentium 4.2 GHz biprocessors with 1 Gb of RDRAM memory) with a Gigabit Ethernet switch, on four subdomains, the \mathbb{P}_0 -DGTD required 570 s on the coarse mesh MS1 and 2326 s on the finer mesh MS2, while the \mathbb{P}_1 -DGTD required 5492 s on the coarse mesh MS1. We obtain that in this case the \mathbb{P}_1 -DGTD method is ten times more costly for the given mesh. However, the \mathbb{P}_0 -DGTD would have required a very fine mesh and a huge computation time to obtain the same kind of accuracy.

4.5. Scattering of a plane wave by an aircraft

We consider a problem involving a complex geometry and a fully unstructured tetrahedral mesh with the numerical simulation of the scattering of plane wave by an aircraft geometry (Airbus A318, courtesy of EADS). A partial view of the surfacic mesh is shown in Figure 12. The underlying mesh is made of 225 326 vertices and 1 305 902 tetrahedra. The minimum, maximum and average lengths of the mesh edges are respectively given by $l_{\min} = 0.706$ mm, $l_{\max} = 0.119$ m and $l_{\text{avg}} = 0.037$ m. The frequency of oscillation of the plane wave

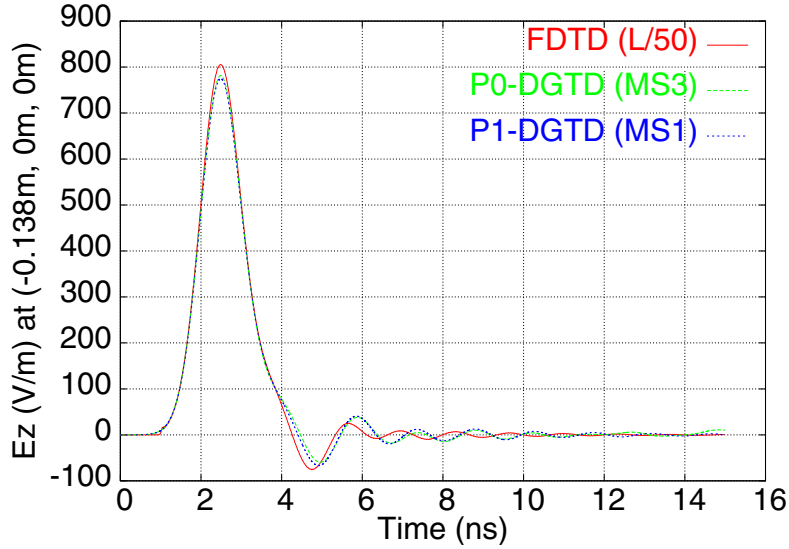


FIGURE 11. Gaussian pulse past a dielectric sphere: time evolutions of E_z .

TABLE 1. Scattering by an aircraft: performance results (for time integration over a period).

Method	N_p	CPU	REAL	% CPU
\mathbb{P}_0 -DGTD	24	214 s	235 s	91%
\mathbb{P}_1 -DGTD	24	2536 s	2727 s	86%

is 565 MHz which corresponds to a wavelength $\lambda = 0.531$ m. Note that $\lambda/l_{\text{avg}} \approx 14$. The polarization of the plane wave is such that the wave vector is $\vec{k} = k\vec{e}_x$ (with $\vec{E} \times \vec{e}_z = 0$). Finally, the absorbing boundary is defined by a cube $[-1.4 \text{ m}, 1.4 \text{ m}]^3$. For this problem, it has not been possible to obtain a sufficiently converged solution with the \mathbb{P}_0 -DGTD method. The solution obtained with the \mathbb{P}_1 -DGTD method is shown in Figure 12 in terms of the contour lines of the quantity $j_{\text{max}}(\vec{x}_i) = \max_{t \in [t_0, t_e]} |\vec{n} \times \vec{H}(\vec{x}_i, t)|$ for mesh vertices located on the surface of the aircraft. Finally, performance results are given in Table 1.

5. CONCLUSION AND FURTHER WORKS

We presented a Discontinuous Galerkin formulation applied to the time domain Maxwell equations. The method is based on centered numerical fluxes and a leap-frog time-scheme. It can be seen as a simplified version of methods found in the literature [5] and an extension of centered finite volume schemes [22]. Beyond its convergence properties, the framework proposed herein presents two other important features. First, some discrete equivalent of the electromagnetic energy is conserved (away from absorbing boundaries) on any kind of mesh and discontinuous elements, including hp -type or non-conformal refinement. This also leads to the proof of the stability under some CFL-type stability condition, including for simple absorbing boundary conditions. Second, for some particular choices of discontinuous finite elements, some discrete divergence preservation property is achieved, justifying the transient solution of the hyperbolic part of Maxwell equations only.

