Accuracy of finite element approximations for two-dimensional time-harmonic electromagnetic boundary value problems involving non-conducting moving objects with stationary boundaries

Praveen Kalarickel Ramakrishnan∗ Mirco Raffetto†

September 17, 2017

Abstract

An analysis on the accuracy of the results computed using a finite element code in the presence of axially moving cylinders is presented. It seems that no results of this type is available in the open literature.

Any material in motion is perceived as a bianisotropic medium. This generates a scattered field having two components: one has the same polarization as the incident field and the other presents the orthogonal polarization. The results on the accuracy of the co-polarized field are new but are similar to those obtained in the presence of motionless media. The outcome on the accuracy of the results related to the orthogonal polarization seems to be more interesting, especially for the information content this component of the field could provide on the axial velocity profile. In particular, using a finite element simulator based on double precision arithmetic it is shown that the range of axial velocity values over which it is possible to obtain very accurate approximations can span nine or even more decades. This allows the use of the simulator, even when the more difficult components of the field are required to be accurate, for a set of applications ranging from astrophysics to medicine.

Keywords: electromagnetic scattering; time-harmonic electromagnetic fields; bianisotropic media; moving media; finite element method; error analysis; reconstruction of velocity profiles.

1 Introduction

The interaction of electromagnetic waves with moving bodies plays a role in many applications [1]. Among these one can consider many important applications involving only axially moving cylinders. In particular, one can refer to axially moving plasma columns [2], [3], [4], ionized meteor trails [5], jet exhausts [6] or mass flows in pneumatic pipes [7].

In most cases of interest, which could involve multiple cylinders having irregular shapes, inhomogeneous constitutive parameters and non-constant velocity profiles, numerical methods are required to try to approximate the solutions of interest [8].

∗Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, Via Opera Pia 11a, I–16145, Genoa, Italy, email: pravin.nitc@gmail.com
†Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, Via Opera Pia 11a, I–16145, Genoa, Italy, email: raffetto@dibe.unige.it
Notwithstanding the difficulty related to the presence of materials in motion [1], [9], [10], determining the appearance of bianisotropic materials in any reference frame in which the media themselves are not at rest, very recently the first results related to the well posedness and the finite element approximability for these problems have been deduced [11].

When simulators are exploited, the results related to the convergence of numerical approximations are not the only aspects of practical interest, however. Error estimates are important, too (particularly, a-priori error estimations), as clearly pointed out in [12] (p. 114). These estimations are not available so far, to the best of authors’ knowledge, for problems involving axially moving cylinders, like the ones of interest in this paper.

Then, in order to understand what can be expected in practice, we have to perform numerical experiments. Unfortunately, in the presence of axially moving cylinders no numerical analysis seems to be available. This is the reason why, in this paper, we present a lot of numerical results. They could suggest what can be expected in other cases and could be considered as benchmarks for any error estimate the research community will be able to deduce.

All numerical results which will be presented refer to a simple problem involving a moving cylinder. For such a problem an independent truncated-series solution can be found [3] (see also [13]). In this way we can evaluate the accuracy of our finite element approximations in terms of absolute and relative errors. The effect of all parameters involved in the definition of the simple problem of reference are studied.

As it has already been pointed out, in the presence of a time-harmonic illumination, the axial motion of cylinders determines, in any reference frame in which the media are not at rest, a bianisotropic effect and this, in turn, is responsible for the presence of a scattered field having both polarizations: the same as that of the incident field and the orthogonal one. For the co-polarized component of the field the outcome is absolutely stable. The errors for this component are, in particular, almost independent of the axial velocity and, then, assume almost the same values we get in the presence of motionless media. For the cross-polarized scattered wave the relative errors are very stable, as well, even though it is a-priori known that the previous results cannot be duplicated in this case, due to the fact that this component of the scattered field is known to go uniformly to zero as the axial speed of the scatterers becomes smaller and smaller [14]. Anyway, the errors on this component of the scattered field are indeed stable for a huge range of axial velocity values. By using double precision arithmetic this range can span nine or even more decades, so allowing the use of finite element simulators for velocities varying from a few centimeters per second to many thousands of kilometers per second. That means that the considered finite element method can be reliably used for a set of applications ranging from astrophysics to medicine. The same results show, in particular, that the indicated simulator can be exploited as a reliable solver of forward scattering problems in imaging procedures aiming at the reconstruction of axial velocity profiles [14] and this, by the way, was the initial motivation for our study.

The paper is organized as follows. In Section 2 the mathematical formulation of the problems of interest is recalled together with some of the results available in the open literature. Some new considerations on the properties of the finite element matrices in the presence of moving media are provided as well. In order to carry out the error analysis of interest, the definition of a test case is necessary. This is done in Section 3, where, in addition, a complete set of relevant absolute and relative errors is defined. The main section of the manuscript, dealing with the error analysis, is Section 4. Finally, before concluding the paper, some considerations on the convergence of two well known iterative methods are provided.
2 Mathematical formulation of the problem

The electromagnetic problems of interest in this paper are those in which axially moving cylinders (having parallel axes) are illuminated by a time-harmonic source or field. This class of problems has been studied in [11] and we refer to that paper for the definition of all details. Here we recall just the main points to let the readers understand the developments which will be presented in the final part of the section and in the next ones.

All our problems will present a cylindrical geometry and we denote by $z$ the axis of such a geometry. The time-harmonic sources and the inhomogeneous admittance boundary conditions involved are assumed to be independent of $z$, too, so that our problems can be formulated in a two-dimensional domain $\Omega$ contained in the $(x, y)$ plane. $\Gamma$ denotes the boundary of $\Omega$. $\mathbf{n}$ and $\mathbf{l}$ are the unit vectors orthogonal (pointing outward) and tangential to $\Gamma$, respectively. We have $\mathbf{n} \times \mathbf{l} = \hat{z}$.

The media involved in our problems can move in the axial direction with respect to the chosen reference frame. In such a frame a velocity field $v_z$, assumed to be time-invariant, is naturally defined, even though we will often refer to it in terms of the usual [15] (p. 525) real-valued normalized field $\beta = \frac{v_z}{c_0}$, being $c_0$ the speed of light in vacuum. Different linear, time-invariant and inhomogeneous materials can be modelled in our problems.

Under the indicated conditions all fields in all media will be time-harmonic, as the considered sources, and a factor $e^{j\omega t}$, common to all fields of interest, is assumed and suppressed.

Any material involved is isotropic in its rest frame and is there characterized by its relative permittivity $\varepsilon_r$, its relative permeability $\mu_r$ and its electric conductivity $\sigma$. In the following, any reference to $\varepsilon_r$, $\mu_r$ or $\sigma$ of a moving medium should be interpreted as a reference to the corresponding quantity when the medium is at rest. All moving media will be considered in any case to have $\sigma = 0$ (in order to avoid the difficulties related to the convective currents which could also become surface electric currents and to avoid difficulty related to the no-slip condition which, ultimately, prevents the possibility of using pure two-dimensional models [11]).

