A three-dimensional self-adaptive \(hp\) finite element method for the characterization of waveguide discontinuities

Ignacio Gomez-Revuelto, Luis E. Garcia-Castillo, Sergio Llorente-Romano, David Pardo

**Abstract**

We propose the use of a highly-accurate three-dimensional (3D) fully automatic \(hp\)-adaptive finite element method (FEM) for the characterization of rectangular waveguide discontinuities. These discontinuities are either the unavoidable result of mechanical/electrical transitions or deliberately introduced in order to perform certain electrical functions in modern communication systems. The proposed numerical method combines the geometrical flexibility of finite elements with an accuracy that is often superior to that provided by semi-analytical methods. It supports anisotropic refinements on irregular meshes with hanging nodes, and isoparametric elements. It makes use of hexahedral elements compatible with high-order \(H(\text{curl})\) discretizations. The 3D \(hp\)-adaptive FEM is applied for the first time to solve a wide range of 3D waveguide discontinuity problems of microwave communication systems in which exponential convergence of the error is observed.

1. Introduction

Waveguides are a fundamental part of any modern communication system [1]. When the translational symmetry of the waveguide is interrupted, we refer to this situation as a waveguide discontinuity. In some cases, discontinuities are due to mechanical defects, transition between waveguides, the need of bends, junctions, etc. In these cases, the main goal is to minimize the power reflection produced by the discontinuity, thereby, maximizing the amount of transferred power. In other cases, the translational symmetry of the waveguide is deliberately interrupted by introducing, for example, posts, irises, and stubs. In these cases, the goal is to produce a significant amount of reflection waves such that a certain electrical function is performed, for example, impedance matching and frequency filtering networks [2,3]. A large number of microwave passive networks are simply “waveguide discontinuities” that are typically characterized by the power relations between their waveguide ports. This paper focuses on rectangular waveguide discontinuities, i.e., discontinuities in which the ports are rectangular waveguides. The rectangular waveguide technology is used in many modern communication systems (specifically, satellite and other systems on the upper microwave and millimeter wave frequency bands) due to the simplicity of its geometry, low losses, and the capacity to handle high powers.

Several computational techniques can be used for the analysis of rectangular waveguide discontinuities. A first classification can be made between analytic (or semi-analytic) techniques and pure discretization methods, such as FEM [4,5], and finite differences (FD) [6]. Techniques in the first group are computationally very efficient, and they can be employed to design several devices with given specifications using optimization procedures [7,8].

However, the main advantage of pure discretization methods approaches is their flexibility to model arbitrary discontinuities. Moreover, even for devices designed with semi-analytic methods, they need to be finally validated with a pure numerical technique that can analyze the structure as a whole, taking into account the effect of round corners due to finite radii of milling tools, the influence/sensibility to position of tuning elements (posts, screws, etc.), losses, and so on. In this context, a FEM has demonstrated to be a powerful and versatile tool. However, high accuracy is not always easy to achieve for a conventional FEM when dealing with field singularities, high-contrast material data, etc. Thus, semi-analytic
results (when available) are typically considered as a reference for the engineering analysis.

In [13], we described a two-dimensional (2D) fully automatic hp-adaptive finite element method (FEM) for the analysis of rectangular waveguide discontinuities with the features mentioned above. The analysis was restricted to those structures that are invariant in one given direction, that is, the so-called H-plane and E-plane discontinuities. In this paper, we extend our previous 2D work [13] to the 3D case for the analysis of general rectangular waveguide discontinuities as those shown in Fig. 1. In addition to the characterization of 3D waveguides that cannot be analyzed with the 2D code, we also consider some H-plane and E-plane discontinuities for verification purposes.

This 3D work enables the analysis of more general structures than those considered in our previous 2D work [13]. Indeed, the numerical method used in 3D computations is also different than that employed in 2D, both in terms of mathematical formulation and computer implementation. The number of lines of the 3D code has considerably grown (more than double) with respect to 2D code described in [13]. While mathematical and computational complexity of a general 3D FE code is at least one order of magnitude higher than in 2D, in the case of hp adaptive computations, this complexity increase from 2D to 3D is even higher. Although conceptually the 3D adaptive strategy is, up to some point, an extension of the one in 2D, the algorithm and its implementation presents many differences, mainly due to three reasons: (a) in the hp-adaptive procedure there exist a higher number of “choices” available to decide the next hp-mesh, (b) ensuring continuity when anisotropic hp-refinements are supported becomes a highly technical and difficult issue to implement, and (c) the computational order of complexity for polynomials of order p involved in the computations is significantly higher in 3D than in 2D. Thus, a straightforward extension of the 2D hp adaptive algorithms and routines to the 3D case leads to a failure in the analysis of practical problems. All the factors mentioned above make the development of a 3D hp FE and automatic adaptivity a challenging task in comparison with its 2D counterpart.

The 3D hp-adaptivity presented here has been the result of a research and development process over the last 20 years. After the development of several preliminary implementations of 3D hp-codes, a version for electromagnetics was presented in [14]. It is worth noting that the electromagnetic modeling has specific issues in contrast to other physics, [15]. The 3D self-adaptive hp-strategy was first developed for elliptic problems in [16]. The resulting code was very complicated but it did provide a proof of concept for different aspects, including parallelization [17]. Its version for electromagnetics was mainly applied to scattering problems in open domains [18–20].

This hp-adaptive strategy led to a second-version of the 3D hp-algorithm and a re-implementation from scratch in [21] (see also [22] for the elliptic case), where in addition to removing a number of bugs and implementation problems, it was for the first time applied, as a proof of concept, to one simple waveguide problem. All these contributions are summarized in [23].