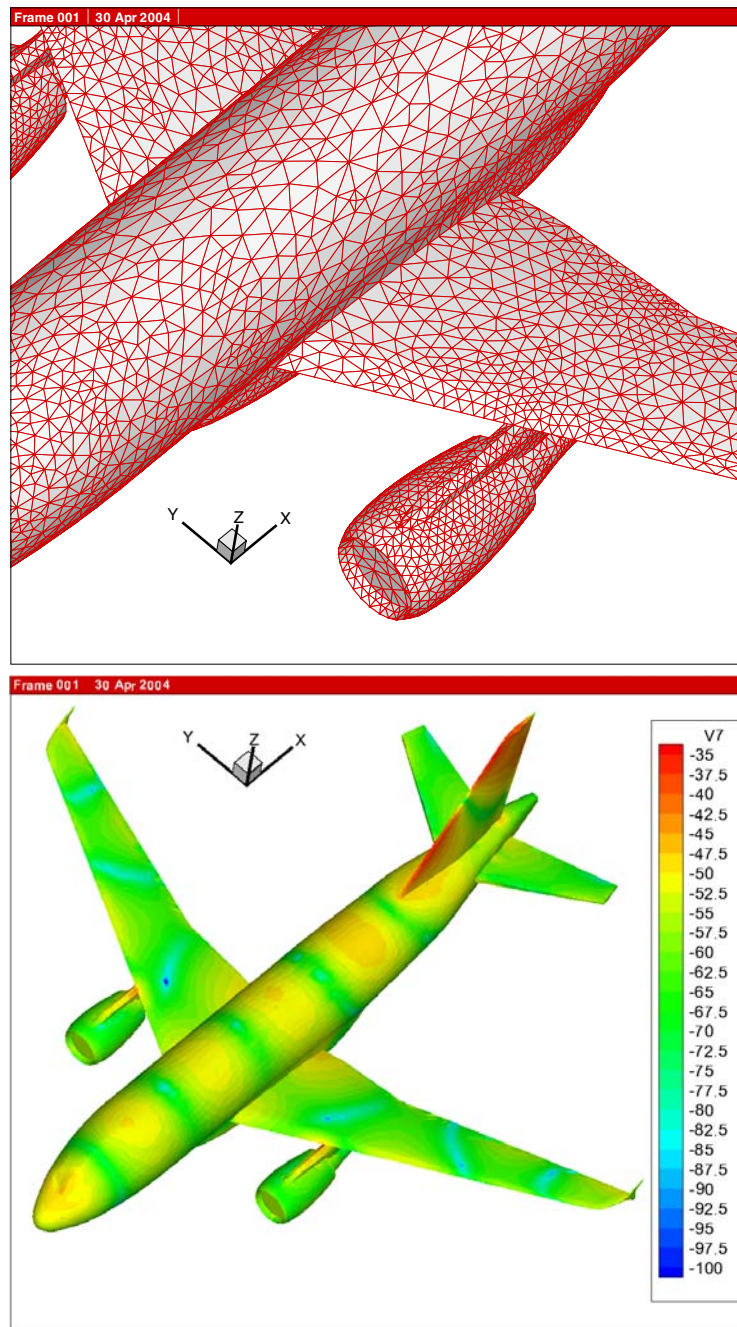


FIGURE 12. Scattering by an aircraft: surfacic mesh and contour lines of $j_{\max}(\vec{x}_i)$.

The three-dimensional numerical results presented here, obtained with a parallel implementation of the \mathbb{P}_1 -DGTD method on unstructured tetrahedral meshes, are encouraging and should be developed in a near future in several interesting directions. For some particular geometries (with very refined devices), non-conforming local refinement could be used as was done for antenna design [2]. Also, local time-stepping should be investigated.

It has been shown it is possible to build totally-explicit energy-conserving partitioned procedures allowing time-stepping [20]. The use of leap-frog-type time-scheme makes it really difficult, compared to partitioned explicit-implicit Runge-Kutta schemes. Finally, and most importantly, extensions to higher accuracy in time (using a fourth-order energy-conserving time-scheme [20]) and space (using modal or nodal basis functions [11] not leading to ill-conditioned mass and stiffness matrices) are to be considered.