By using the subscript "t" to denote the field quantities transverse to the $z$ direction, the constitutive relations for the media in motion are [11]:

\[
D_t = \frac{1 + \mu_r \varepsilon_r - \zeta_1}{\varepsilon_0 \mu_0 \mu_r} E_t + \frac{\zeta_2}{\varepsilon_0 \mu_0 \mu_r} \hat{z} \times B_t, \tag{1}
\]

\[
D_z = \varepsilon_0 \varepsilon_r E_z, \tag{2}
\]

\[
H_t = \frac{\zeta_1}{\mu_0 \mu_r} B_t + \frac{\zeta_2}{\varepsilon_0 \mu_0 \mu_r} \hat{z} \times E_t, \tag{3}
\]

\[
H_z = \frac{1}{\mu_0 \mu_r} B_z, \tag{4}
\]

where

\[
\zeta_1 = \frac{1 - \mu_r \varepsilon_r \beta^2}{1 - \beta^2}, \tag{5}
\]

\[
\zeta_2 = \frac{\beta (\mu_r \varepsilon_r - 1)}{1 - \beta^2}. \tag{6}
\]

In order to be able to define the problems of interest and to talk of their finite element approximations, it is necessary to introduce some additional notations. $(L^2(\Omega))^n$ is the usual Hilbert space of square integrable vector fields on $\Omega$ with values in $\mathbb{C}^n$, $n = 2, 3$, and with scalar product given by $(\mathbf{u}, \mathbf{v})_{0, \Omega} = \int_{\Omega} \mathbf{v}^* \mathbf{u} \, dS$, where $\mathbf{v}^*$ denotes the conjugate transpose of the column
vector $\mathbf{v}$. For a given three-dimensional complex-valued vector field $\mathbf{A} = (A_z, A_y, A_x) \in (L^2(\Omega))^3$ we consider the operators $\text{curl}_{2D}$ and $\text{grad}_{2D}$, defined according to

$$\text{curl}_{2D}\mathbf{A}_t = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y},$$  

$$\text{grad}_{2D} A_z = \begin{pmatrix} \frac{\partial A_z}{\partial x} \\ \frac{\partial A_z}{\partial y} \end{pmatrix}.$$  

(7)

(8)

The transverse parts of the electric and magnetic fields will be in the Hilbert space

$$U_{2D} = \{ \mathbf{A}_t \in (L^2(\Omega))^2 \mid \text{curl}_{2D}\mathbf{A}_t \in L^2(\Omega) \text{ and } \mathbf{A}_t \cdot \mathbf{1} \in L^2(\Gamma) \},$$

(9)

whose inner product is given by

$$(\mathbf{u}_t, \mathbf{v}_t)_{U_{2D}} = (\mathbf{u}_t, \mathbf{v}_t)_{0,\Omega} + (\text{curl}_{2D}\mathbf{u}_t, \text{curl}_{2D}\mathbf{v}_t)_{0,\Omega} + (\mathbf{u}_t \cdot \mathbf{1}, \mathbf{v}_t \cdot \mathbf{1})_{0,\Gamma}.$$  

(10)

The axial components of the same fields are in the Hilbert space

$$H^1(\Omega) = \{ A_z \in L^2(\Omega) \mid \text{grad}_{2D} A_z \in (L^2(\Omega))^2 \},$$

(11)

whose inner product is

$$(u_z, v_z)_{1,\Omega} = (u_z, v_z)_{0,\Omega} + (\text{grad}_{2D} u_z, \text{grad}_{2D} v_z)_{0,\Omega}.$$  

(12)

Overall the electric and magnetic fields are in the Hilbert space

$$U = U_{2D} \times H^1(\Omega)$$

(13)

with inner product given by

$$(\mathbf{u}, \mathbf{v})_U = (\mathbf{u}_t, \mathbf{v}_t)_{U_{2D}} + (u_z, v_z)_{1,\Omega}.$$  

(14)

$\| U \|$ will denote the corresponding norm. The norms of the different spaces so far introduced will be of particular interest in establishing the accuracy of the results of the finite element simulator considered.

After some cumbersome deductions, from the usual problem formulation [11], based on Maxwell’s equations, boundary conditions and constitutive relations, one can deduce the equivalent variational formulation [11]: given $\omega > 0$, the electric and magnetic current densities $\mathbf{J}_e, \mathbf{J}_m \in (L^2(\Omega))^3$ and the known terms, $f_R, f_I \in L^2(\Gamma)$, involved in the admittance boundary conditions considered, find $\mathbf{E} \in U$ such that

$$a(\mathbf{E}, \mathbf{w}) = l(\mathbf{w}) \quad \forall \mathbf{w} \in U,$$

(15)

where $a$ is the following sesquilinear form

$$a(\mathbf{u}, \mathbf{w}) = \left( \frac{\zeta_1}{\mu_r} \text{grad}_{2D} u_z, \text{grad}_{2D} w_z \right)_{0,\Omega} + \frac{1}{\mu_r} (\text{curl}_{2D}\mathbf{u}_t, \text{curl}_{2D}\mathbf{w}_t)_{0,\Omega} +$$

$$+ j \frac{\omega}{c_0} \left( \frac{\zeta_2}{\mu_r} \mathbf{u}_t, \text{grad}_{2D} w_z \right)_{0,\Omega} - j \frac{\omega}{c_0} \left( \frac{\zeta_2}{\mu_r} \text{grad}_{2D} u_z, \mathbf{w}_t \right)_{0,\Omega} +$$

$$- \frac{\omega^2}{c_0^2} (\varepsilon_r u_z, w_z)_{0,\Omega} - \frac{\omega^2}{c_0^2} \left( 1 + \varepsilon_r \mu_r - \frac{\zeta_1}{\mu_r} u_z, \mathbf{w}_t \right)_{0,\Omega} +$$

$$+ j \omega \mu_0 \left( \gamma_0 u_z, \gamma_0 w_z \right)_{0,\Gamma} + j \omega \mu_0 \left( Y(\mathbf{u}_t \cdot \mathbf{1}), \mathbf{w}_t \cdot \mathbf{1} \right)_{0,\Gamma}.$$  

(16)
for all \( u, w \in U \) and \( l \) is the following antilinear form

\[
\begin{align*}
\ell(w) = & -j\omega \mu_0 \left( J_{ez}, w_z \right)_{0, \Omega} - \left( \frac{\zeta_1}{\mu_r} \hat{z} \times \mathbf{J}_{nt}, \text{grad}_{2D} w_z \right)_{0, \Omega} - j\omega \mu_0 \left( f_{Rz}, \gamma_0 w_z \right)_{0, \Gamma} + \\
& - \left( \frac{1}{\mu_r} J_{mz}, \text{curl}_{2D} w_t \right)_{0, \Omega} - j\omega \mu_0 \left( J_{et}, w_t \right)_{0, \Omega} + j\frac{\omega}{c_0} \left( \frac{\zeta_2}{\mu_r} \hat{z} \times \mathbf{J}_{nt}, w_t \right)_{0, \Omega} + \\
& - j\omega \mu_0 \left( f_{Rt}, w_t \right)_{0, \Gamma} \end{align*}
\]

for all \( w \in U \).

In [11] we showed that under some non-restrictive hypotheses any electromagnetic problem of the class considered is well posed. In particular, for any problem of interest we can find a unique solution \((E, B, H, D)\) belonging to \( U \times (L^2(\Omega))^3 \times U \times (L^2(\Omega))^3 \) and depending continuously on \( J, J_m \in (L^2(\Omega))^3 \) and on \( f_{Rz}, f_{Rt} \in L^2(\Gamma) \). It is worth mentioning that, due to the bianisotropic behaviour of the media in motion, all unknown fields will have, in general, all three components, even if the scatterers are illuminated by simple fields.