Despite these advances described in [23], a number of technical and implementation problems were still unresolved, including a so-called “deadlock” problem. This prevented the solution of more complex rectangular waveguide problems. In here, we further developed the 3D hp-adaptive strategy by removing a number of faults and implementation problems, and we extensively applied it to solve general rectangular waveguide applications. In this work, we describe specific formulations corresponding to this class of problems, and apply for the first time hp-finite elements and adaptivity to solve a wide range of 3D waveguide discontinuity problems of microwave communication systems in which exponential convergence of the error is observed.

The 3D self-adaptive hp-strategy supports anisotropic refinements on irregular meshes with hanging nodes, and isoparametric elements. It makes use of hexahedral elements (which are particularly well-suited for the modeling of structures in rectangular waveguide technology) and it is compatible with H(curl) higher order discretizations.

The organization of the paper is as follows. The discontinuity problem in rectangular waveguide technology is introduced in Section 2, together with the mathematical and variational formulation of the problem. Procedures to compute the scattering parameters of the structure are also described in that section. The 3D hp finite element discretization and automatic adaptivity strategy are described in Section 3.1 and 3.2, respectively. Numerical results are shown in Section 4. Finally, some conclusions are given in Section 5.

2. Analysis of rectangular waveguide discontinuities

Engineering characterization of waveguide discontinuities in terms of the so called scattering parameters is introduced next.
The mathematical formulation of the problem is presented in Section 2.2. The details can be found in [13,24], or in any of the numerous books in the subject (e.g., [25–29]).

2.1. Characterization by $S$-Parameters

Rectangular waveguides, and more generally any homogeneous waveguide, support TE ("transverse electric") and TM ("transverse magnetic") modes with respect to the translational symmetry axis (z in Fig. 2). Those modes experiment a cut-off effect, i.e., propagation occurs only above a certain non-zero frequency. Typically, the modes are ordered according to their cutoff frequency $f_c$ in ascending order. The first propagating mode is referred to as the fundamental mode. When $a > b$, this first fundamental mode of a rectangular waveguide is called TE$_{10}$ mode.

The field of the TE$_{10}$ mode points towards the $\eta$-direction (i.e., parallel to the narrow edge of the waveguide) and it is constant in that direction, and of the form one "half-sine"/"half-cosine" along the $\xi$-direction (wide edge).

In general, the discontinuity region excites all modes in the waveguides of the structure going outwards the region in order for the boundary conditions at the discontinuity to be satisfied. Thus, part of the energy is reflected back into the excited port and the rest of the energy is transmitted to the other ports. In microwave engineering the waveguide operates typically in the monomode region (the excitation frequency is within the cut-off frequencies of the fundamental and second modes). Thus, if the ports are placed at a certain distance from the discontinuity or further away, then only the TE$_{10}$ mode is present with a non-negligible amplitude at the ports. The discontinuity can be characterized by the ratio of power carried by the reflected TE$_{10}$ wave (for the excited port) or by the transmitted TE$_{10}$ wave (for the remaining ports) to the power carried by the incident TE$_{10}$ wave (at the incident port). These ratios are directly related to the so-called $S$-parameters, which are organized in matrix form (scattering matrix $S$).

In electrical engineering terminology, the inward wave is referred to as the incident wave. Typically, during simulations only one port is excited at a time. Thus, the outward wave is referred to as the reflected wave for the excited port and as the transmitted wave for the remaining ports. The complete characterization of the discontinuity requires several simulations exciting each time a different port of the structure. By exciting the $i$th port of a $N$-port linear microwave network and implementing an absorbing boundary conditions in all other ports (e.g., Cauchy boundary conditions as in (4)), we obtain $S_{ij}$, $j = 1 \ldots N$ (the $i$th column of $S$). Note that

$S_{ij}$ are reflection coefficients and $S_{ji}$, $j \neq i$, are transmission coefficients. Specifically, $|S_{ii}|^2$ represents the ratio between the mean powers of the reflected and incident waves. The argument of $S_{ii}$ represents the phase shift between the reflected and incident waves. Analogously, $|S_{ji}|^2$, $j \neq i$ relates the power relation between the transmitted wave at port $j$ and the incident wave at port $i$; the argument of $S_{ji}$ gives the phase shift between transmitted and incident waves. In general, the complete characterization of a $N$-port microwave network requires the field solution of the structure with $N$ different excitations (exciting a different port each time), i.e., the same matrix system of equations with $N$ different right hand sides. Obviously, the knowledge of some properties of the structure under analysis (reciprocal or not, lossless or lossy, the presence of some symmetries, etc.) allows to infer some of the $S$ parameters in terms of others, so the $S$ matrix may be obtained with a lower number of experiments.

Thus, the electrical behavior of a complex microwave system may be analyzed by a cascade connection of blocks, each block being characterized by its $S$ matrix.

The $S$ matrix is symmetric if and only if the circuit is reciprocal.\footnote{In the simplest sense, a reciprocal system is such that the response of the system to a source remains constant when source and measure are interchanged. It can be proved that any system in which the media are linear with permittivity and permeability characterized by symmetric tensors is reciprocal.} This is the case of the structures analyzed in this paper. Non-reciprocal behavior would require a non-symmetric tensor for the electric permittivity and/or magnetic permeability. For a lossless structure the $S$ matrix is unitary (see also [28]).

