Domain decomposition method for electromagnetic modelling and quantitative microwave imaging in large-scale three-dimensional configurations

Ivan Voznyuk

Ivan Voznyuk. Domain decomposition method for electromagnetic modelling and quantitative microwave imaging in large-scale three-dimensional configurations. Electromagnetism. Aix-Marseille Université, 2014. English. NNT : . tel-01279590

HAL Id: tel-01279590
https://amu.hal.science/tel-01279590

THÈSE DE DOCTORAT

Pour obtenir le grade de
Docteur d'Aix Marseille Université

École doctorale n°352
Discipline : Physique et Sciences de la Matière
Spécialité : Optique, Photonique et Traitement d'Image

Présentée par

Ivan Voznyuk

Méthodes de décomposition de domaines pour la modélisation électromagnétique et l'imagerie micro-onde quantitative dans des configurations tridimensionnelles de grande taille

---

Soutenance le 29 septembre 2014
devant le jury composé de :

M.	Stéphane LANTERI	Directeur de Recherches INRIA	Rapporteur
M.	Christophe GEUZAINE	Professeur, Univ. de Liège	Rapporteur
Mme	Hélène ROUSSEL	Professeur, Univ. Pierre et Marie Curie	Examinatrice
M.	Alain REINEIX	Directeur de Recherches CNRS	Examinateur
M.	Lionel PICHON	Directeur de Recherches CNRS	Examinateur
M.	André NICOLET	Professeur, Univ. Aix-Marseille	Examinateur
Mme	Amélie LITMAN	Maître de conférences, Univ. Aix-Marseille	Directrice de thèse
M.	Hervé TORTEL	Maître de conférences, Univ. Aix-Marseille	Co-Directeur de thèse

Financement : Bourse Ministérielle

Acknowledgements

I would like to write this chapter, so important to me, in French, in a language which allowed me to open so many doors and never let me feel a foreigner here.

Je n'aurais jamais pu réaliser cette thèse scientifique sans le soutien d'un grand nombre de personnes dont la bonne humeur et la générosité m'ont permis de progresser et terminer ce travail.

Je tiens tout d'abord à adresser mes remerciements les plus vifs et sincères à Amélie Litman et Hervé Tortel pour avoir dirigé cette thèse et m'avoir permis de la réaliser dans les meilleures conditions. Leurs compétences et rigueur scientifique m'ont beaucoup appris. Amélie et Hervé, je tiens aussi à vous remercier pour la confiance énorme que vous m'avez accordée, elle m'a permis d'apprendre à travailler tout seul, en autonome, quand il fallait.

J'exprime tous mes remerciements à l'ensemble des membres de mon jury : Madame Hélène Roussel et Messieurs Stéphane Lanteri, Christophe Geuzaine, Alain Reineix, Lionel Pichon et André Nicolet.

Mes remerciements s'adressent ensuite à toute l'équipe HIPE, notamment : Pierre Sabouroux, Christelle Eyraud, Jean-Michel Geffrin et Anar Rahimov. Je remercie énormément beaucoup Jessica Benedicto pour avoir partagé son expérience avec moi et m'avoir soutenu à des moments très durs pour moi. Puis, je ne peux pas oublier nos meilleurs stagiaires : Clara, Mathieu et Nicolas.

Ces remerciements seraient incomplets si je n'en adressais pas à plusieurs membres de l'équipe CLARTE : Stefan Enoch, Redha Abdeddaim, Nicolas Bonod et Gérard Tayeb pour toutes les discutions pertinentes que nous avons eues pendant ces 3 ans.

Il m'est impossible d'oublier l'administration de l'Institut Fresnel. Pas seulement pour avoir géré toute la partie administrative (qui était, du coup, la partie la plus dure dans ma thèse), mais aussi pour le soutien personnel. Donc, Nelly, Claire, Laëtitia, Fréderic, Marc, Nadège, Guylene et Magali, ce merci est pour vous ! Je remercie également tous les doctorants et les autres membres de l'Institut Fresnel, notamment ceux avec qui j'ai eu l'occasion de travailler et les autres simplement pour les bons moments partagés.

Mention spéciale à Elodie Georget qui m'a supporté et soutenu pendant ma thèse, qui m'a aidé dans tous les aspects de la vie. Je te remercie, Élodie, pour m'avoir appris cette langue tellement belle et à la fois très difficile. Je voudrais te dire merci pour ton soutien pendant mes périodes de doutes.

Je remercie également toutes les personnes qui m'ont permis d'oublier momentanément le travail dans des soirées, repas, sport ou dans la vie quotidienne : Ilya, Romain et Victoria.

Enfin, les mots les plus simples étant les plus forts, ma reconnaissance va à ceux qui m'ont soutenu le plus et qui ont cru à moi malgré tout – à ma famille et en particulier à ma maman et mon papa. Je vous aime.

J'en oublie certainement encore et je m'en excuse...

Je dédie cette thèse à ma femme.
Désolé, je suis en retard, comme d'habitude.

Table des matières

CHAPITRE 1

General introduction

Electromagnetic analysis has become part and parcel of a big amount of scientific and engineering research work since *J.C. Maxwell* completed the electromagnetic theory in 1873 [1]. Examples of such work are radar and antenna simulations, geo-, bio- and medical-electromagnetics, optics, and so on. The predictive power of the Maxwell's theory is valid from the static to optical regimes and from subatomic to intergalactic length scales, as proven over the years. The modern electromagnetic analysis consists in solving a set of Maxwell's equations which can only be resolved analytically for very few simple cases, while an accurate and complete analysis of complex ones can be accomplished only through techniques that solve the involved differential equations numerically with the help of modern computers.

The general form of the Maxwell's equations can be formulated as a hyperbolic system of partial differential equations, and then can be solved with Godunov-type methods [2, 3], Finite-difference time-domain (FDTD) methods [4–7] and so on. In this work, we aim at considering the simulation of *time-harmonic* electromagnetic wave propagation problems. A field is referred to be time-harmonic when the field quantities in Maxwell's equations are harmonically oscillating functions with a single frequency. We can point out the first- and second-order form of these equations. The first-order form of time-harmonic Maxwell's equations represents a system of two equations which couples both the electric and magnetic field. Recently, discontinuous Galerkin methods [8] have been considered in order to solve this set of equations and have showed their efficiency in various domains in scientific computing [9]. In this work, we focus on the second-order form of time-harmonic Maxwell's equations, that is represented by a *Helmholtz* differential equation describing the evolution of the electric or the magnetic field. The analytical solution of this equation when the field is generated by a dipole is called a *Green* function [10], and as for the Maxwell's theory, it exists unfortunately for very few cases.

There are many different numerical techniques for obtaining approximate solutions to boundary-value problems of mathematical physics. One can note Volume integral equation methods, such as the Method of Moments (MoM) [11, 12] with its different variations [13, 14] and Boundary Element methods (BEM) [15], which are well adapted to homogeneous configurations. These methods profit from the existence of the Green function in closed form.

When we deal with a complex case whose Green function is unknown, the *Finite Element Method* (FEM) has totally proven its efficiency during a history of about 60 years. FEM was firstly proposed in the 1940s and its use begun in the 1950s for aircraft design. After this, the method was applied extensively to problems of structural analysis and increasingly to problems in other fields. Nowadays, the FEM method has become a powerful numerical technique for computations in a large amount of scientific domains, including computational electromagnetics [16].