Under some additional assumptions [11], we also found that a finite element method, based on the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines the above variational formulation and exploiting a first-order Lagrangian approximation for the axial component and a first-order edge element approximation for the transverse part, determines.

In [11] no specific considerations related to the implementation of finite element codes were provided. These considerations could be useful for our next developments and, for this reason, we present here the main points. Suppose that for any mesh adopted, fixing a given \( U_h \), we order the degrees of freedom by placing those related to \( E_{hz}, [e_z] \in \mathbb{C}^{nn} \), in the first part of the vector \([e] \in \mathbb{C}^{nn+ne} \) of the unknowns while those related to \( E_{ht}, [e_t] \in \mathbb{C}^{ne} \), are in its second and last part. In the previous formulas, \( nn \) (respectively, \( ne \)) refers to the number of nodes (respectively, edges) of the mesh considered. With this convention, by firstly considering the \( nn \) test functions like \( w_h = (0, 0, w_{hz}) \) and, then, the \( ne \) test functions like \( w_h = w_{ht} \), one can easily deduce that the general form of the final matrix equation obtained from (15), (16) and (17), with \( E, u, w \) and \( U \) replaced, respectively, by \( E_h, u_h, w_h \) and \( U_h \), is

\[
[A] [e] = [l],
\]

where the entries of \([l]\) are given by \(\ell(w_h)\), for all test functions considered. The order for its entries is given by the order considered for the test functions. \([A]\) is given by

\[
[A] = \begin{bmatrix}
[A_{zz}] & [A_{zt}] \\
[A_{tz}] & [A_{tt}]
\end{bmatrix}.
\]

In this formula, \([A_{zz}], [A_{zt}], [A_{tz}] \) and \([A_{tt}]\) are complex matrices whose entries are deduced directly from (16)

\[
[A_{zz}]_{ij} = \left( \frac{\zeta_1}{\mu_r} \text{grad}_{2D} w_{hzj}, \text{grad}_{2D} w_{hzi} \right)_{0, \Omega} + \\
\frac{\omega^2}{c_0} (\varepsilon_r w_{hzj}, w_{hzi})_{0, \Omega} + j\omega \mu_0 \left( Y (\gamma_0 w_{hzj}), \gamma_0 w_{hzi} \right)_{0, \Gamma}, \ i, j \in 1, \ldots, nn,
\]
results shown are computed by using the much slower but converging CGNE \[20\] (p. 308). The effects of the motion on their convergence behaviours. When the BiCG fails to converge the

\[ [A_{zt}]_{ij} = \frac{j \omega}{c_0} \left( \frac{\zeta_2}{\mu_r} w_{htj} \cdot \text{grad}_2D w_{hti} \right) \quad , \quad i \in 1, \ldots, mn, \quad j \in 1, \ldots, ne, \quad (21) \]

\[ [A_{tz}]_{ij} = -\frac{j \omega}{c_0} \left( \frac{\zeta_2}{\mu_r} \text{grad}_2D w_{hzt} \cdot w_{hti} \right) \quad , \quad i \in 1, \ldots, ne, \quad j \in 1, \ldots, mn, \quad (22) \]

\[ [A_{tl}]_{ij} = \left( \frac{1}{\mu_r} \text{curl}_2D u_{htj} \cdot \text{curl}_2D w_{hti} \right)_{0,\Omega} + \]

\[ -\frac{\omega^2}{c_0^2} \left( 1 + \varepsilon_r \mu_r - \xi_l \right) w_{htj} \cdot w_{hti} + j \omega \mu_0 Y \left( w_{htj} \cdot l \right) \quad , \quad i, j \in 1, \ldots, ne. \quad (23) \]

It is important to note that the second and third addends in the right-hand side of (20) are independent of \( \beta \). The same is true for the first and third addends in the right-hand side of (23). Taking account of (6) one easily deduces that, for small values of the maximum of \( |\beta| \), \( \zeta_2 \simeq \beta(\mu_r, \varepsilon_r - 1) \) and all entries of \( [A_{zt}] \) and \( [A_{tz}] \) have magnitudes which are smaller or equal to numbers proportional to the maximum of \( |\beta| \), under the same conditions. Finally, the addends of the right-hand sides of (20) and (23) which depend on \( \beta \), again for small values of the maximum of \( |\beta| \), involve \( \zeta_l \simeq 1 + (1 - \mu_r \varepsilon_r) \beta^2 \) and \( 1 + \varepsilon_r \mu_r - \xi_l \simeq \varepsilon_r \mu_r - (1 - \mu_r \varepsilon_r) \beta^2 \). Thus, the considered quantities are only slightly affected by the motion, if the maximum of \( |\beta| \) is small.

Another consequence is that, whenever \( \beta = 0 \) everywhere in \( \Omega \), the continuous and discrete problems (15) and (18) (see also (16), (17), (19), (20), (21), (22) and (23)) split into two disjoint problems, one for the axial component \( E_z \) (in the continuous case) or \( \varepsilon_{iz} \) (in the discrete case), the other for the transverse part \( E_t \) or \( \varepsilon_{it} \). The two continuous and disjoint problems are the traditional, two-dimensional variational problems formulated in terms of \( E_z \) or \( E_t \) in the presence of only motionless isotropic media. Analogously, the two discrete and disjoint problems are the same ones by using traditional, two-dimensional finite element algorithms based on first-order Lagrangian and edge elements, when only motionless isotropic media are involved.

Some kind of losses are necessary in order to deal with well posed problems \[11\]. This is a requirement also in the simple case in which all media are at rest (see, for example, \[16\]) and is not a problem in practice. For example, when scattering problems are of interest we always have some type of absorbing boundary conditions on \( \Gamma \) \[11\]. The presence of losses prevent the complex matrix \([A]\) from being Hermitian symmetric (for example, when we have losses on \( \Gamma \) we have \( \text{Re}(Y) \geq CY_{\text{in}} > 0 \) \[11\] and the diagonal terms of \([A_{zz}]\) and \([A_{tt}]\) involving \( Y \) have non-trivial imaginary parts and this is enough to deduce that \([A]\) is not Hermitian symmetric.

In order to solve the algebraic linear system (18) when \([A]\) is a non-Hermitian complex matrix, several strategies can be adopted. By confining ourselves to iterative methods, which are very often adopted in finite element codes \[17\] (pp. 382, 383, 396-405), \[18\] (p. 334), from \[19\] (p. 57) we deduce that the conjugate gradient on the normal equation (CGNE)

\[ [A]^H [A] [\varepsilon] = [A]^H [l], \quad (24) \]

can be used. Other approaches, like the biconjugate gradient method (BiCG), are very popular in finite element codes \[20\] (p. 308) and converge very quickly, at least for traditional problems, not involving moving media. It is also known that it is susceptible to breakdown \[19\] (p. 22).