2.2. Mathematical formulation

By combining the time-harmonic source free form of Maxwell's curl equations, it is obtained

$$
V \times \frac{1}{\varepsilon_0} V \times H - k_0^2 \mu_0 H = 0
$$

(1)

where $H$ stands for the magnetic field intensity vector, $\varepsilon_0 = \varepsilon/\varepsilon_0$, $\mu_0 = \mu/\mu_0$ being $\varepsilon$, $\mu$ the electric permittivity and magnetic permeability, respectively, of the medium, and $k_0 = \omega/\sqrt{\varepsilon_0 \mu_0}$. The subindex "0" is used to particularize to vacuum medium. Notice that it is possible to further re-scale the above equation (c.f., [30,31]), although in our applications a further re-scaling has no additional advantages.

The imposed boundary conditions are:

$$
\mathbf{n} \times H = 0, \quad \text{at perfect magnetic conductors (}\Gamma_D) \tag{2}
$$

$$
\mathbf{n} \times \frac{1}{\varepsilon_0} V \times H = 0, \quad \text{at perfect electric conductors (}\Gamma_N) \tag{3}
$$

$$
\nabla \times \mathbf{n} \times H + j \omega \beta_{10} \mathbf{n} \times \mathbf{n} \times H = \mathbf{0}^n, \quad \text{at the port boundaries (}\Gamma_p) \tag{4}
$$

with $\mathbf{0}^n$ defined as

$$
\mathbf{0}^n = 2jk_0^2 \beta_{10} \mathbf{n} \times \mathbf{n} \times H^n
$$

(5)

and $H^n$ being the incident magnetic field at the port corresponding to the TE$_{10}$ mode. $\mathbf{n}$ stands for the outward unit vector normal to $\Gamma$. Notice that perfect magnetic conductors are not present physically in the structure under analysis but they may be used in certain cases as symmetry boundary conditions for the FEM analysis, in order to reduce the number of unknowns.

The "absorbing" boundary conditions at the ports (4) are exact provided that the only mode present at the ports is the TE$_{10}$ mode. In practice, and assuming the monomode propagation (typical
mode of operation), boundary condition (4) can be made arbitrarily accurate by simply truncating the FEM domain at a certain distance from the discontinuity. This is due to the exponential decrease of the amplitude of the high-order modes as we travel away from the discontinuity. For instance, the lengths of the waveguide sections that connect the discontinuity to the ports of the structure have been typically chosen around 0.85 wavelengths. This assures an error level below approximately −65 dB. Lower error levels can be obtained by placing the ports further away. Another approach is the use of boundary conditions that take into account the presence of several modes at the ports, i.e., absorbing multimode boundary condition. The higher the number of modes considered the closer the port can be placed to the discontinuity. However, the discretization of the boundary condition yields then non-sparse matrices. The interested reader is referred to [32, Section 11.1]. Although other type of truncation methods such as PML are also possible, it is customary to use absorbing boundary conditions in this context.

The variational formulation is obtained by multiplying (1) with a test function \( \mathbf{F} \). Integration by parts (first vector Green's theorem) is applied to the double curl term and a boundary term appears in the variational formulation:

\[ c(\mathbf{F}, \mathbf{H}) = l(\mathbf{F}), \quad \forall \mathbf{F} \in \hat{\mathcal{W}} \]  \hspace{1cm} (6)

where bilinear form \( c(\mathbf{F}, \mathbf{H}) \) and linear form \( l(\mathbf{F}) \) are given by:

\[ c(\mathbf{F}, \mathbf{H}) = \int (\mathbf{V} \times \mathbf{F}) : \int \frac{1}{\varepsilon} (\mathbf{V} \times \mathbf{H}) \, d\Omega - k_0^2 \int \mathbf{F} : \mu \mathbf{H} \, d\Omega \]

\[ + \frac{k_0^2}{\varepsilon_0} \int \int \mathbf{n} \times (\mathbf{F} \times \mathbf{H}) \, d\Gamma \]

\[ l(\mathbf{F}) = 2j \int \int \mathbf{n} \times (\mathbf{H} \times \mathbf{n}) \, d\Gamma \]

and

\[ \hat{\mathcal{W}} := \{ \mathbf{A} \in \mathbf{H}(\text{curl}, \Omega), \mathbf{n} \times \mathbf{A} = 0 \text{ on } \Gamma_D \} \]  \hspace{1cm} (8)

For simplicity of notation, the relative permittivity \( \varepsilon \) and permeability \( \mu \) are considered as scalar magnitudes in the above expressions. In practice, the code supports the definition of permittivity and permeability as tensor magnitudes (i.e., matrices). We note that the above formulation is given in terms of the magnetic field, although a dual formulation in terms of the electric field can also be derived in a similar way.

3. hp Finite elements and automatic adaptivity

In this section, we describe a self-adaptive hp-Finite element method (FEM) that provides a sequence of solutions delivering exponential convergence rates of the error in the energy norm vs. the problem size for all the problems considered above. Neither the \( h \)- nor the \( p \)-version of the FEM is capable of delivering such fast convergence rates when analyzing rectangular waveguide discontinuities. They only produce algebraic convergence rates, leading to less accurate solutions.

3.1. hp-Finite elements (FE)

With each finite element, we associate element size \( h \) and order of approximation \( p \). In the \( h \)-adaptive version of FEM, element size \( h \) may vary from element to element, while order of approximation \( p \) is fixed (usually \( p = 1, 2 \)). In the \( p \)-adaptive version of the FEM, \( p \) may vary locally, while \( h \) remains constant throughout the adaptive procedure. Finally, a true hp-adaptive version of the FEM allows for local variations of both \( h \) and \( p \).