One of the main advantages of this method is that it allows to treat complex media (such as anisotropic or inhomogeneous) and different types of geometry. The oscillatory nature of the high-frequency electromagnetic scattering problems requires a great number of grid points per wavelength (typically 10 to 15 for the linear finite element method) [17], and therefore leads to large-scale system of equations. For such problems, solving the main system of linear equations by a direct method entails memory and CPU requirements that rapidly overwhelm even the largest currently available resources. Furthermore, the convergence of classical iterative methods, such as GMRES [18], BICG [19] or BICGSTAB [20] for example, which are well suited for complex matrices which are neither Hermitian nor definite (even if they are symmetric in general) is chaotic or failing [17, 21]. In addition, finding a good preconditioner to solve these systems of linear equations is often treated as a combination of art and science, and is very often not evident [21–23]. All this makes the numerical resolution of the Helmholtz equation and related optimal control problems in heterogeneous media at high wave number a challenging problem for the classical Finite Element Method.

A way to overcome these difficulties can be, for example, to apply high-order finite-elements [24], or *Domain Decomposition* (DD) technique. Indeed, time harmonic wave propagation phenomena has become a privileged field of application of *Domain Decomposition Methods* (DDMs) during the last decade. The idea of the Domain Decomposition was firstly introduced by Lions in the end of the 1980s for the Laplace equation [25, 26]. But solving the Helmholtz equations describing wave propagation in the frequency domain does not have the same nice properties than the solution of more standard elliptic equations such as Laplace equation, which makes impossible to apply the idea of overlapping Schwartz algorithm [27]. This probably explains why a big amount of research work on DDM for such models started a bit later (in the middles 1990s in fact), since Després extended the DD idea for the two-dimensional Helmholtz equations [28–33]. The principal DD idea is to split the entire computational domain into smaller *non-overlapping subdomains* and to solve a sequence of similar subproblems on these subdomains. The boundary conditions are adjusted iteratively by ad hoc transmission conditions between neighboring subdomains. Després also introduced a relaxation scheme [34] that greatly accelerated the iterative process. Later Stupfel extended this method by using an "onion-like" partition of the computational domain improving the efficiency of the transmission condition and overall performance of this DDM [35].

Among a variety of DDMs, the *Finite Element Tearing and Interconnecting* (FETI) method based on DD idea with *Lagrange multipliers* is showed to be an amazingly powerful technique with an excellent numerical scalability [36–38]. Essentially, a FETI method can be viewed as an iterative substructuring method where Lagrange multipliers are introduced at the internal interfaces between subdomains in order to enforce the continuity of the electric field. The role of the Lagrange multipliers consists in representing the unknown boundary condition between the subdomains. Over the past years, FETI is proved to be suitable for the fast sequential and parallel iterative solution of large-scale systems of equations arising from the finite element discretization of partial differential equations [39]. This method, originally proposed by Farhat and Roux [40] for computational mechanics and acoustics problems, has been combined with the Dual-Primal (DP) idea [39] and applied to electromagnetic problems [41]. As a vector-element implementation of the FETI-DPH method [42], the resulting one named FETI-DPEM [41], distinguishes itself through the introduction of one Lagrange multiplier per internal interface and through the following construction of an *explicit interface equation*. The advantage, here is that, contrary to the classical FEM approach, the resulting linear system can be solved by an iterative method. The FETI-DPEM method was shown to be scalable with respect to the size of finite elements, the number of subdomains and conditionally scalable with respect to the frequency [41]. In order to accelerate the convergence of the FETI-DPEM method and to make it fully scalable with respect to the frequency, in 2007 a new approach, named FETI-DPEM2 method, was extended by Li and Jin [43] for the electromagnetic problems. This method combines the DP idea with an implementation of two Lagrange multipliers per interface [36] in order to obtain a mixed (Neumann- and Robin-type) boundary condition for the internal interfaces.

Indeed, one of the key-points in the Domain Decomposition methods is the way the boundary conditions are imposed. In order to accelerate DDMs, there are some research work which aim at improving the transmission boundary condition between subdomains [44–46], i.e. on internal boundaries. One can also find some research work devoted to the extension of the Domain Decomposition algorithms by prescribing on external boundaries a new Absorbing Boundary Conditions (ABC) [47] of the first- and the second-order [35, 48]. To deal with subdomains with nonconformal interfaces, one has to introduce two sets of unknown Lagrange multipliers and develop special DDM in order to couple the subdomains. Among this type of DDM, the most known are the mortar element method [49] and the cement element method [50], which have been investigated in computational mechanics and applied mathematics communities respectively. Different Domain Decomposition methods based on the cement elements have been successfully applied to large-scale electromagnetic simulations [51, 52] and become a base for creating a new hybrid finite element-boundary integral DDM [53]. We start from a mesh for the entire problem and then, thanks to a greedy-like algorithm provided by an automatic meshing partition software METIS [54], we divide the computational domain into a set of non-overlapping subdomains whose interfaces are always conformal.

The ***first goal*** of this doctoral thesis is thus to investigate and develop a powerful tool for an efficient simulation of two- and three-dimensional direct electromagnetic problems for large-scale configuration with conformal meshes. For this aim, we do not consider cement type methods, but we propose here to restrict ourselves to the classical FETI-DPEM2 method by enforcing different types of boundary condition not only on the degrees of freedom (DOFs) related to the internal interfaces, but also to the ones related to the *corner* DOFs (we denote by "corner" the geometrical entities which belong to more than two subdomains). This method, called the FETI-DPEM2-full [55], will be applied to large scale electromagnetic and, in particular, scattering problems with complex geometries, which will allow to study its efficiency from an applied point of view. We aim as well at presenting all the implementation issues related to the algorithm.

The ***second goal*** of this PhD is to extend this method for the solution of a large (from a physical point of view) number of electromagnetic problems containing whether inhomogeneous or anisotropic objects. The main challenges that we are facing, will be considered here from a mathematical, physical and numerical viewpoints.

The application area of fast and accurate forward solvers is very large. At the beginning of this introduction we already discussed some of them, nevertheless there is one domain on which we would like to pay most attention. Indeed, a fast forward simulator is a key-point in every *Electromagnetic inverse scattering problems*. Microwave imaging techniques have become an important issue with a large number of applications, such as non-destructive testing, biomedical imaging and geophysical exploration. The term electromagnetic inverse scattering refers to techniques and processes used to *determine the properties of unknown targets* (shape, position, permittivity and permeability) from the knowledge of their response to an exterior electromagnetic excitation. This problem is known to be non-linear and ill-posed. One can distinguish two classes of inverse problems.