In our simulations we will consider these two methods and we will investigate what are the effects of the motion on their convergence behaviours. When the BiCG fails to converge the results shown are computed by using the much slower but converging CGNE \[20\] (p. 308).
3 Definition of a test case and of the relevant errors

In order to deduce some results on the accuracy of finite element solutions in the presence of axially moving cylinders we need to consider simple problems which allow the calculation of the fields of interest with other reliable tools. For this reason we consider single canonical cylinders moving in the axial direction. Analogous studies have been performed under the same type of simplifying assumptions related to the inhomogeneity of the media involved, the particular shapes of the scatterers or the illuminating field (see, for example, [18] (p. 188) or [21]). For this reason, we consider the case of a circular cylinder hosted in vacuum and illuminated by a uniform plane wave. In particular, the cylinder axis is assumed to be the $z$ axis and the cylinder cross-section will have a radius $R \leq 0.2$ m. The medium inside the cylinder is assumed to be homogeneous and, in its rest frame, isotropic and not dispersive. It will be characterized by $\mu_r = 1$. We assume that such a medium is in uniform motion along the $z$ axis. Finally, we will consider a TM-polarized incident plane wave impinging orthogonally on the cylinder and defined by $E_{z^{inc}} = E_0 e^{j2\pi f \sqrt{\mu_0 \varepsilon_0} y}$, being the frequency always equal to 1 GHz except for one case when $f = 500$ MHz will be considered. The choice of the simple canonical problem just described was motivated not only by the possibility of finding semi-analytical solutions by using other tools, but also by the possible application of this study to the reconstruction of $\beta$ profiles, as it will be explained later on. In the following we will consider several different values for the normalized axial speed $\beta$ of the cylinder, for its relative permittivity $\varepsilon_r$ (in its rest frame) and for its radius $R$.

For problems of this class an efficient semi-analytical procedure, able to compute very good approximations of their solutions was proposed by Yeh [3] (see also Remark 5 of [13]).

The scattering problems just defined are numerically studied by using a finite element simulator based on the considerations reported in the previous section. The domain of numerical investigation we have adopted is in any case a polygon approximating a circle in the $(x, y)$ plane, whose center is at the origin and whose radius is equal to 0.4 m. Such a numerical domain is discretized by using several meshes. In particular, all these meshes are obtained by using $n$ concentric circles and, starting from the center, the innermost circle is divided into 6 segments, the next one in 12 and so on. The domain is thus divided almost uniformly into $6n^2$ triangles, with $1 + 3n + 3n^2$ nodes, $3n + 9n^2$ edges and $6n$ boundary edges. An example of one mesh of this type can be found in Figure 1 of [21] (the reader has to consider just the upper base of the three-dimensional cylinder shown in that figure). In the following $n$ will be equal to 20, 40, 80, 120, 160 or 200. Correspondingly, we will get a mesh characterized, respectively, by $h$ equal to 0.0285874, 0.143858, 0.0721629, 0.00481608, 0.00361402 or 0.00289216 m, with, respectively, 1261, 4921, 19441, 77281, 120601 nodes, 2400, 9600, 38400, 86400, 153600, 240000 elements, 3660, 14520, 57840, 129960, 360600 edges, and 120, 240, 480, 720, 960, 1200 boundary edges. All the indicated values of $n$ can be used to discretize scatterers whose radius $R$ is a multiple of 2 cm (respectively, 1 cm if we avoid using $n = 20$).

It is very important to point out that, in order to keep the analysis as simple as possible, we avoided considering meshes made up of curved triangles. This means, in particular, that, since the scatterer cross-section is not a polygon, in all our simulation we suffer from a kind of non-conformity [12] (p. 209). As a matter of fact, the scatterer has not the shape of the numerical scatterer and, moreover, the domain of numerical investigation and the numerical scatterer itself change their shapes for different values of $n$. By the same token, we adopted the semi-analytical procedure defined by Yeh [3] (the series are truncated after the first 60 terms for these calculations) to compute the piecewise constant data $f_{Rlh}$ and $f_{Rlh}$ enforcing the inhomogeneous terms in the admittance boundary conditions considered on $\Gamma$ for the discretized problem (the admittance $Y$ is set to $Y_0 = \sqrt{\mu_0 \varepsilon_0}$ is any case). In this way, we get another
violation of conformity, according to [12] (p. 183). The reader should note that these violations of conformity were not considered in [11] and formally the convergence results we deduced there could not be applied. In the following, however, according to a well established approach, we neglect this technical problem and assume that our convergence results do apply to the cases considered. The numerical results we will show provide a heuristic proof of this statement.

An evaluation of the numerical errors of the finite element solutions could now be performed. However, we introduce an additional simplification which allows us to find good estimates of the errors by exploiting finite element calculations. In particular, we will call $E_{z,h,\text{analytic}}$ and $H_{z,h,\text{analytic}}$ the first-order Lagrangian element expansions which are deduced by evaluating their degrees of freedom with the semi-analytical procedure proposed by Yeh [3]. Analogously, $E_{t,h,\text{analytic}}$ will refer to the first-order edge element expansion which is deduced by evaluating its degrees of freedom with Yeh’s procedure. In the above three calculations the series involved in Yeh’s procedure are truncated after the first 40 terms. The reader should observe that $H_{z,h,\text{analytic}}$ is not related to $E_{t,h,\text{analytic}}$ by the usual Maxwell’s curl equation (while $H_z$ is, by definition, equal to $-\frac{1}{j\omega \mu_0 \mu_r} \text{curl}_2D E_{ht}$; see equation (3.3) and the considerations below Theorem 5.3 at the end of Section 5 of [11]). As a matter of fact, it is very well known that the curl of a first-order edge element field is piecewise constant while $H_{z,h,\text{analytic}}$ is, as it has already been pointed out, a first-order Lagrangian element field. The decision to consider $H_{z,h,\text{analytic}}$ and $E_{t,h,\text{analytic}}$ not related by the usual Maxwell’s curl equation was made for the possible application of our results to inverse problem techniques aiming at the reconstruction of $\beta$ profile, as it will be further clarified later on.

The previous definitions allow us to introduce a set of (estimates of) absolute errors on $E_z$, $E_t$ and $H_z$ by using different relevant norms or seminorms. Thus we have

$$e_{z,a,l2} = \|E_{z,h,\text{analytic}} - E_{hz}\|_{0,\Omega},$$  
$$e_{z,a,h1} = \|E_{z,h,\text{analytic}} - E_{hz}\|_{1,\Omega},$$  
$$e_{z,a,semi} = \|\text{grad}_2D E_{z,h,\text{analytic}} - \text{grad}_2D E_{hz}\|_{0,\Omega},$$  
$$e_{t,a,l2} = \|E_{t,h,\text{analytic}} - E_{ht}\|_{0,\Omega},$$  
$$e_{t,a,\Gamma} = \|(E_{t,h,\text{analytic}} - E_{ht}) \cdot l\|_{0,\Gamma},$$  
$$e_{a,hz} = \|H_{z,h,\text{analytic}} - H_{hz}\|_{0,\Omega}.$$  

In the presence of a constant relative permeability $\varepsilon_{a,hz}$ can be considered as an estimate of 

$$e_{t,a,semi} = \|\text{curl}_2D E_{t,h,\text{analytic}} - \text{curl}_2D E_{ht}\|_{0,\Omega},$$  

so that we can also consider

$$e_{t,a,u2d} = \sqrt{e_{t,a,l2}^2 + e_{t,a,\Gamma}^2 + e_{t,a,semi}^2}$$  

as an estimate of the $U_{2D}$ norm error.

Relative errors are important, too, especially for the problems of interest, due to the huge variation of some of the quantities involved. For this reason, we consider the following relative errors, $e_{z,r,l2}$, $e_{z,r,h1}$, $e_{z,r,semi}$, $e_{t,r,l2}$, $e_{t,r,\Gamma}$ and $e_{r,hz}$, which are defined by dividing the corresponding absolute error by the norm of the “analytic” part involved in the definition of the absolute error itself. Once more, under the indicated condition, $e_{r,hz}$ can be considered as an estimate of $e_{t,r,semi}$. 