In our hp-adaptive FEM, which is an extension of the implementation described in [21,23], we support local anisotropic refinements in both \( h \) and \( p \). Support of anisotropic refinements dramatically increases the implementation complexity, but it is essential to achieve optimal convergence rates in the case of edge singularities or in the presence of boundary layers. Both situations occur in our waveguide problems of interest. To ensure continuity, we enforce the 1-irregularity rule and use the constrained approximation [33]. The simulator utilizes hexahedral edge (Nédélec) elements of variable order of approximation to discretize \( \mathbf{H}(\text{curl}) \). FE spaces associated to those elements satisfy the commutativity of the so-called de Rham diagram when equipped with appropriate projection based interpolation operators (see [15,34]). This commutativity property is essential for guaranteeing convergence and stability of the FE formulation for electromagnetics [15] and, in particular, to avoid the presence of unphysical spurious modes [35].

As advertised above, the main advantage of the hp-FEM resides on the proof (see [36–42]) showing that it can achieve exponential convergence for elliptic problems with a piecewise analytic solution, whereas \( h \)- or \( p \)-FEM converge at best algebraically. But to attain that goal, one needs to develop an hp-adaptive strategy in order to orchestrate an optimal (or quasi-optimal) distribution of element size \( h \) and polynomial order of approximation \( p \) throughout the computational grid.

There exist several hp-adaptive strategy schemes, which have been recently summarized in [43]. Most of them are based on an \( h \times p \)-technique (e.g. [44]), where one \( h \)-adaptive step is followed by a \( p \)-adaptive step. These \( h \times p \) methods do not incorporate any criteria to decide if a given element is more optimally refined via an \( h \)-refinement vs. a \( p \)-refinement. hp methods in which \( h \) and \( p \) refinements may be performed simultaneously in one step of the iterative process, can be further divided into two groups: (a) those based on (a possibly automatic) “detection” of singularities (e.g. [45]), and (b) those based on the use of a reference solution function to decide optimal hp-refinements. While the first group of algorithms typically provide optimal results in a variety of academic problems, its performance on complex real-world problems is uncertain, mainly because general 3D anisotropic \( h \)- and \( p \)-refinements are typically unsupported.

In this work, we focus on hp-adaptive algorithms intended to maximize the reduction of the energy-norm interpolation error computed using a global reference solution with respect to the added number of unknowns. Specifically, our 3D adaptive strategy builds upon the hp-adaptive strategies described in [33,21], and it is a 3D extension of the 2D hp-adaptive strategy for waveguide problems described in [13].

3.2. Fully automatic 3D hp-adaptivity

The 3D \( hp \)-self-adaptive refinement strategy is an iterative process intended to maximize the error decrease rate measured in the energy norm with respect to the added number of unknowns. Given a coarse \( hp \)-grid, several optimal refinements are performed to maximize the error decrease rate. After those refinements are executed, a new \( hp \)-mesh is delivered, and the entire process is iterated until the prescribed error tolerance is satisfied.

The energy-norm \( \| \cdot \|_E \) used for electromagnetic problems is the \( \mathbf{H}(\text{curl}) \) norm defined as:

\[ \| \mathbf{H} \|_E = \sqrt{\| \mathbf{H} \|_{L^2}^2 + \| \mathbf{\nabla} \times \mathbf{H} \|_{L^2}^2} \]  \hspace{1cm} (9)

In the above definition, it is possible to include material properties in the form of weights and the frequency (using \( k_0^2 \) as weight factor.
of $|H|_{p^2}$; see (7)), and even the boundary conditions at the ports by using a norm obtained from the bilinear form of the problem. These variations have not reported significant variations in the behavior of the adaptivity for the waveguide discontinuities analyzed in the paper.

Within each iteration of the $hp$-adaptive strategy, one needs to estimate the error associated to each element. Since there exist multiple (sometimes thousands of) possibilities in which one element can be refined in $h$ or $p$, conventional a posteriori error estimators providing only one number per element cannot be used to select optimal refinements. To overcome this problem, we globally refine the given $hp$-mesh both in $h$ and $p$ to yield a $h/2, p + 1$ fine mesh, by breaking each element into eight element sons and raising the order of approximation uniformly by one. Then, the problem of interest is solved on the fine mesh, which in 3D contains approximately $8 + (p + 2)^2/(p + 1)^2$ more degrees of freedom than the corresponding coarse mesh.\footnote{In practice, when the fine grid problem is too large, its solution can be approximated with a geometric two-grid solver and partially converged solutions, as described in [46].} The difference between fine and coarse grid solutions is an approximation of the error function over the coarse grid, and it can be used to determine the error over each element.

However, one should not be tempted to refine those elements delivering maximum error, since in the case of wave propagation problems, the solution over a highly refined element may still contain a large error due to dispersion (pollution) produced by poorly discretized elements in other regions of the mesh. To overcome this problem, one should evaluate for each element the energy norm of the fine grid solution minus its projection over the coarse grid. This number expresses the distance of each coarse grid FE space to the fine grid solution. Those FE space with maximum distance are the ones that should be refined.