In the first class, inverse scattering problems can be formulated as linear or linearized problems under restrictive conditions, such as the Born or Rytov approximations, or if one is only interested by the induced currents characterisation. This class of inverse problems is called to be *qualitative*. Various methods exist in order to give some idea of the shape, location and number of objects inside an investigation domain. Among the different techniques yielding the shape of the scatterer, one can cite the diffraction tomography algorithm if the measurement is performed in far-field [56–59] or the back-propagation algorithm [60]. Several sampling methods have also been derived, such as the multiple signal classification (MUSIC) method [61, 62], the linear sampling method (LSM) [62–64] or the factorization method (FM) [65]. Finally, if one is only interested in the localization of point-like scatterers, the DORT (Décomposition de l'Opérateur de Retournement Temporel) method is particularly well suited [66–68]. The qualitative imaging does not aim at providing quantitative information. But nevertheless, it is an indefeasible part of solving every inverse problem. Indeed, we can use a solution of the linearized problem in order to initialise and reduce the computational burden of the *quantitative* inverse scattering problem, which belongs to the second class of inverse problems.

The quantitative inverse scattering problems provide a detailed information about all the relative parameters that determine the interaction of objects with electromagnetic fields. In other words, this class of problems aims at providing not only the aforementioned support of the objects, but they also provide the values of its electromagnetic material parameters. Unlike qualitative methods, quantitative ones solve the exact non-linear electromagnetic scattering inverse problem, which therefore requires the solution of a system of coupled equations. For this aim, an iterative optimisation procedure is generally applied. In the frame of this procedure, a cost function depending on the permittivity distribution is defined and iteratively minimized. We would like to report the existence of two general approaches here. In the first approach, called the contrast source inversion method (CSI), one of the unknowns is not the electrical field, but the current distribution. This method has been extended for various problems [69–72]. In this doctoral thesis we focus on the second, so-called "conventional" approach, where the electrical field comes into play in order to satisfy the Helmholtz equation at each iteration. One can note various Newton-type schemes [73–77] and a few global optimisation techniques [78, 79].

Thus, the ***third goal*** of this work consists in taking advantages of the proposed FETI-DPEM2-full method into a quantitative inversion algorithm investigated in our laboratory. Indeed, it is of great interest to render the inversion process more flexible in terms of required memory. This tool will give us an opportunity to treat problems on a larger scale. We will discuss as well different effective numerical strategies while implementing the proposed method in the inversion scheme. In order to validate the new approach from a practical point of view, we will test it against experimental data acquired with real world targets from the Fresnel database [80, 81].

The organisation of this thesis is as follows. In Chapter (2) the basic concepts and equations of the electromagnetic theory are briefly reviewed and introduced in order to position the problem. Chapter (3) introduces a new variation of the FETI-DPEM2 method, with improved transmission conditions between subdomains. We put some emphasis on the explanation of the difference between the various DDMs and discuss these methods based on the use of the Lagrangian theory. We also detail the specificities of the proposed FETI-DPEM2-full method. Chapter (4) is devoted to test this FETI-DPEM2-full method for various two-dimensional configurations and to show its efficiency with respect to the classical FETI-DPEM2 method. In Chapter (5) we discuss the difficulties that we faced while solving the Interface equation related to the FETI-DPEM2-full method with both direct and iterative methods in the case of 3D configurations. In this chapter, we mostly focus our attention on the convergence behaviour of the iterative algorithm and we show that it is strongly affected by the presence of anisotropic materials. After discussing the efficiency of the FETI-DPEM2-full method, we compare the results of simulations with Scattered field measurements. Finally, Chapter (6) explores the inversion part of this PhD thesis, i.e. the inversion algorithm that we use where an efficient implementation of the FETI-DPEM2-full method has been performed. In this chapter, reconstructions of various targets from measured scattered fields, extracted from the Fresnel database, are performed in order to validate the entire inversion algorithm and show the effectiveness of the proposed methodology.

Chapitre 2

Mathematical formulation

Contents

2.1 Introduction

Electromagnetic analysis has been an indispensable part of many engineering and scientific studies during the past century. Its success would not have been possible except for the existence of an accurate and complete theory, which J.C. Maxwell finalized in 1873 [1]. The problem of electromagnetic analysis is actually a problem of solving a set of Maxwell's equations subject to given boundary conditions. As it was mentioned in the general introduction, in this work we deal with the time-harmonic regime of Maxwell's equations [82] directly in terms of the electric and magnetic fields. For this, it is necessary to derive from Maxwell's equations, which involve both electric and magnetic fields, the governing differential equations involving only one of them [16]. We will focus on the electric field which can be found as a solution of the Helmholtz equation.

In this chapter, we introduce briefly the basic concepts and equations of the electromagnetic theory used in this PhD thesis, in particular the different types of electromagnetic fields [83] and its formulations in two- and three-dimensional cases [84], as well as the different types of associated boundary conditions from a mathematical and physical point of view [85]. Finally, we introduce some auxiliary techniques in the electromagnetic theory, such as the Near-to-Far-field transformation based on the Huygen's principle [86] and the Perfectly Matched Layer [87].

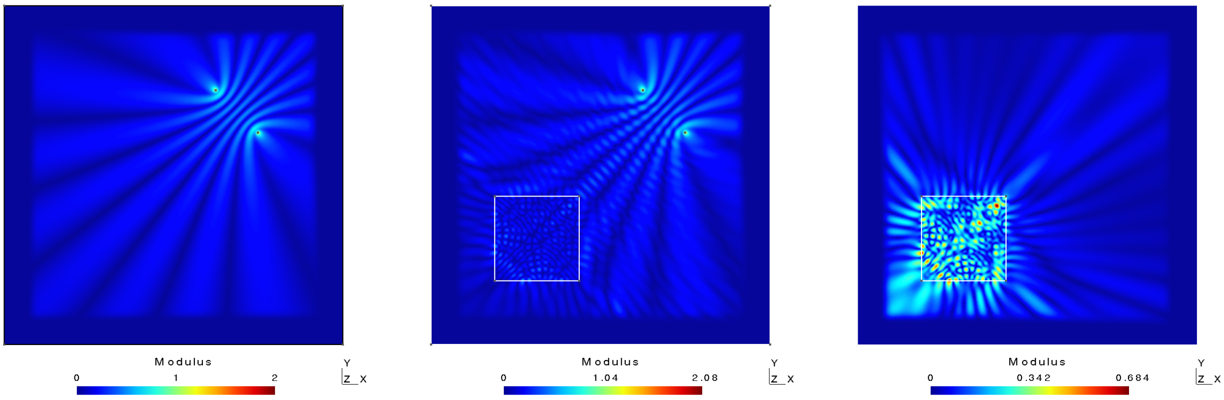
For a complete presentation of electromagnetic theory, the reader is encouraged to study references [16,82,83,88,89] for an in-depth discussion of this material. This chapter may be skipped if the reader is familiar with the theory.

2.2 2D electromagnetic problems

In order to explain the different types of electromagnetic fields, we will start with a simple example. Imagine that there are some sources that radiate in a bounded domain Ω filled with free space, i.e. the space whose relative permittivity and permeability are equal to 1. The field that corresponds to this configuration is called the *Incident* field $\mathcal{E}^{\text{inc}}$ (Figure 2.1a). Imagine right now, that we want to add to the domain Ω some objects with a relative permittivity which is different from 1. In this case we obtain the *Total* field $\mathcal{E}^{\text{tot}}$ (Figure 2.1b). From the linearity of Maxwell's equations, the *Scattered* field can be calculated as a subtraction of the incident field $\mathcal{E}^{\text{inc}}$ from the total field $\mathcal{E}^{\text{tot}}$ (Figure 2.1c) as follows

$$\mathcal{E}^{\text{sc}} = \mathcal{E}^{\text{tot}} - \mathcal{E}^{\text{inc}} \tag{2.1}$$

In this work we distinguish two main ways for obtaining the scattered field. In the next subsections we will discuss these cases.