For the TM-polarized incident field considered, it is important to emphasize that, for a motionless cylinder, the solution presents \( E_t = 0 \). Moreover, for small values of \(|\beta|\) it is known that \( \|E_t\|_{L^2} \) is small [14]. Thus, all relative errors related to the transverse part of the electric field (that is, \( e_{t,r,\Omega}, e_{t,r,\Gamma}, e_{r,hz}, e_{t,r,semi} \)) are expected to become larger and larger as \(|\beta|\) goes to zero. At the same time, it could be important to analyze the behaviours of these errors, especially under the indicated conditions, because, on the one hand, good electromagnetic imaging techniques, able to recover the profile of the axial speed, exploits only, for the indicated incident polarization, data related to \( H_z = \frac{1}{\omega \mu_0} \text{curl}_2 \Omega E_t \) [14]. On the other hand, the finite element method we are studying can be exploited to provide approximate values of \( H_z \) at the measurement points for any trial solution for the \( \beta \) profile considered by the inverse procedure itself [14].

For this reason, in the next section, a part of our numerical analysis will be devoted to considerations related to the reliability of finite element solutions in terms of \( H_z \) and, in particular, of \( e_{a,hz} \) and \( e_{r,hz} \). This part of the analysis was, actually, the initial motivation for this study.

**Remark 1.** In many applications the reconstruction of the profiles of \( \varepsilon_r \) and \( \beta \) are of interest [7]. For the indicated polarization of the incident field, the axial component of the electric field is the most important quantity for the reconstruction of \( \varepsilon_r \) while the axial component of the magnetic field is crucial for the reconstruction of \( \beta \) [14]. In particular, under some non-restrictive hypotheses, the reconstruction of \( \varepsilon_r \) can be carried out neglecting any movement and by using data related to \( E_z \) only. The estimated \( \varepsilon_r \) is then adopted as an input data for the reconstruction of the axial speed profile. For the indicated reasons, a finite element code based on a formulation expressed in terms of \( E \) was considered. In our previous considerations we focused in particular on the generation of reliable data for the second step of the reconstruction process, devoted to the estimate of the \( \beta \) profile, simply because the reconstruction algorithms adopted for determining \( \varepsilon_r \) have been studied for decades [22] while those adopted to recover \( \beta \) are not so standard in the framework of microwave imaging techniques.

**Remark 2.** \( e_{r,hz} \) could also be referred to a proper subdomain \( \Omega_m \) of \( \Omega \). We could use \( e_{r,hz,\Omega_m} \) as an alternative symbol in this case. The subdomain can also be of zero measure (e.g., made up of curves or points) because the involved quantities (\( H_{z,h,analytic} \) and \( H_{hz} \)) are continuous in \( \Omega \). However, in this case we have to change the norm in the definition of the error (the \( L^2(\Omega) \) norm is not meaningful anymore).

The results provided in [11] can be applied to the problems here considered if some conditions involving \( \varepsilon_r \) and \( \beta \) are satisfied (see, in particular, Section 7 of [11]). In particular, in order to show some examples we can say that in the presence of a cylinder having \( \varepsilon_r = 2 \) the problem is well posed and the convergence of finite element approximations is guaranteed (neglecting the conformity violations already pointed out) whenever \(|\beta| \leq 0.264308\). For other cylinder media, for example when \( \varepsilon_r = 1.1 \) or \( \varepsilon_r = 8 \), the upper bounds for \(|\beta|\) are 0.777053 or 0.0826784, respectively. The reader should notice that the corresponding upper bounds for the axial speed values are really impressive (equal to \( \approx 232954629 \) m/s when \( \varepsilon_r = 1.1 \), \( \approx 79237545 \) m/s when \( \varepsilon_r = 2.0 \) and \( \approx 24786361 \) m/s when \( \varepsilon_r = 8.0 \)) and that they are not much smaller than the values of the speed of light in the media considered (given by \( \frac{1}{\sqrt{v}} \approx 0.953463 \), \( \frac{1}{\sqrt{2}} \approx 0.707107 \) and \( \frac{1}{\sqrt{8}} = \frac{1}{2\sqrt{2}} \approx 0.353355 \), respectively).

### 4 Numerical results

As it was already pointed out, for the defined test case we consider several values of \( \varepsilon_r, \beta, R \) and \( n \). The first results we show are all related to cases involving cylinders characterized by \( \varepsilon_r = 2 \).
and $R = 0.2$ m. In particular, in Figure 1 we report the absolute errors related to the axial component of the electric field, $\varepsilon_{z,a,h1}$, $\varepsilon_{z,a,l2}$ and $\varepsilon_{z,a,semi}$, versus $h$ for two values of $\beta$: $\beta = 0$ and $\beta = 0.25$. As the reader can easily check, it is not possible to distinguish the results of the $\beta = 0$ case from those obtained when $\beta = 0.25$. Many other simulations have been performed, for $\beta = 5 \times 10^{-m}$, $m = 1, \ldots, 15$. These additional results are not reported in Figure 1 because the plots would be the same as those already shown. In the figure two plots proportional to $h$ and $h^2$ are provided, too, in order to be able to determine the rate of convergence of the results as functions of $h$.

As it was pointed out in Section 2 the results for the $\beta = 0$ case can be obtained by using a traditional two-dimensional finite element algorithm based on first-order Lagrangian elements (dealing with isotropic media at rest). There is nothing new in the results shown for this case, as it is very well known [17]; we can simply observe that the absolute errors related to $E_z$ behaves like $h^2$ and that $\varepsilon_{z,a,semi}$ is much larger than $\varepsilon_{z,a,l2}$ so determining almost completely $\varepsilon_{z,a,h1}$.

The results corresponding to $\beta = 0.25$, a value which is close to the upper bound for $\beta$ reported in Section 3, has a significant effect on $E_z$ (these effects are not shown in a figure for space reasons; we simply observe that, for example, $Re(E_{z,z}) = -0.0544$ V/m at the origin in the $\beta = 0$ case while $Re(E_{z,z}) = -0.132$ V/m at the same point when $\beta = 0.25$. Both values are obtained by using a mesh with $n = 200$) and corresponds to a huge axial speed of $(74948114.5$ m/s), show that the finite element capability of approximating the axial component of the true solution is the same as in the case all media are at rest. In particular, the convergence rate remains quadratic and, taking account of all simulations performed, we can say that this property is completely independent of $\beta$. For the same reason, no figure related to the relative errors on $E_z$ is provided.

In Figure 2 some results related to $E_t$ are shown. The case $\beta = 0$ is not meaningful for the present analysis because, for the test case considered, the solution has $E_t = 0$ in $\Omega$ and the finite
element method is able to compute $E_{ht} = 0$ in $\Omega$ for any mesh considered, as explained in Section 2. For this reason the results refer to cases in which $\beta \neq 0$. A complete set of results is shown in the indicated figure for $\beta = 0.25$. One can observe that $e_{t,a,semi}$ is much larger than $e_{t,a,l2}$ and $e_{t,a,\Gamma}$ so that $e_{t,a,u2d} \approx e_{t,a,semi}$. Moreover, one can observe that the convergence rate is $O(h)$ for all errors considered.