As mentioned above, the projection of the fine grid solution over the coarse grid and, more generally, the projection over those grids associated to all possible competitive refinements, is used to compute the error decrease rates. Ideally, each projection involves simply the solution of a new global finite element problem, leading to a prohibitively expensive method. To overcome this problem, we employ the projection based interpolation operator $\Pi_{p+1}$ described in [47–49,34]. This operator is defined as the sum of a Whitney interpolant with edge, face, and element interior contributions. The Whitney interpolant $H_{i}$ is built as the lift (using Whitney shape functions) of the average value over each edge of the tangential component of the interpolated function. Intuitively, each edge contribution $H_{e}$ is intended to minimize the norm of the curl of $H_{e} - (H - H_{0})$ and such that $H_{e} - (H - H_{0})$ is discrete divergence free in the proper inner-product. Once all edge contributions are summed into $H_{i}$, each face contribution $H_{f}$ aims at minimizing the norm of the curl of $H_{f} - (H - H_{0} - H_{e})$ and such that $H_{f} - (H - H_{0} - H_{e})$ is discrete divergence free in the proper inner-product. The element interior contribution is built in a similar way. The proper inner products and norms mentioned above are $L^2$ for the element interiors, $H^{1/2}$ for the faces, and $H^{1/4}$ for the edges, where $\epsilon$ is a small positive number. Notice that these norms are different than those needed for the 2D case, where $L^2$ was utilized for the faces and $H^{1/2}$ for the edges. The precise mathematical formulation of the projection based interpolation operator can be found in [34]. This operator has the following properties:

- Conformity: the union of element interpolants is globally conforming.
- Optimality: the interpolation behaves asymptotically, both in terms of $h$ and $p$, in the same way as the actual approximation error.

Furthermore, the de Rham diagram equipped with this operator commutes, which is essential to prove stability of the resulting FE formulation.

Mathematically, within each iteration, given a coarse $hp$-grid and the reference solution computed in the fine grid $H = H_{h/2,p+1}$, the objective is to find a new optimal $hp$-grid such that it maximizes the following decrease rate of the interpolation error:

$$hp = \arg \max_{hp} \frac{||H - \Pi_{hp} H||_{L^2} - ||H - \Pi_{hp} H||_{L^2}}{N_{hp} - N_{0p}},$$

where $N_{hp}$ is the number of dofs of a given $hp$ grid. Notice that $hp$ has to be coarser than $h/2, p + 1$. However, in some instances, the order of approximation associated to the optimal $hp$-grid for a particular element edge, face, or interior could be lower than that of the original $hp$-grid. This occurs, for instance, when an edge, face, or interior is $h$-refined, and the order of approximation is lowered in some (or all) new element edges, faces, or interiors. As a result, we obtain a sequence of non-nested $hp$-meshes.

Additionally, to make this optimization problem tractable from the implementation point of view, and following the construction of the projection-based interpolation operator, we divide it into three simpler sequential problems: (1) optimal refinement of edges, (2) optimal face refinements for given edge refinements, and (3) optimal refinements of element interiors for given edge and face refinements.

Within each of the three steps (namely, edge, face, and interior refinements), we would like to compute the decrease rate of the interpolation error (10) for each possible $hp$-grid up to the level $h/2, p + 1$. However, the number of possible 3D refinements for each element interior is large, which makes this approach prohibitively expensive for practical applications. Thus, we compute an ordered sequence of refinements that only includes those combinations that are most likely to provide a larger decrease rate of the interpolation error. For example, in the case of edge refinements, we first evaluate the error decrease rate for a $p$-refinement. Then, for an $h$-refinement, we start by considering two edge sons with $p = 1$. After evaluating the error decrease rate for this case, we raise $p$ only on that edge son delivering maximum error, and we iterate the process. Thus, we avoid evaluation of the error decrease rate for most combinations of $p$ in the two edge sons. For the case of 3D element interiors, this algorithm provides dramatic savings with respect to the evaluation of all possible refinement combinations.

After a local edge-based competition determines the optimal $hp$-refinement for each edge, we actually perform refinements only on those edges that are within a percentage $X$ of the maximum error decrease rate (global competition). Once optimal edge refinements have been performed, we proceed with the face refinements in a similar manner. At this point, the starting grid topology has been determined, and the minimum order of approximation over the face is dictated by the already selected order of approximation over the edges and the minimum rule. This rule establishes that an order of approximation for an edge should always be smaller or equal than the minimum of the orders of approximation of the adjacent faces. Similarly to the case of edge optimization, we compute the decrease rate of the interpolation error for each face, using a local competition between $h$- and $p$-refinements. In general, each face could be refined into two or four element sons, but if a refinement is inherited from the edge
 optimization step, the number of possible local refinements decreases. After the local competition takes place and the best possible refinement is selected for each face, we refine those faces whose error decrease rate is within the X percentage of the maximum error decrease rate over all faces (global competition). For the refinement of element interiors, our starting point are the already selected edge and face refinements along with the minimum rule, which set the starting topology for the local 3D optimization problem. Again, a local h- and p-refinement competition based on the error decrease rate is followed by a global competition that refines only those interiors whose error decrease rate is within the X percentage of the maximum error decrease rate over all elements. In order to illustrate the effect of X in the error convergence of the hp-adaptivity consider the plots of Fig. 3. The plots show how the adaptive algorithm is quite insensitive to the refinement threshold X. However, the optimum value of X is problem dependent. Fig. 3(a) shows the convergence of the error of a problem with a smooth field solution while Fig. 3(b) correspond to the case of a singular solution. It is observed that a relative low value of X performs better (for low error levels) in the case of the smooth solution while the opposite is true for the non-smooth solution.

Based on these observations, the X factor is chosen typically within 33% and 50%.

Once all optimal refinements have been performed, additional undesired h-refinements may be needed to maintain the 1-irregularity rule. At this point, the order of approximation p of the new element interiors arising as a result of these undesired refinements need to be determined. For that purpose, we first need to set up a minimum p distribution over the element interiors ensuring that at least the coarse grid error level is satisfied, that is, the numerator of (10) should be non-negative. Then, the aforementioned local competition is used to determine the optimal order of approximation p for element interiors. Finally, element faces and element edges orders of approximation associated to undesired refinements are determined using the minimum rule.