(a) Incident field (b) Total field (c) Scattered field

FIGURE 2.1 – Modulus maps ($20 \cdot log_{10}|\mathcal{E}|$) of the different types of electric fields obtained in a 2D domain of size $21 \times 21\lambda^2$.

2.2.1 Incident, total and scattered field formulations

We consider the electromagnetic scattering problem where an incident electromagnetic wave impinges on an inhomogeneous medium. Let us assume that we are in a 2D configuration and that the non-null component of the electric field is along the invariance axis, the z-axis. The term $\mathcal{E}^{\text{inc}} = \mathcal{E}_z^{\text{inc}}(\vec{r})$ in this 2D configuration corresponds to the z-component of the incident field, and can be found as a solution of the following 2D Helmholtz equation in a bounded domain Ω, when the s polarisation case is considered, with a time convention in $\exp(-j\omega t)$:

$$-\operatorname{div}\left(\frac{1}{\mu_r^{\text{inc}}}\operatorname{grad}\mathcal{E}^{\text{inc}}\right) - k_0^2\varepsilon_r^{\text{inc}}\mathcal{E}^{\text{inc}} = jk_0Z_0J, \text{ in } \Omega \tag{2.2}$$

where k_0 and Z_0 are the wavenumber and the intrinsic impedance respectively of free space, the quantity μ_r^{inc} (resp. $\varepsilon_r^{\text{inc}}$) corresponds to the relative permeability (resp. permittivity) of the medium when it is isotropic. If the media are anisotropic, permeability tensors $\bar{\bar{\mu}}_r^{\text{inc}}(\vec{r})$ and permittivity tensors $\bar{\bar{\varepsilon}}_r^{\text{inc}}(\vec{r})$ also come into play. $J = J_z(\vec{r})$ is a given current distribution which produces the field.

In the presence of inhomogeneities, the field is perturbed and is now denoted as the total field $\mathcal{E}^{\text{tot}}$, which satisfies a similar Helmholtz equation

$$-\operatorname{div}\left(\frac{1}{\mu_r^{\text{tot}}}\operatorname{grad}\mathcal{E}^{\text{tot}}\right)-k_0^2\varepsilon_r^{\text{tot}}\mathcal{E}^{\text{tot}}=jk_0Z_0J,\ \text{in }\Omega \tag{2.3}$$

where the relative tensors of permeability $\bar{\bar{\mu}}_r^{\text{tot}}(\vec{r})$ and permittivity $\bar{\bar{\varepsilon}}_r^{\text{tot}}(\vec{r})$ only differ from $\bar{\bar{\mu}}_r^{\text{inc}}(\vec{r})$ and $\bar{\bar{\varepsilon}}_r^{\text{inc}}(\vec{r})$ where the inhomogeneities are located.

In the framework of this PhD thesis we will consider two main possibilities for finding the scattered field. To start with, according to Eq. (2.1), we represent the scattered field as $\mathcal{E}^{\text{sc}}=\mathcal{E}^{\text{tot}}-\mathcal{E}^{\text{inc}}$. That means that we can find the scattered field $\mathcal{E}^{\text{sc}}$ in three steps

1/ At first, by calculating numerically the total field $\mathcal{E}^{\text{tot}}$ solving Eq. (2.3)
2/ Then, by calculating the incident field $\mathcal{E}^{\text{inc}}$ solving Eq. (2.2)
3/ And, finally, by subtracting one from another

According to Eq. (2.1), we can also present the total field as a summation of the scattered and incident fields. Then, by substituting this relation into Eq. (2.3) we obtain Eq. (2.4)

$$-\operatorname{div}\left(\frac{1}{\mu_r^{\text{tot}}}\operatorname{grad}\ \left[\mathcal{E}^{\text{inc}}+\mathcal{E}^{\text{sc}}\right]\right)-k_0^2\varepsilon_r^{\text{tot}}\left[\mathcal{E}^{\text{inc}}+\mathcal{E}^{\text{sc}}\right]=jk_0Z_0J,\ \text{in }\Omega \tag{2.4}$$

We now subtract Eq. (2.2) from Eq. (2.4) and obtain Eq. (2.5) for the scattered field.

$$-\operatorname{div}\left(\frac{1}{\mu_r^{\text{tot}}}\operatorname{grad}\mathcal{E}^{\text{sc}}\right)-k_0^2\varepsilon_r^{\text{tot}}\mathcal{E}^{\text{sc}}=J^{\text{sc}},\ \text{in }\Omega \tag{2.5}$$

where the induced currents are defined as

$$J^{\text{sc}}=\operatorname{div}\left(\left[\frac{1}{\mu_r^{\text{tot}}}-\frac{1}{\mu_r^{\text{inc}}}\right]\operatorname{grad}\mathcal{E}^{\text{inc}}\right)+k_0^2\left[\varepsilon_r^{\text{tot}}-\varepsilon_r^{\text{inc}}\right]\mathcal{E}^{\text{inc}} \tag{2.6}$$

As one can see, the right part of this equation J^{SC} serves as secondary sources which differ from 0 if $\frac{1}{\mu_r^{\text{tot}}}\neq\frac{1}{\mu_r^{\text{inc}}}$ or $\varepsilon_r^{\text{tot}}\neq\varepsilon_r^{\text{inc}}$ and when the incident field is non null.

The incident field $\mathcal{E}^{\text{inc}}$ involved in Eq. (2.6) can be computed either by solving Eq. (2.2) with a help of a numerical method, or by using the exact solution if it can be calculated analytically These two approaches for calculating the scattered field are complementary and will be both exploited in the following.

2.2.2 Boundary conditions

While there are many functions that satisfy the differential equations given above in the domain of interest, only one of them is the real solution to the problem. To determine this solution, one must know the *boundary conditions* associated with the considered domain. In other words, a complete description of an electromagnetic problem should include the information about both differential equations and boundary conditions.

In this subsection we present some boundary conditions in order to guarantee the unicity of the solution in the domain Ω.