Taking account of these results and of those related to $E_z$ we conclude, as expected [11], that the finite element approximation is converging to the true solution (notwithstanding the two violations of conformity pointed out in Section 3).

A few other results are shown in the same figure. They refer to $10^s C e_{t,a,semi}$ for $\beta = 5 \times 10^{-m}$, where $C = 5 \times 10^7$, $m = 9, \ldots, 12$, $s = m - 9$. The factor $10^s C$ have been introduced to avoid plotting a figure with a huge range along the vertical axis, which determines a potential loss of information. This is because, for small values of $|\beta|$, all parts of the $U_2D$ norm of $E_t$ are proportional to $|\beta|$ (see [14]). With the indicated coefficient, on the contrary, we get almost the same values for the plotted quantities since $10^s C$ is able to balance the drop of $|\beta|$ as $m$ get larger and larger. In particular, $10^s C 5 \times 10^{-m} = 0.25$ for all $m$ of interest, which is the largest value for $\beta$ considered in the figure.

One can observe that for $m = 9$ we obtain exactly the same values we get for $\beta = 0.25$. We avoided reporting other intermediate results ($m = 1, \ldots, 8$, which are all available, as it has already pointed out) for the same reason.

For $m = 10$ we get the same results for $h$ equal to 0.0285874, 0.0143858, 0.00721629, and 0.00481608 m, corresponding to $n = 20, 40, 80$ and 120, respectively, but a slight deviation occurs for $h = 0.00361402$ and 0.00289216 m (corresponding to $n = 160$ and $n = 200$, respectively). As $m$ is increased the deviation is larger and begins for larger values of $h$ (i. e., for smaller values of $n$). For example, for $m = 11$ a deviation can be observed for $h = 0.00481608$ m (i. e., $n = 120$)
and for $m = 12$ even for $h = 0.00721629$ m (i.e., $n = 80$).

These results should not be interpreted as a failure of convergence, however. Since in Figure 2 only absolute errors are considered, this behaviour is not related to the small norm of $E_t$, either. It is likely, on the contrary, that it is a consequence of the round-off errors introduced by our finite element code (based on double-precision arithmetic) and, possibly, of the progressive reduction of the condition number of the final finite element matrix as we consider values of $h$ which are smaller and smaller. In particular, $h$ values much smaller than the smallest values usually exploited in applications for the considered problems (remember that the frequency value is 1 GHz, giving $\lambda_0 \simeq 0.3$ m, that the materials involved are not very dense and, finally, that for the more critical $\beta$ values the wavelength is only slightly changed).

In order to have an idea of the importance of these numerical errors it is necessary to refer to relative errors. This is important also for the application of the finite element code here studied to inverse scattering procedures aiming at the reconstruction of the profiles of $\beta$.

In Figure 3 the behaviour of $e_{r,hz}$ versus $h$ is reported, for $\beta = 0.25$ and $\beta = 5 \times 10^{-m}$, $m = 9, \ldots, 15$. Since we plot only $e_{r,hz}$ we extend the range of $m$ in order to complete the analysis, without any risk to reduce the readability of the figure. The reader can easily find the deviations from the ideal behaviour. However, it is believed that the most significant comment is related to the capability of the finite element code to get good approximations in terms of relative errors on $H_z$. As a matter of fact, one can observe that it is possible to get $e_{r,hz} \leq 0.03$ (respectively, 0.02) for $h \leq 0.00721629$ m (respectively, $h \leq 0.00481608$ m), corresponding to $n = 80$ (respectively, $n = 120$) for all $m \leq 11$. For $m = 12$ we can get results just above 3% for the same range of $h$ and it is not possible anymore to get $e_{r,hz}$ below 2%. Finally, for $m \geq 13$ the reliability in terms of the considered relative error is completely lost for small values of $h$ ($m = 13$) or independently of $h$ ($m = 14$ or $m = 15$).

The reader should observe, however, that the reliability we get independently of $h$ and $\beta$ at
The cylinder is assumed to be made up of a material having $\varepsilon_r = 2$ at rest. The numerical solutions are computed by using a mesh with $n = 120$.

Figure 4: Behaviour of $e_{z,r,h1}$ versus $R$ for different values of $\beta$ ($\beta = 0, 5 \times 10^{-8}, 5 \times 10^{-5}, 0.05, 0.25$). The cylinder is assumed to be made up of a material having $\varepsilon_r = 2$ at rest. The numerical solutions are computed by using a mesh with $n = 120$.

least for $m \leq 10$ means that, for the case analyzed which involves a cylinder having $\varepsilon_r = 2$ in its rest frame, the relative error on $H_z$ is under control when the magnitude of the axial speed is larger than or equal to $5 \times 10^{-10}$, which is approximately $15 \times 10^{-2}$ m/s, or $0.15$ m/s.

Before considering variations of $\varepsilon_r$ we have to take account of possible variations of $R$. As for the errors related to $E_z$, no difference with respect to the previous case are worth mentioning. In particular, in Figure 4 we report $e_{z,r,h1}$ versus $R$ for different values of $\beta$ when $\varepsilon_r = 2$ and $n = 120$. Once more the performances related to the axial component of the electric field are not significantly affected by $\beta$ (the differences in this case can be observed simply because in Figure 4 we focus on a very small range of values along the vertical axis) and we have verified that all comments reported for the $R = 0.2$ m case and concerning $E_z$ could be duplicated here.

In Figure 4 we do not consider $R$ smaller than 5 cm because a further reduction of $R$ could increase the effects of the first non-conformity pointed out in Section 3. As a matter of fact, the difference between the polygonal scatterers considered in finite element simulations and the circular ones adopted for the computations of the reference solutions and of the boundary conditions becomes larger and larger as $R$ gets smaller and smaller values.

An equivalent effect can be obtained, however, by reducing $f$. This change is not considered in Figure 4 since the effects of the movement on all errors related to $E_z$ are negligible. On the contrary, it is used in the analysis of the relative error $\varepsilon_r$. Since a reduction of $R$ could determine a reduction of the bianisotropic effect of the scatterer, at least for small values of $R$, such a reduction could restrict the width of the range of $\beta$ of numerical reliability, in terms of the indicated error, of the finite element code. This effect is in some way confirmed by the results shown in Figure 5. As a matter of fact, the smaller the radius the sooner we get a deviation from the ideal (“flat”) behaviour when we start reducing $\beta$, independently of the frequency value. Such an effect is not so important, however. If it was, a reduction of the frequency and of the radius, determining a significant effective reduction of the extension of the scatterer and, then,
Figure 5: Behaviour of $e_{r,hz}$ versus $\beta$ for different values of the scatterer radius $R$ and of the frequency $f$. The values of $\beta$ which are particularly critical for $e_{r,hz}$ are considered. The cylinder is assumed to be made up of a material having $\varepsilon_r = 2$ at rest. The numerical solutions are computed by using a mesh with $n = 120$.

of its depolarizing effect, should worsen the situation but this is not the case. On the contrary, the well known [17], [18] (p. 344) and contrasting effect related to the error reduction due to the reduction of the frequency, when the same mesh is adopted, is sufficient to widen the range of stability of the results: at 500 MHz it seems that, for the indicated conditions, we can guarantee reliable results provided that $\beta \geq 5 \times 10^{-12}$, while the lower bound for $\beta$ seems to be $5 \times 10^{-11}$ for $f = 1$ GHz.