The described adaptive strategy is very general, and it applies to $H^1$, $\mathbf{H}(\text{curl})$, and $\mathbf{H}(\text{div})$-conforming discretizations. Moreover, since the mesh optimization process is based on minimizing the interpolation error rather than the residual, the algorithm is problem independent and it can also be applied to nonlinear and eigenvalue problems.

For the analysis of rectangular waveguide discontinuities, we have implemented the boundary conditions described on the mathematical formulation as well as several postprocessing routines for computing the S-parameters, graphics, and so on. We have also fixed a few problems and accelerated a number of key features within the software such as the computation of the projections during the hp-adaptivity. Additionally, we have fixed a problem related to the choice of the optimum p for face and interiors with 3D vector-field problems. Now, the algorithm uses a fast Fourier transform like approach to estimate the direction of maximum variation and set the, in general, anisotropic choice of p accordingly. Moreover, a “deadlock” problem arising only in 3D geometries that disabled further refinements in a particular set of four elements have been identified and fixed by adding a flag that prevented that particular situation from arising in our waveguide problems of interest.

The 3D hp-adaptive algorithm is implemented in FORTRAN90, and its level of complexity is at least one order of magnitude larger than the corresponding 2D algorithm described in [13]. Since the 3D algorithm does not only require an additional step for optimization of 3D element interiors, but it also incorporates more computationally efficient implementation in order to be competitive in practical waveguide applications. Specifically, since the number of possible competitive refinements for a given element dramatically increases from 2D to 3D, the search algorithm to decide the optimal selected refinement has been significantly improved to minimize the computational time invested during this operation.

4. Numerical results

The fully automatic hp-adaptive FE strategy described above has been applied to the analysis of a wide variety of discontinuity problems as well as more complex structures obtained by combining several discontinuities. A selection of those results are presented in the following.

$\text{TE}_{10}$ mode excitation has been used in all the structures. In most cases, the ratio of the broad dimension $a$ to the narrow dimension $b$ of the rectangular waveguide sections is set to $a/b = 2$. In general, results correspond to a given frequency which is selected to be in the middle of the monomode region. In some cases, the analysis is performed for a large number of frequencies within a given frequency region in order to characterize its frequency response.

Quite coarse meshes are used as initial grids in order to assess the robustness of the hp strategy in the context of real engineering analysis in which the initial mesh has to be as coarse as possible to
simplify the mesh generation process. The refinement threshold $X$ mentioned in Section 3.2 is set to 33%.

The scattering parameters obtained using the $hp$-FEM are compared with values computed with the semi-analytic mode matching (MM) method. The FEM scattering parameter results delivered from the $hp$ adaptivity are comparable in accuracy (and in many cases more accurate) than semi-analytic results. In this context, we remark that these semi-analytic results are typically considered as a reference for the engineering analysis of discontinuities in rectangular waveguide technology, such as the structures shown below. In addition, the $hp$-FE technology enables modeling of more complex structures which cannot be solved using semi-analytic techniques.

The computer system used to run the $hp$-adaptivity is a single computer with 24 GB of available RAM for computations. The computer has two dual-core CPUs. A direct multi-frontal solver MUMPS [50] is used to solve the system matrix of equations.

4.1. Waveguide section

The first structure is a simple rectangular waveguide section. The ratio of the broad dimension $a$ to the narrow dimension $b$ is set to $a/b = 2$. The length of the waveguide section is $l = 5a$. Since there is no discontinuity in the translational symmetry for this structure, the field solution is known analytically (the field is the one of the incident TE$_{10}$ mode) and hence, $S_{11} = S_{22} = 0$ and $S_{21} = S_{12} = \exp(-j\beta l)$, where $l$ denotes the waveguide section length. Note that the structure is suitable to be analyzed by the 2D formulation, either as a H-plane or an E-plane structure. This structure has been selected as a first verification of the code, specifically to test the boundary conditions at the ports and the control of the dispersion error.

The convergence history (up to an energy error below 0.0001%) achieved with $hp$ adaptivity is displayed in Fig. 5 together with results of $h$-adaptivity for several values of uniform order $p$. Logarithmic scales for the error and the number of degrees of freedom ($N_{\text{dof}}$) are used. According to FEM convergence theory, straight lines for $h$-adaptivity are expected. This is clearly observed in Fig. 5. Since the field solution is smooth for this case, the slopes are increasing as $p$ augments. The exponential convergence of $hp$-adaptivity is indicated by a slight curvature corresponding to the $hp$ case. The initial mesh is shown in Fig. 4(a). Different colors indicate, according to the scale on the right, the order of approximation $p$ of the elements. Due to the smoothness of the field solution, a sequence of meshes with increasing order of approximation is expected. An intermediate $hp$-mesh during the adaptive process is shown in Fig. 4(b), specifically, the 4th mesh, which corresponds to an error below 0.01%. In addition to the use of higher order elements, refinements (basically of uniform type in $h$ and $p$) take place only along the $H$-plane of the structure due to the null variation of the field along the direction perpendicular to that plane.

The computed values of $S_{11}$ and $S_{21}$ with $hp$-adaptivity are shown (for the first $hp$ meshes) in Table 1, together with the analytical ones. It is emphasized the low dispersion error achieved (error in $\text{arg}(S_{21})$).