2.2.2.1 Mathematical formulation

Generally speaking, in electromagnetism, we deal with three main types of boundary conditions from a mathematical point of view :

1. Dirichlet-type boundary condition.
   We will start with the Dirichlet-type boundary condition across the surface Σ_1, which in the general form is
   $$\mathcal{E} = \theta_1, \text{ on } \Sigma_1 \tag{2.7}$$
2. Neumann-type boundary condition.
   Now we will consider the Neumann-type boundary conditions across the surface Σ_2, which in the general form is presented as
   $$\frac{1}{\mu_r}\frac{\partial \mathcal{E}}{\partial n} = \theta_2, \text{ on } \Sigma_2 \tag{2.8}$$
3. Robin-type boundary condition.
   The third type used in this work is the Robin-type boundary conditions across the surface Σ_3, which in the general form is presented as
   $$\frac{1}{\mu_r}\frac{\partial \mathcal{E}}{\partial n} + \alpha\mathcal{E} = \theta_3, \text{ on } \Sigma_3 \tag{2.9}$$

2.2.2.2 Physical sense

1. Dirichlet-type boundary condition.
   From an electromagnetic viewpoint, metal is characterized as having a very high conductivity. So in the microwave regime, it can be assumed as a perfect conductor [13]. It means that at the interface of this type of material, the tangential component of the electric field is null. In the specific case of polarization considered here, it corresponds indeed to $\mathcal{E} = 0$ [85,90].
2. Neumann-type boundary condition.
   The Neumann-type boundary condition occurs whenever the normal component of the electric field intensity is zero. Physically it corresponds to a perfect magnetic conductor in the considered polarization case, or can be found when there is a plane of symmetry for the electric field [90].
3. Robin-type boundary condition.
   From a physical point of view, the Robin-type boundary condition corresponds to the knowledge of the relation between some specified potentials and the intensity of the electric field on the specific boundary. As an example, we can refer to a case where the outer boundary of the domain recedes to infinity. Such domain is called unbounded or open. A condition must be specified at this infinite boundary to obtain an unique solution for the problem. Such a condition is referred to as a Sommerfeld radiation boundary condition [91]. Assuming that all the sources and objects are immersed in free space and located within a finite distance from the origin of the coordinate system, the electric field $\mathcal{E}(\vec{r})$ is required to satisfy
   $$\lim_{\rho\to\infty}\sqrt{\rho}\left[\frac{\partial}{\partial\rho}\mathcal{E}(\rho) - jk_0\mathcal{E}(\rho)\right] = 0 \tag{2.10}$$
   where $\rho = \sqrt{x^2+y^2}$.
   This equation is exactly valid at infinity. But as we reduce the size of the computational domain to a finite one, we have to replace Eq. (2.10) by its finite approximation at an outer boundary Σ_{out}. Thus, we present the *first*-order approximation of the radiation boundary condition which is a representative of the Robin-type boundary conditions (Eq. 2.11).
   $$\frac{1}{\mu_r}\frac{\partial \mathcal{E}(\vec{r})}{\partial n} - jk_0\mathcal{E}(\vec{r}) = 0, \text{ on } \Sigma_{out} \tag{2.11}$$
   It can be noticed here that this kind of boundary conditions can also be found in the case of a high conductivity approximation.

2.3 3D electromagnetic problems

Similarly, a 3D scattered field formulation of the problem can be derived from the incident and total field formulations. We consider a three-dimensional electromagnetic scattering problem, where a known monochromatic incident electromagnetic wave is impinging on an inhomogeneous target $D \subset \Omega$, whose relative permittivity varies with respect to the surrounding. The relative permittivity and permeability distribution in absence (resp. in presence) of the scatterer are denoted as previously by $\bar{\bar{\varepsilon}}_r^{\text{inc}}(\vec{r})$ and $\bar{\bar{\mu}}_r^{\text{inc}}(\vec{r})$ (resp. $\bar{\bar{\varepsilon}}_r^{\text{tot}}(\vec{r})$ and $\bar{\bar{\mu}}_r^{\text{tot}}(\vec{r})$) with

$$\begin{array}{ll} \bar{\bar{\varepsilon}}_r^{\text{tot}}(\vec{r}) \neq \bar{\bar{\varepsilon}}_r^{\text{inc}}(\vec{r}) & \quad \forall \vec{r} \in D \\ \bar{\bar{\varepsilon}}_r^{\text{tot}}(\vec{r}) = \bar{\bar{\varepsilon}}_r^{\text{inc}}(\vec{r}) & \quad \forall \vec{r} \in \Omega \setminus D \end{array}$$

Unlike the 2D case, for the 3D formulation, we have to consider the 3D vectorial Helmholtz equation for the incident or the total field :

$$\nabla \times \left(\frac{1}{\bar{\bar{\mu}}_r} \nabla \times \vec{\mathcal{E}} \right) - k_0^2 \bar{\bar{\varepsilon}}_r \vec{\mathcal{E}} = j k_0 Z_0 \vec{J} \text{ in } \Omega \tag{2.12}$$

Depending on the value of relative permittivity and permeability, the solution $\vec{\mathcal{E}} = (\mathcal{E}_x(\vec{r}), \mathcal{E}_y(\vec{r}), \mathcal{E}_z(\vec{r}))^T$ can correspond to the incident or total field.

From the linearity of Maxwell's equations, the total field (associated to the permittivity distribution $\bar{\bar{\varepsilon}}_r^{\text{tot}}$) can be decomposed into the incident field $\vec{\mathcal{E}}^{\text{inc}}$ (associated to $\bar{\bar{\varepsilon}}_r^{\text{inc}}$) and the scattered field $\mathcal{E}^{\text{sc}}$ which satisfies a Helmholtz equation

$$\nabla \times \left(\frac{1}{\bar{\bar{\mu}}_r^{\text{tot}}} \nabla \times \vec{\mathcal{E}}^{\text{sc}} \right) - k_0^2 \bar{\bar{\varepsilon}}_r^{\text{tot}} \vec{\mathcal{E}}^{\text{sc}} = \vec{J}^{\text{sc}} \text{ in } \Omega \tag{2.13}$$

where k_0 is the vacuum wavenumber. The induced currents are defined as

$$\vec{J}^{\text{sc}} = -\nabla \times \left(\left[\frac{1}{\bar{\bar{\mu}}_r^{\text{tot}}} - \frac{1}{\bar{\bar{\mu}}_r^{\text{inc}}} \right] \nabla \times \vec{\mathcal{E}}^{\text{inc}} \right) + k_0^2 \left[\bar{\bar{\varepsilon}}_r^{\text{tot}} - \bar{\bar{\varepsilon}}_r^{\text{inc}} \right] \vec{\mathcal{E}}^{\text{inc}} \tag{2.14}$$

2.3.1 Finite domain and associated boundary conditions

As for the 2D case (Section 2.2.2), we need to look for the correct solution among all the functions which satisfy the differential equation (2.13). To determine this solution, one must know the boundary conditions associated with the domain. The idea of the boundary conditions is the same as for the 2D case (Section 2.2.2). There is only a difference in the equations.

For the 3D case we consider :

— Dirichlet-type boundary conditions (Eq. 2.15)

$$\vec{n} \times \vec{n} \times \vec{\mathcal{E}} = \theta_1, \text{ on } \Sigma_1 \tag{2.15}$$

— Neumann-type boundary conditions (Eq. 2.16)

$$\vec{n} \times \frac{1}{\mu_r} \nabla \times \vec{\mathcal{E}} = \theta_2, \text{ on } \Sigma_2 \tag{2.16}$$

— Robin-type boundary conditions (Eq. 2.17) given as :

$$\vec{n} \times \left(\frac{1}{\mu_r} \nabla \times \vec{\mathcal{E}} \right) + \alpha\, \vec{n} \times \vec{n} \times \vec{\mathcal{E}} = \theta_3 \text{ on } \Sigma_3 \tag{2.17}$$

We can notice here that the physical meaning of each type of boundary conditions is the same as the one encountered in Section (2.2.2) for the 2D case.