We have to verify what happen when the cylinder is made up of other materials. In the following, in order to complete our analysis, we will consider $\varepsilon_r = 1.1$ and $\varepsilon_r = 8$. These values were chosen taking account, on the one hand, of the meshes we have considered and, on the other hand, of equation (6) which, ultimately, determines the bianisotropy of the homogeneous scatterer. For $\varepsilon_r = 8$ we have $\zeta_2 \simeq \beta (\mu_r \varepsilon_r - 1) = 7\beta$, while for $\varepsilon_r = 1.1$ we get $\zeta_2 \simeq 0.1\beta$. These values of $\zeta_2$ should be compared with $\zeta_2 \simeq \beta$, which is the result we get for the $\varepsilon_r = 2$ case studied before.

The results related to $E_z$ are not affected at all by these considerations related to the bianisotropy, due to the independence of $\beta$ of the approximations already pointed out and retained in all new cases considered (see Figure 6). The convergence rate remains $O(h^2)$, as well. However, for any given mesh, the relative errors on $E_z$ are larger for the $\varepsilon_r = 8$ case, due to the reduction of the wavelength in the scatterer by a factor approximately equal to 2 (if the comparison is done with respect to the case in which $\varepsilon_r = 2$). This effect is very well known and is related to the one considered above and concerning a frequency variation [17], [18] (p. 344).

As for the errors related to $E_t$, the results shown in Figure 7 for $\varepsilon_r = 1.1$ and $\varepsilon_r = 8$ should be compared with the ones reported in Figure 3.

The comparison clearly indicates that the best results are those achieved when $\varepsilon_r = 2$. In comparing the plots related to $\varepsilon_r = 1.1$ and $\varepsilon_r = 8$ with the ones related to $\varepsilon_r = 2$ two contrasting
Figure 6: Behaviour of $e_{z,r,h1}$ versus $h$ for different values of $\beta$ and $\varepsilon_r$. In particular, for any value of $\varepsilon_r$ one of the largest possible normalized axial speed values is considered in addition, for both cases, to $\beta = 0$. The scatterer radius is $R = 0.2$ m.

effects have to be considered, as it was the case for the results of Figure 5. The case with $\varepsilon_r = 1.1$ (respectively, $\varepsilon_r = 8$) presents the smallest (respectively, the largest) discretization errors but also, for a given value of $\beta$, the smallest (respectively, the largest) norm of the field having orthogonal polarization with respect to the one of the incident field. As it was already pointed out, the latter effect could increase (respectively, reduce) the importance of round-off errors in any evaluation of relative errors related to $E_t$ or $H_z$.

The results reported in Figures 7 and 3 show that the second effect is dominant with respect to the first one when $\varepsilon_r = 1.1$. It is difficult to understand what happen from a quantitative point of view. Qualitatively, one can observe that the magnitude of the orthogonal polarization could be ten times smaller than in the case $\varepsilon_r = 2$ (because, as it has already been pointed out, $\zeta_2$ is ten times smaller) while the wavelength change in the scatterer is limited to a $\approx 26\%$ ($\approx \frac{\sqrt{2} - 1}{\sqrt{8} - 1}$). The opposite happens when $\varepsilon_r = 8$ and this could be a consequence of the fact that the scatterer wavelength changes by a $50\%$ ($\approx \frac{1}{\sqrt{8} - 1}$) while the magnitude of the orthogonal polarization could be only seven ($< 10$) times larger than in the case $\varepsilon_r = 2$.

Since we have done no attempt to find the best possible case we can assume that it is possible to find a value of $\varepsilon_r$ of the scatterer giving even better results than for the case $\varepsilon_r = 2$. This is not really interesting, however, since we are much more interested in the worst case. In this sense, we could say that from the previous considerations we can expect that the situation can get worse by considering $\varepsilon_r > 8$ or $1 < \varepsilon_r < 1.1$. It should also be worse when $\varepsilon_r < 1$ and, at the same time, $\varepsilon_r \approx 1$. Sooner or later, however, we should find a threshold value $T < 1$ such that for $0 < \varepsilon_r < T$ we should get improvements because the two effects considered are no more in contrast (a reduction of $\varepsilon_r < 1$ determines an increase of the wavelength in the scatterer and an increase of the depolarization effect).
Figure 7: Behaviour of $\varepsilon_{r,hz}$ versus $h$ for different values of $\beta$ and $\varepsilon_r$. In particular, when $\varepsilon_r = 1.1$ we consider $\beta = 0.75$ and $\beta = 5 \times 10^{-m}$, $m = 9, \ldots, 12$, while for $\varepsilon_r = 8.0$ we consider $\beta = 0.08$ and $\beta = 5 \times 10^{-m}$, again with $m = 9, \ldots, 12$. Two plots proportional to $h$ and $h^2$ are provided, too.

In addition to the comments related to the comparative analysis, we can say that the asymptotic behaviour of $\varepsilon_{r,hz}$ remains $O(h)$ in both cases, when $\beta$ is large enough. Moreover, the reliability of our finite element code in terms of all errors considered is guaranteed, for the cases considered in Figures 6 and 7, for $m \geq 10$ provided that we avoid using $n = 200$ or for $m \geq 9$ without any restriction in terms of $n$.

As a final comment, it is interesting to observe that, according to our results, we could consider the finite element method here studied as a reliable numerical tool in the presence of axially moving media for a huge range of $\beta$ values, even when the most critical errors are evaluated (as is the case of the errors related to $H_z$, in the presence of a TM-polarized incident field and of numerical methods considering the electric field as their primary unknown field), provided that the materials in motion are not characterized by $\mu_r \varepsilon_r$ too close to 1 in their rest frames. Such a huge range of $\beta$ values could cover applications of interest in astrophysics (astrophysical jets can have $|\beta| \geq 10^{-1}$ [23]), plasma physics (where $|\beta|$ could be as large as $10^{-3}$ [2]), engineering (bulk solids or water in pneumatic pipelines can move at $|\beta| \approx 10^{-7}$ [7]) or medicine (blood speed can have $|\beta| \approx 10^{-9}$ on average, in large blood vessels [24]). As a matter of fact, when the motion takes place with small axial velocities the materials involved have usually $\mu_r \varepsilon_r$ very different from 1. On the contrary, for example in plasma physics, we can have $\mu_r \varepsilon_r$ closer to 1 than before but much larger $|\beta|$ values are usually considered.

It could happen, however, that one be interested in applications involving materials with $\mu_r \varepsilon_r$ so close to 1 and moving with a so small axial velocity that the effects of motion are very hardly perceived. In these cases any double-precision finite element code could be not enough precise to evaluate the very small effects of the motion. Quadruple-precision or octuple-precision finite element codes could be considered in these cases because in infinite precision arithmetic convergence of the finite element approximation is in any case guaranteed [11].
5 Solution of the algebraic linear system

As it was already pointed out, iterative methods are very often exploited to compute the solution of the algebraic linear systems determined by electromagnetic finite element codes [17] (pp. 382, 383, 396-405). In particular, the BiCG [17] (pp. 396-405) is known to be particularly efficient for time-harmonic electromagnetic problems [20] (p. 308) involving only traditional media.