4.2. H-plane right angle bend

An H-plane 90° bend is analyzed next. The structure is shown in Fig. 6. Being a H-plane structure, it may be analyzed more efficiently (actually, it was) using the 2D formulation of [13]. The reason to include it here is for illustration and comparison purposes.

The structure is analyzed by exciting port 1 (on the left). The bend is a common part of microwave circuits. The initial mesh used for the analysis is shown in Fig. 7(a). Despite the coarseness of this mesh, the $hp$-strategy achieves an error below 0.5% after just 3 iterations. The convergence history (up to an error as low as 0.01%) is shown in Fig. 8 together with results of $h$-adaptivity for several values of $p$. Note that log scale for the energy error (in percent of the energy norm) in the ordinate axis is used while a scale corresponding to $N_{\text{dof}}$ is employed in the abscissa axis. This is the type of plot that will be used in the remainder of the paper. Note that the abscissas scale corresponds to $N_{\text{dof}}$ while abscissas axis ticks should be read as $N_{\text{dof}}$ in the plots. According to [41] and references therein, an approximate straight line should appear with such scales, indicating the theoretical exponential convergence that can be achieved with an optimal $hp$ adaptive strategy. This is observed in the figure. In contrast, a curvature (towards

![Fig. 4. hp-meshes of the waveguide section.](image)

![Fig. 5. Convergence history for the waveguide section.](image)

<table>
<thead>
<tr>
<th>$S_{11}$</th>
<th>$S_{21}$</th>
<th>$\text{arg}(S_{21})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iter. 1: 0.50317992</td>
<td>0.357711764</td>
<td>52.5598924</td>
</tr>
<tr>
<td>Iter. 2: 0.066255241</td>
<td>0.997207739</td>
<td>91.0591254</td>
</tr>
<tr>
<td>Iter. 3: 0.000025470</td>
<td>0.999998601</td>
<td>73.7793384</td>
</tr>
<tr>
<td>Iter. 4: 0.000000002</td>
<td>0.999999995</td>
<td>73.7694414</td>
</tr>
<tr>
<td>Analytical: 0.0</td>
<td>1.0</td>
<td>73.7694410</td>
</tr>
</tbody>
</table>
higher error levels) is observed for the \( h \)-adaptive cases. The curvature denotes the algebraic character of the convergence of the \( h \)-adaptive procedures.

With respect to the meshes delivered by the \( hp \)-adaptivity, the final \( hp \)-mesh is displayed in Fig. 7(b). Refinements take place only on the \( H \)-plane and no refinements occur along the direction of the narrow edge of the waveguide in which there is no variation of the field. Thus, the automatic \( hp \)-adaptivity in 3D provides for this case meshes that are basically the extrusion of the ones delivered by its counterpart in 2D. With respect to the refinements on the \( H \)-plane, heavy \( h \)-refinements of the mesh around the corner are observed. The \( hp \)-strategy in this case tends also to increase \( p \). Actually, most of the elements of the final mesh that are not close to the corner have reached a very high (even the maximum) \( p \) order.

As in the previous case, \( S_{21} = S_{12} \) and \( S_{22} = S_{11} \). The \( hp \)-results for \( S_{11} \) and \( S_{21} \) are shown (for some of the \( hp \) meshes) in Table 2. The scattering parameters computed with the \( hp \)-FEM are compared with those obtained with a MM technique. Only five significant digits are shown in the table as the MM results are presumed to have no more than five digits of accuracy.\(^5\) \( hp \)-FEM results agree with those provided by MM; the difference is below 1% after the second iteration. After the fourth/fifth iteration, the FEM results seem to be more accurate than those provided by the MM, as implied by the convergence pattern shown in the table.

4.3. \( E \)-plane step

A change in the height of a rectangular waveguide has been chosen to illustrate the behavior of the automatic \( hp \)-adaptivity with \( E \)-plane structures.

\(^5\) This is concluded after increasing the number of modes in the modal expansions and observing fluctuations in the sixth digit of the S-parameters.
Table 2
Scattering parameters for the H-plane 90° bend.

<table>
<thead>
<tr>
<th></th>
<th>$S_{11}$</th>
<th>$S_{21}$</th>
<th>$\arg(S_{11})$</th>
<th>$\arg(S_{21})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iter. 1</td>
<td>0.23576</td>
<td>0.80093</td>
<td>1.292°</td>
<td>96.029°</td>
</tr>
<tr>
<td>Iter. 3</td>
<td>0.41563</td>
<td>0.90953</td>
<td>26.866°</td>
<td>-63.134°</td>
</tr>
<tr>
<td>Iter. 6</td>
<td>0.41556</td>
<td>0.90956</td>
<td>26.865°</td>
<td>-63.135°</td>
</tr>
<tr>
<td>Iter. 8</td>
<td>0.41554</td>
<td>0.90957</td>
<td>26.866°</td>
<td>-63.133°</td>
</tr>
<tr>
<td>MM</td>
<td>0.41554</td>
<td>0.90957</td>
<td>26.866°</td>
<td>-63.133°</td>
</tr>
</tbody>
</table>

$S_{22}$; however, $|S_{22}| = |S_{11}|$.\(^6\) The results for $S_{22}$ are shown for some of the hp meshes in Table 3. The scattering parameters computed with the hp-FEM method are of comparable accuracy to those obtained with a MM technique.

The “dual” structure to the E-plane step is the H-plane step, i.e., a change of width in the waveguide. Similar conclusions to the ones given here for the E-plane step hold for the H-plane step and its results have been omitted.