2.4 Near-field/Far-field transform

The sources and the receivers are not necessarily located in the vicinity of the target. When the sources are far from the target, one way to compute the field scattered by the object in its vicinity is to use the scattered field formulation. When the receivers are located far from the object it is not possible to mesh the whole space containing both the scatterers and the receivers. Thus, in this work, we will use the Near-to-Far-Field transformation based on Huygen's principle. The main idea of this approach [16] is that if we know the electric and magnetic fields $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ on the surface Σ enclosing all the scatterers in free space, we are able to calculate the field everything outside Σ thanks to the following equation :

$$\begin{aligned}\vec{\mathcal{E}}(r) = \oiint_{\Sigma} \{ -jw\bar{\bar{\mu}} \left[\hat{n}' \times \vec{\mathcal{H}}(r') \right] G_0(r,r') + \left[\hat{n}' \cdot \vec{\mathcal{E}}(r') \right] \nabla' G_0(r,r') + \\ + \left[\hat{n}' \times \vec{\mathcal{E}}(r') \right] \times \nabla' G_0(r,r') \} \, dS\end{aligned} \tag{2.18}$$

where $\vec{\mathcal{H}}(r')$ is the Magnetic field at r', $\hat{n}'$ denotes the unit vector normal to Σ at r' and pointing toward the exterior region. $G_0(\mathrm{r,r}')$ is the 3D free-space scalar Green's function, given by

$$G_0(\mathrm{r,r}') = \frac{e^{jk_0|r-r'|}}{4\pi|r-r'|} \tag{2.19}$$

2.5 Perfectly Matched Layer

Instead of using a Robin-type boundary condition in order to obtain the solution of the electromagnetic problem in an unbounded domain, we can use a *perfectly matched layer* (PML). The concept of PML was proposed by Berenger in 1994 to alleviate some problems with fictitious absorbers [87]. A perfectly matched layer is a medium which is added and bounds the computational domain, it is built such as all impinging waves emanating from Ω are transmitted and damped in the PML without any reflections whatever the angle or the polarisation are. In order to ensure numerically that the domain is physically unbounded, we will surround the external boundary of Ω with a cartesian PML.

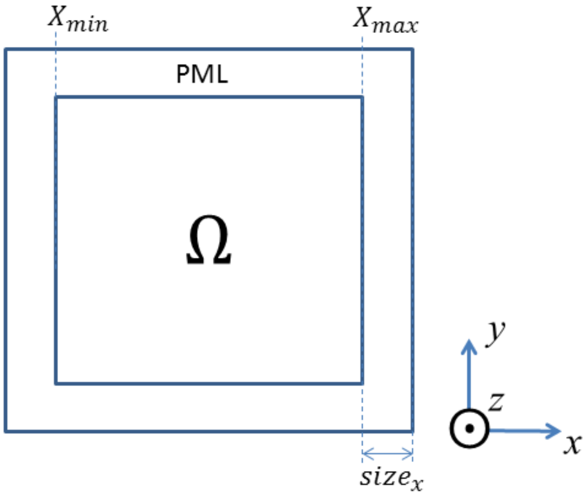
There are various ways to apply the PML. In particular we can introduce a PML by stretching the coordinate system [92]. But in this work we associate it with some kind of anisotropic material. The anisotropic absorber model of PML was first derived by Sacks et al. [93]. Following this idea, we present the coefficients of the relative permeability and permittivity associated to PML (resp. μ_{PML} and $\varepsilon_{\mathrm{PML}}$) as a matrix-multiplication of tensors in the following form :

$$\bar{\bar{\mu}}_{\mathrm{PML}} = \bar{\bar{\mu}}_r \bar{\bar{\Lambda}} \quad \text{and} \quad \bar{\bar{\varepsilon}}_{\mathrm{PML}} = \bar{\bar{\varepsilon}}_r \bar{\bar{\Lambda}} \tag{2.20}$$

where the tensors $\bar{\bar{\mu}}_r$ and $\bar{\bar{\varepsilon}}_r$ are the relative permeability and permittivity arising in Eqs. (2.5) and (2.13). In the Cartesian coordinate system, the new tensor is defined as follows :

$$\bar{\bar{\Lambda}} = \begin{bmatrix} s_y s_z/s_x & 0 & 0 \\ 0 & s_z s_x/s_y & 0 \\ 0 & 0 & s_x s_y/s_z \end{bmatrix} \tag{2.21}$$

where the coefficient s_x is calculated according to the rule which is presented in Figure (2.2).



```
if ($x > X_{max}$)
    $d = |x - X_{max}|/size_x$
    $\chi_x = 1 + \chi\ d^n$
    $\sigma_x = \frac{\sigma}{\omega\varepsilon_0}\ d^n$
    $s_x = \chi_x + j\ \sigma_x$
elseif ($x < X_{min}$)
        $d = |x - X_{min}|/size_x$
        $\chi_x = 1 + \chi\ d^n$
        $\sigma_x = \frac{\sigma}{\omega\varepsilon_0}\ d^n$
        $s_x = \chi_x + j\ \sigma_x$
    else
        $s_x = 1$
    end if
end if
```

(a) Investigation domain Ω with PML

(b) PML computation

FIGURE 2.2 – (a) Schematic map of a domain Ω containing PML area with (b) the algorithm of the PML computation, investigated in this work.

x is the current value of the barycentre of the considered element, X_{max} and X_{min} are the limits of the PML boundaries, $size_x$ is the size of the PML area in the x direction, σ is the relative conductivity, ω is the angular frequency, $\varepsilon_0 = 8.85 \cdot 10^{-12}\mathrm{F} \cdot \mathrm{m}^{-1}$, χ is a parameter which is equal to 1 in our work, and finally, n - is a numerical parameter which controls at which speed the wave is attenuated. In our work it is equal to 3.

The other two coefficients (s_y and s_z) are determined in a similar way.

2.6 Conclusion

In this chapter we introduced briefly the basic concepts and equations of the electromagnetic theory used in this PhD thesis. We have considered in particular the different types of electromagnetic fields and their formulations in two- and three-dimensional cases. Then we discussed the different types of the associated boundary conditions from a mathematical and physical point of view. As a next step we are going to solve the presented equations in a numerical way, with methods which are going to be introduced in the next chapter.

Chapitre 3

Numerical methods

Contents

3.1 Introduction

Unfortunately, Maxwell's equations and, in particular, Helmholtz equations for can only be solved analytically for a very few primary cases, such as mono-chromatic homogeneous problems without anisotropy. Whereas a variety of approximate analytical techniques have been developed for relatively simple problems in terms of geometry and physics, accurate and complete analysis of complex cases, especially inhomogeneous configurations containing anisotropy, can be accomplished only through a numerical method that solves differential equations with the help of powerful computers.

In particular, the Finite-Element Method (FEM) [94] is a numerical technique used to obtain approximate solutions to boundary-value problems and is applied in computational electromagnetism for treating mostly complex geometry and physics. The system of linear equations resulting from a FEM discretization is highly sparse and can be solved using efficient solution techniques for sparse matrices based on either direct methods [23, 95], or iterative methods [22, 23]. Direct methods have an advantage that multiple right-hand sides can be treated efficiently with an excellent precision. However, storing the factorized matrix is very challenging in terms of memory for large-scale problems. Iterative methods are much less memory expensive. Nevertheless, solving the indefinite Helmholtz equation with an iterative method is a difficult task in terms of convergence process [21] and it is not so easy to find a good preconditioner for the problem matrix [21, 23].