In Section 2 we analyzed the structure of the matrix defined by finite element codes for two-dimensional time-harmonic electromagnetic boundary value problems in the presence of axially moving cylinders. The effects of the motion on the magnitudes of the entries of the different submatrices were discussed, too. In summary, we can say that the main diagonal square submatrices are slightly affected by β, unless relativistic velocities are involved. On the contrary, the two off-diagonal matrices are much more sensitive to the magnitude of β. It is very difficult, however, to understand how this change of the matrix structure with β could affect the performances of iterative solvers like BiCG. This difficulty is worsened by the well known fact that the magnitude of the entries of finite element matrices depends on h in a different way according to the presence of spatial derivatives [25] (p. 141) and such a presence is different in the four submatrices $[A_{zz}]$, $[A_{zt}]$, $[A_{zt}]$ and $[A_{zt}]$. Thus, in order to provide indications on the performances of the iterative solvers of interest, specific investigations are usually required [26]. In the following we limit ourselves to give some data related to the performances of the BiCG and CGNE, as it was pointed out in Section 2.

In Section 4 we observed that we are interested in having a good control of the errors of the numerical solutions. For this reason we have to control the errors that we admit when we stop the iterative solver. The stopping criterion we have adopted is “criterion 2” of [19] (p. 60). For the reader convenience we recall here the main points. If the algebraic linear system to be solved is $[A][e] = [l]$, as reported in Section 2, we initially calculate the euclidean norm $||[l]||$ of $[l]$ and do not stop the iterative solver until the approximate solution $[e]_i$ at iteration $i$ satisfies $||[A][e]_i - [l]|| < \delta ||[l]||$, with $\delta$ always equal to $10^{-5}$, $p \in \{10, \ldots, 16\}$. One should avoid using larger values of $p$ when double-precision arithmetic is used [19] (p. 58). Another key factor allowing very often to improve the control of the numerical errors is preconditioning [20] (p. 313), [19] (p. 39). A large number of preconditioning techniques is available, however [19] (pp. 39-55). For this reason, in order to avoid a too detailed discussion on this point, in all our simulations we have considered iterative solvers without any preconditioner or with the so-called point Jacobi preconditioner [19] (p. 41), also known as diagonal preconditioner [20] (p. 313). In the following we will refer to the point Jacobi preconditioned BiCG (respectively, CGNE) by using BiCGJP (CGNEJP; preconditioning is meant to be applied before forming the normal equation, so deducing, instead of equation (24), $[A]^H([M]^{-1})^H[M]^{-1}[A][e] = [A]^H([M]^{-1})^H[M]^{-1}[l]$, where $[M]$ is a diagonal matrix whose entry $[M]_{ii}$ is the same as $[A]_{ii}$).

All the results shown in Section 4 were calculated by using BiCG, BiCGJP, CGNE or CGNEJP. In order to avoid presenting unreliable results we computed them at least twice, for two consecutive values of $p$. When the outcomes were different we considered the next larger value of $p$ and did not stop this process until, for two consecutive values of $p$, we got the same result (with a tolerance equal to 0.1 %).

Independently of the reliability of the outcome, we find that the BiCGJP is able to compute the solutions for all values of β and $p$ considered and all $n \leq 80$. For finer meshes the convergence of biconjugate iterative solvers is not guaranteed anymore. In particular, when $n = 120$, BiCGJP does not converge, for example, when $\varepsilon = 2$, $R = 0.2$ m and $f = 1$ GHz, for $\beta \geq 5 \times 10^{-5}$ (for this specific example the smallest residual value seems to be $\approx 0.26$ which is by far too large to stop the iterations). BiCG fails to converge, under the same conditions, for $\beta \geq 5 \times 10^{-3}$ (the smallest residual value is $\approx 0.026$). For a mesh obtained by using $n = 160$ BiCG (respectively, BiCGJP)
does not converge, again under the conditions indicated above, for $\beta \geq 5 \times 10^{-7}$ (respectively, $\beta \geq 5 \times 10^{-5}$). When $\beta = 5 \times 10^{-7}$ the smallest residual value is $\simeq 3.2 \times 10^{-7}$ (respectively, $\simeq 0.28$ for $\beta = 5 \times 10^{-5}$). In particular, by using $p = 13$, BiCGJP converges in 20214 steps (with the indicated mesh we have 308161 unknowns) for $\beta = 5 \times 10^{-7}$ and in 242643 steps for $\beta = 5 \times 10^{-6}$. Finally, when $n = 200$ we have 481201 unknowns and BiCGJP and BiCG fail to converge for $\beta \geq 5 \times 10^{-7}$ in the presence of the same scatterer and illuminating field, reaching a minimum residual value, for $\beta = 5 \times 10^{-7}$, of $0.45 \times 10^{-8}$ and, respectively, 0.5. In particular, BiCGJP converges in 22956, 20001, 17834 and 13415 steps for $\beta = 5 \times 10^{-6}$, with $m$ respectively equal to 8, 9, 10, 11 for $p = 15$. It is also interesting to point out that it converges in 3556 steps for $\beta = 0$ (and $p = 15$ as before).

By analyzing the data reported above one can understand that it is not easy to deduce a general rule allowing to understand the behaviour of BiCG, with or without the point Jacobi preconditioner, in the presence of moving media. When the mesh is not very fine the solver seems to be able to converge independently of the selected values for all parameters here considered. For finer meshes, however, it fails to converge even for relatively small axial speed values.

From the previous considerations one can also understand that most of our results were calculated by using CGNE or CGNEJP. The convergence of this type of algebraic solver, however, is by far too slow, independently of the use of the preconditioner. Even though this is not a surprise [20] (p. 308), [19] (p. 18), it is instructive to report some data. For example, CGNE, for $p = 15$ and $n = 200$, requires 536957 steps to converge for $\beta = 5 \times 10^{-6}$ and 710410 steps for $\beta = 0.25$ ($\epsilon_r = 2$, $R = 0.2$ m, $f = 1$ GHz). The results are not much better when the point Jacobi preconditioner is used. For example, with the usual values of $\epsilon_r$, $R$ and $f$, for $n = 120$ and $\beta = 0.25$ CGNEJP converges in 220942 steps (173521 unknowns) while CGNE does the same requiring 233866 steps ($p = 13$ in both cases). In terms of CPU time the difference is even lower (equal, more or less, to 4% on the same computer).

In summary, we can say that in the presence of moving media a reliable and fast iterative solver is still to be found and some additional works could be required.

Remark 3. The slow convergence of CGNE and CGNEJP is known to be related to the fact that the condition number of $[A]^H[A]$ can be much larger than that of $[A]$ [19] (p. 18). It is also known that the same property could introduce larger numerical errors. However, we use CGNE or CGNEJP when $\beta$ assume values which are not so small (and prevent, in particular, the convergence of BiCG or BiCGJP). These cases are not the most critical ones in terms of relative errors, especially on $E_t$ or $H_z$.

6 Conclusions

The accuracy of finite element results in the presence of axially moving cylinders is analyzed for the first time, to the best of authors’ knowledge.

The study refers to relative and absolute errors related to two components of the electromagnetic field. The part of the results presented concerning one of the two components is new but the outcome is analogous to the one which is obtained when all media involved are motionless. The second part is related to the field component which is specifically excited by the presence of moving objects. This field component is the most difficult to be approximated. For its information content related to the motion of the objects, it could also be the most important component to be evaluated, at least for some applications.

This study has shown that finite element simulators based on double precision arithmetic could guarantee and extraordinary reliability of all their outcomes. These performances suggest that the indicated simulators can be exploited and could become the reference method for astrophysics, engineering and medical applications involving media in motion.
References