4.4. Filters with dielectric resonators

The structure considered in this section is a dielectric resonator (DR)\([51]\) coupled through slots to two rectangular waveguide sections (see Fig. 12). The slots are an example of waveguide discontinuity. The DR consists of a high permittivity and low losses dielectric material. The high contrast between the value of the permittivity of the dielectric and the surrounding medium (air) makes it behave as a cavity concentrating the electromagnetic field. The DR has a selective frequency response by itself. Note that the DR is placed on top of a pedestal which consists of another circular dielectric (this time of much lower permittivity). Filters based on DRs have become a key element in many filtering applications. Indeed, dielectric-loaded waveguide filters are frequently found in satellite and mobile communication systems, due to their advantages in terms of mass and volume reduction, low losses, and thermal stability\([52,53]\).

The application of the mode matching method (MM) to the analysis of cylindrical DRs is straightforward if each element of the geometry (including the encapsulation) is contained in the coordinate surfaces of a cylindrical coordinate system. The resonator is modelled as a concatenation of radial waveguides, and each cylindrical discontinuity is analyzed with MM using radial modes instead of the plane modes defined in cylindrical waveguides. However, when the DR is encapsulated by a rectangular metallic box, as it is the case of the structure under analysis, the interaction between the radial modes defined at the border of the DR, and the plane modes defined at the walls of the rectangular cavity must be analyzed by a different method. The method applied here is a version of the boundary integral-resonant-mode expansion (BI-RME)\([54]\) that is used only to extract a multimodal circuit model of the transition between rectangular ports defined on the encapsulating box, and the radial waveguide at the border of the cylindrical DR inside the metallic box. Then, this block is combined with the MM analysis of the irises and the MM analysis of the cylindrical DR (see\([55]\)).

The modeling of this structure is a challenge for a numerical method due to the high contrast in the permittivity and the geometry. The frequency response over part of the monomode region of the waveguide sections of the ports is shown in Fig. 15. It exhibits a resonant character around two frequencies: $k_0/k_c = 1.4265$ (11.2 GHz), and $k_0/k_c = 1.6300$ (12.8 GHz) being $k_c$ the

---

\(^6\) It can be shown that for a lossless microwave network, matrix $|S|$ is unitary. From that property it is deduced that $|S_{21}| = |S_{12}|$ and $|S_{11}| = |S_{22}|$. The equality in magnitude and phase between $S_{21}$ and $S_{12}$ is again a consequence of the reciprocity of the problem.
Fig. 9. E-plane step \((b_1 - a/2, b_2 - h_2/2, l_1 - l_2 - 2a)\).

Fig. 10. \(hp\)-meshes of the E-plane step.

Fig. 11. Convergence history for the E-plane step.

wavenumber corresponding to the cut-off frequency. The frequency response is the one provided by the MM-IE technique mentioned above [55]. The results provided by the \(hp\)-adaptivity at
Table 3
Scattering parameters for the E-plane step.

| $|S_{21}|$ | $|S_{11}|$ | $\arg(S_{22})$ | $\arg(S_{12})$ |
|----------|----------|----------------|----------------|
| Ker. 1   | 0.21464  | 0.702688       | 1.292°         | 0.368°         |
| Ker. 3   | 0.36704  | 0.93002        | 26.866°        | 59.284°        |
| Ker. 5   | 0.36945  | 0.92924        | 26.865°        | 58.083°        |
| Ker. 12  | 0.36962  | 0.92918        | 26.866°        | 57.986°        |
| MM       | 0.36962  | 0.92918        | 26.866°        | 57.986°        |

certain frequencies are also displayed. An excellent agreement is observed. The first resonant mode of the DR corresponds to the so-called “magnetic dipole mode”. Its field is similar to the one of a magnetic dipole. This field solution is relatively smooth. However, the second resonance exhibits a field singularity in the air-dielectric interface. The convergence history of the two sequences of $hp$ meshes provided by the adaptivity when running at the first and second resonance frequencies is shown in Fig. 13. A slightly slower convergence is observed for the second resonance mode due to the presence of the singularity. Despite that and the strong contrast in $\varepsilon_r$ present in the structure, a nice exponential convergence pattern is obtained. The analysis have been made using a symmetry plane. The initial mesh in both cases is displayed in Fig. 14(a). Two views of intermediate $hp$-meshes obtained during the adaptive process are displayed: Fig. 14(b) and (c) for the first and second resonances, respectively. It is observed how the field singularity of the second resonance produces a mesh capable of

**Fig. 12.** Slot coupled dielectric loaded resonator ($a = 19.05$ mm, $b = 9.52$ mm, $a_c = 9$ mm, $b_c = 9$ mm, $c_c = 11$ mm, $D_{res} = 5.45$ mm, $H_{res} = 2.208$ mm, $D_{sup} = 6.10$ mm, $H_{sup} = 2.35$ mm, $\varepsilon_{r,1} = 29$, $\varepsilon_{r,2} = 2.53$).

**Fig. 13.** Convergence history for the dielectric resonator filter.
reproducing strong boundary layers at the dielectric-air interface in comparison with the mesh corresponding to the first resonance.

5. Conclusions

A three-dimensional fully automatic hp-adaptive finite element method for the characterization of rectangular waveguide discontinuities has been presented. Its application to a wide variety of discontinuity problems in microwave engineering has demonstrated the versatility and accuracy of the method. Exponential convergence of the error has been obtained.

Fig. 14. Initial mesh and zoom over meshes of 12th iteration for half of the dielectric resonator filter.

Fig. 15. Frequency response of the dielectric resonator filter.

References