Nevertheless, however powerful is the employed linear solvers, finite element methods face major difficulties when the size of the computational domain is large with respect to the wavelength. Indeed, memory storage and computational time increase drastically with the size of the mesh. It is thus compulsory to find new ways for solving these large scale problems.

Over past decade, the Domain Decomposition (DD) technique has been recognized as one of the most important methodologies for constructing efficient parallel computing. Among various Domain Decomposition Methods (DDMs), the Finite Element Tearing and Interconnecting (FETI) method shows great potential to improve the capability of the FEM. The FETI method distinguishes itself through the partitioning of the entire computational domain into non-overlapping subdomains and the construction of the interface problem between them, which can be solved iteratively. Recently, the ElectroMagnetic Dual-Primal FETI method, called FETI-DPEM [41, 43], has been developed for the simulation of three-dimensional electromagnetic problems, with the use of a global preconditioner smartly designed in [39] to significantly improve the convergence of the interface solution. Over past few years, the serial implementation of the FETI-DPEM and other FETI-like methods have been successfully applied to the 2D and 3D electromagnetic problems [43, 45, 55, 96, 97].

In this Chapter we aim ourselves to introduce a new variation of the FETI-DPEM2 method, so-called FETI-DPEM2-full method, with improved transmission conditions between subdomains. Starting from the classical FEM formulation we will introduce the Domain Decomposition principal idea and, then, review briefly the different DD techniques. We will present methods based on the use of the Lagrangian theory and we will explain the main difference between the classical FETI-DPEM2 method and its modification which is going to be proposed. Finally, we will discuss some implementation issues of the proposed method.

3.2 Finite Element Method in 2D

We are interested in solving the following 2D Helmholtz equation in the case of s polarization in a bounded domain Ω (Figure 3.1) with the boundary $\Sigma = \partial\Omega$ which can contains inhomogeneities :

$$-\operatorname{div}\left(\frac{1}{\mu_r}\operatorname{grad}\mathcal{E}\right) - k_0^2\varepsilon_r\mathcal{E} = jk_0Z_0J, \text{ in } \Omega \tag{3.1}$$

where the solution $\mathcal{E}$ can be, for example, the z-component of the total or incident field.

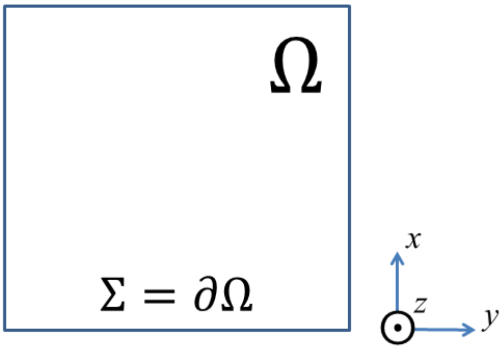


Figure 3.1 – Schematic presentation of the domain Ω in 2D case.

Using the Galerkin-approach [98] the residual of Eq. (3.1) must be orthogonal (in the sense of the $L^2(\Omega)$ space) to the space of the trial functions which we call Φ

$$\int\limits_{\Omega}\left(-\operatorname{div}\left(\frac{1}{\mu_r}\operatorname{grad}\mathcal{E}\right) - k_0^2\varepsilon_r\mathcal{E} - jk_0Z_0J\right) v \, d\Omega = 0, \qquad \forall v \in \Phi \tag{3.2}$$

Next, we convert the first part of Eq. (3.2) using Green's formula

$$\int\limits_{\Omega}\frac{1}{\mu_r}\operatorname{grad}\mathcal{E}\operatorname{grad}v \, d\Omega - \int\limits_{\Sigma}\frac{1}{\mu_r}\frac{\partial\mathcal{E}}{\partial n} v \, dS - \int\limits_{\Omega}k_0^2\varepsilon_r\mathcal{E} \, v \, d\Omega = -\int\limits_{\Omega}jk_0Z_0J \, v \, d\Omega \tag{3.3}$$

As we have full freedom for choosing the space of trial functions, we can in particular select $H_\theta^1(\Omega)$ as Φ. The functions of $H_\theta^1(\Omega)$ are the functions from $L^2(\Omega)$, which equal to θ on the boundary Σ_θ. To solve the problem (3.1) using the Finite Element Method, the domain Ω is subdivided into a number of small domains (such as triangles for example), denoted as $\Omega^s (s = 1, 2, 3, ..., M)$, with M denoting the total number of triangles. To approximate Eq. (3.3) we substitute the infinite space of solution $H_\theta^1(\Omega)$ by its finite-dimensional subspace $V^h(\Omega)$ with the set of basis functions $\{\psi_s(\vec{r})\}$. In the present work, we will limit ourselves to first-order linear basis functions. In this case we can find a function $\mathcal{E}_h$ close to $\mathcal{E}$, using trial functions $v_h(\vec{r})$. We will also approximate the infinite function J by its finite discretized analogue J_h. Taking this into account, we can rewrite Eq. (3.3) as follows :

$$\int\limits_{\Omega}\frac{1}{\mu_r}\operatorname{grad}\mathcal{E}_h\operatorname{grad}v_h \, d\Omega - \int\limits_{\Sigma}\frac{1}{\mu_r}\frac{\partial\mathcal{E}_h}{\partial n} v_h \, dS - \int\limits_{\Omega}k_0^2\varepsilon_r\mathcal{E}_h \, v_h \, d\Omega = -\int\limits_{\Omega}jk_0Z_0J_h \, v_h \, d\Omega \tag{3.4}$$

Every function $\mathcal{E}_h \in V^h(\Omega)$ can be presented as a linear combination

$$\mathcal{E}_h(\vec{r}) = \sum_{j=1}^{n_j} q_j\psi_j(\vec{r}) \tag{3.5}$$

We can rewrite (Eq. 3.4) into a system of linear equations for the vector components q_j in the discretized domain Ω :

$$\begin{aligned}\sum_{j=1}^{n_j}\left(\int\limits_{\Omega}\frac{1}{\mu_r}\mathrm{grad}\,\psi_j(\vec{r})\cdot\mathrm{grad}\,\psi_i(\vec{r})\,d\Omega-\int\limits_{\Omega}k_0^2\varepsilon_r\ \psi_j(\vec{r})\ \psi_i(\vec{r})\ d\Omega\right)q_j=\\ -\int\limits_{\Omega}jk_0Z_0J_h\,\psi_i(\vec{r})\ d\Omega+\int\limits_{\Sigma}\frac{1}{\mu_r}\frac{\partial\mathcal{E}_h}{\partial n}\ \psi_i(\vec{r})\ dS,\qquad i=1,...,n_i\end{aligned}\tag{3.6}$$

This equation is written for every basis function ψ_i $(i=1,...,n_i)$ of the linear combination of trial function v_h. We then rewrite Eq. (3.6) in a matrix form :

$$\mathbf{KE}=\mathbf{f}+\int\limits_{\Sigma}\frac{1}{\mu_r}\frac{\partial\mathcal{E}_h}{\partial n}\ \Psi\ dS\tag{3.7}$$

Where, the FEM system matrix **K** is a summation of the Stiffness and Mass matrix, **f** is a right hand side (rhs) vector. Note that the terms **E** and $\mathcal{E}_h$ are *different* [41, 43]. The former is the discretized finite vector of *weights* $\mathbf{E}=(q_1,...,q_{n_j})^T$, while the latter represents the discretized analogue of the electric field $\mathcal{E}$. They are related through Eq. (3.5).