REFERENCES

- [1] F. Bourdel, P.-A. Mazet and P. Helluy, Resolution of the non-stationary or harmonic Maxwell equations by a discontinuous finite element method. Application to an E.M.I. (electromagnetic impulse) case. *Comput. Method Appl. Sci. Engrg.* (1991) 405–422.
- [2] N. Canouet, L. Fezoui and S. Piperno, A discontinuous galerkin method for 3d maxwell’s equation on non-conforming grids, in *Sixth International Conference on Mathematical and Numerical Aspects of Wave Propagation*. G.C. Cohen Ed., Springer, Jyväskylä, Finland (2003) 389–394.
- [3] P. Ciarlet, *The Finite Element Method for Elliptic Problems*. North Holland-Elsevier Science Publishers, Amsterdam, New York, Oxford (1978).
- [4] J.-P. Cioni, L. Fezoui, L. Anne and F. Poupaud, A parallel FVTD Maxwell solver using 3D unstructured meshes, in *13th annual review of progress in applied computational electromagnetics*, Monterey, California (1997) 359–365.
- [5] B. Cockburn, G.E. Karniadakis and C.-W. Shu, Eds., Discontinuous Galerkin methods. Theory, computation and applications. *Lect. Notes Comput. Sci. Eng.* **11** (2000).
- [6] B. Cockburn, F. Li and C.-W. Shu, Locally divergence-free discontinuous galerkin methods for the maxwell equations. *J. Comput. Phys.* **194** (2004) 588–610.
- [7] A. Elmkies and P. Joly, Éléments finis d’arête et condensation de masse pour les équations de Maxwell: le cas de dimension 3. *C. R. Acad. Sci. Paris Sér. I Math.* **325** (1997) 1217–1222.
- [8] R. Eymard, T. Gallouët and R. Herbin, The finite volume method, *Handbook Numer. Anal.*, North Holland-Elsevier Science Publishers, Amsterdam, New York, Oxford **7-3** (2000).
- [9] R.S. Falk and G.R. Richter, Explicit finite element methods for symmetric hyperbolic equations. *SIAM J. Numer. Anal.* **36** (1998) 935–952.
- [10] J. Hesthaven and C. Teng, Stable spectral methods on tetrahedral elements. *SIAM J. Sci. Comput.* **21** (2000) 2352–2380.
- [11] J. Hesthaven and T. Warburton, Nodal high-order methods on unstructured grids. I: Time-domain solution of Maxwell’s equations. *J. Comput. Phys.* **181** (2002) 186–221.
- [12] J. Hesthaven and T. Warburton, High-order nodal discontinuous galerkin methods for the maxwell eigenvalue problem. *Philos. Trans. Roy. Soc. London Ser. A* **362** (2004) 493–524.
- [13] J.M. Hyman and M. Shashkov, Mimetic discretizations for Maxwell’s equations. *J. Comput. Phys.* **151** (1999) 881–909.
- [14] P. Joly and C. Poirier, A new second order 3D edge element on tetrahedra for time dependent Maxwell’s equations, in *Fifth International Conference on Mathematical and Numerical Aspects of Wave Propagation*, A. Bermudez, D. Gomez, C. Hazard, P. Joly, J.-E. Roberts, Eds., SIAM, Santiago de Compostella, Spain (2000) 842–847.
- [15] D.A. Kopriva, S.L. Woodruff and M.Y. Hussaini, Discontinuous spectral element approximation of Maxwell’s equations, in *Discontinuous Galerkin methods. Theory, computation and applications.*, B. Cockburn and G.E. Karniadakis, C.-W. Shu, Eds. *Lect. Notes Comput. Sci. Eng.* **11** (2000) 355–362.
- [16] D. Kröner, M. Rokyta and M. Wierse, A Lax-Wendroff type theorem for upwind finite volume schemes in 2-D. *J. Numer. Math.* **4** (1996) 279–292.
- [17] S. Lohrengel and M. Remaki, A FV scheme for Maxwell’s equations: Convergence analysis on unstructured meshes, in *Finite Volumes for Complex Applications III*, R. Herbin, D. Kröner, Eds., Hermes Penton Science, London, Porquerolles, France (2002) 219–226.
- [18] F. Poupaud and M. Remaki, Existence and uniqueness of the Maxwell’s system solutions in heterogeneous and irregular media. *C. R. Acad. Sci. Paris Sér. I Math.* **330** (2000) 99–103.
- [19] S. Piperno, L^2 -stability of the upwind first order finite volume scheme for the Maxwell equation in two and three dimensions on arbitrary unstructured meshes. *ESAIM: M2AN* **34** (2000) 139–158.
- [20] S. Piperno, *Schémas en éléments finis discontinus localement raffinés en espace et en temps pour les équations de Maxwell 1D*. INRIA Research report **4986** (2003).
- [21] S. Piperno, M. Remaki and L. Fezoui, A non-diffusive finite volume scheme for the 3D Maxwell equations on unstructured meshes. *SIAM J. Numer. Anal.* **39** (2002) 2089–2108.
- [22] M. Remaki, A new finite volume scheme for solving Maxwell’s system. *COMPEL* **19** (2000) 913–931.

- [23] J. Shang and R. Fithen, A comparative study of characteristic-based algorithms for the Maxwell equations. *J. Comput. Phys.* **125** (1996) 378–394.
- [24] A. Taflove, *Re-inventing electromagnetics: supercomputing solution of Maxwell's equations via direct time integration on space grids*. AIAA paper 92–0333 (1992).
- [25] J.-P. Vila, Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws. I. Explicit monotone schemes. *RAIRO Modél. Math. Anal. Numér.* **28** (1994) 267–295.
- [26] T. Warburton, Application of the discontinuous Galerkin method to Maxwell's equations using unstructured polymorphic *hp*-finite elements, in Discontinuous Galerkin methods. Theory, computation and applications, B. Cockburn, G.E. Karniadakis, C.-W. Shu, Eds. *Lect. Notes Comput. Sci. Eng.* **11** (2000) 451–458.
- [27] K.S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE T. Antenn. Prop.* **AP-16** (1966) 302–307.