The FEM system matrix and the rhs vector are determined as follows :

$$\mathbf{K}\ =\ \mathbf{G}-\mathbf{M}_{vol}\tag{3.8}$$

$$\mathbf{G}\ =\ \int\limits_{\Omega}\frac{1}{\mu_r}\ \nabla\Psi\cdot(\nabla\Psi)^T\ d\Omega\tag{3.9}$$

$$\mathbf{M}_{vol}\ =\ \int\limits_{\Omega}k_0^2\varepsilon_r\ \Psi\cdot\Psi^T\ d\Omega\tag{3.10}$$

$$\mathbf{f}\ =\ -\int\limits_{\Omega}jk_0Z_0\ J_h\cdot\Psi\ d\Omega\tag{3.11}$$

The term Ψ denotes a column vector containing the first-order linear basis functions ψ_i, $i=1,...,n_i$. The number of unknowns in this case corresponds to the number of the points in the discretized domain Ω. More precisely, the reader is referred to [99] in order to complete its knowledge on this type of basis functions.

3.2.1 Imposition of the boundary conditions

In order to make the problem (Eq. 3.7) well posed, we need at first to impose boundary conditions on the boundary Σ (Figure 3.1). In other words, we need to determine the last term

$$\int\limits_{\Sigma}\frac{1}{\mu_r}\frac{\partial\mathcal{E}_h}{\partial n}\ \Psi\ dS\tag{3.12}$$

which is related to the external boundary Σ. In this subsection we will talk about the implementation of the different types of boundary conditions presented in Sections (2.2.2 and 2.3.1).

1. Dirichlet-type boundary condition.
   The term (3.12) can not be defined exactly from the Dirichlet boundary conditions. But if we choose a special type of basis functions which are equal to θ_1 on the boundary Σ_1 (see Eqs. 2.7 and 2.15), it will automatically cancel out this term.

2. Neumann-type boundary condition.
   The application of the Neumann-type boundary conditions is also simple. When we substitute Eq. (2.8) into Eq. (3.12) we obtain the term $\int\limits_{\Sigma_2} \theta_2 \; \Psi \; dS$ for all the basis functions across the boundary Σ_2 which is fully-determined and can be calculated.[1]
3. Robin-type boundary condition.
   When we substitute the Robin-type boundary condition defined as Eq. (2.9) into Eq. (3.12) the resulting term can be divided into two parts :

$$\int_{\Sigma_3} \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h}{\partial n} \; \Psi \; dS = \int_{\Sigma_3} \theta_3 \; \Psi \; dS - \int_{\Sigma_3} \alpha \; \Psi \; \Psi^T \; dS \tag{3.13}$$

   The first integral is associated to the rhs vector, while the second term, the so-called Mass-matrix of the Robin-type boundary condition (Eq. 3.14), makes a contribution to the FEM system matrix $\mathbf{K}$.

$$\mathbf{M}_\Sigma(\alpha) = \int_{\Sigma_3} \alpha \; \Psi \; \Psi^T \; dS \tag{3.14}$$

3.2.2 Final linear system

We suppose that on the *external* boundary Σ the electrical field $\mathcal{E}$ satisfies the first-order approximation of the Sommerfeld radiation boundary conditions[2] written in the form of Eq. (2.11) in order to guarantee the unicity of the solution in the bounded domain Ω. Taking into account the Robin-type boundary condition as previously described in Section (3.2.1), Eq. (3.7) can be rewritten as follows :

$$\tilde{\mathbf{K}}\mathbf{E} = \mathbf{f} \tag{3.15}$$

where :

$$\tilde{\mathbf{K}} = \mathbf{G} - \mathbf{M}_{vol} - \mathbf{M}_\Sigma(jk_0) \tag{3.16}$$

As a rule, the FEM matrix $\tilde{\mathbf{K}}$ resulting from the finite element discretization of the computational domain even if sparse, can tend to be enormous. Especially in the domain of high-frequency electromagnetic scattering problems, where fine meshes are required for obtaining an accurate solution. For such problems, solving Eq. (3.15) with a direct method entails memory and CPU requirements that rapidly overwhelm even the largest resources that are currently available. In order to make this process less expensive in terms of memory requirement, we could also provide an iterative solution to Eq. (3.15). However, for high frequencies and therefore large values of the wavenumber k, the finite-element matrix $\tilde{\mathbf{K}}$ is usually indefinite, which poses serious challenges to the analysis, implementation and performance of iterative solvers [21].

3.3 Domain decomposition technique in 2D

Over the past decade, the domain decomposition technique has been recognized as one of the most effective methodologies for solving large-scale problems. There are many different methods which are based on the Domain Decomposition technique. In the following, we will exploit a specific version of one of these methods, named FETI. First, a tour of the various available methods is provided. The 2D case will be mainly considered, and the specificities related to the 3D configuration will be detailed later on.

---

[1] We use the same technology for the 3D case.

[2] We also would like to note that we could impose the other types of boundary conditions described in Section (2.2.2) and it would not change the principals of the FETI or DDM techniques.

3.3.1 Brief overview of DD-based methods

Let us start with the principal Domain Decomposition idea, which consists in dividing the domain Ω into a set of non-overlapping subdomains Ω^i $(i = 1, .., N_s)$ with external and internal boundaries Σ^i and Γ^i $(i = 1, .., N_s)$ respectively (Figure 3.2), in order to look for the N_s-tuple of the solutions $E = \{E^1, \dots, E^{N_s}\}$ in an independent way, instead of looking for the solution $\mathbf{E}$ in the entire domain Ω.

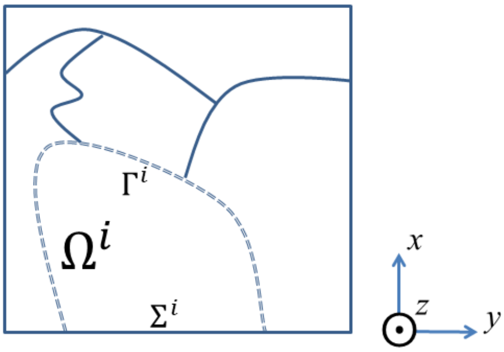


FIGURE 3.2 – Schematic map of the non-overlapping domain decomposition of the 2D domain Ω. Notations of the external and internal boundaries of the subdomain Ω^i.

Let us now focus on the given subdomain Ω^i. After applying the Galerkin approach as it has been done previously in Section (3.2) and after finite element discretization, we obtain the following system of linear equations which we will write in the matrix form for the subdomain Ω^i :

$$K^i E^i = f^i + \int\limits_{\Gamma^i} \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n} \Psi^i \, dS \tag{3.17}$$

In the above equation, the matrix K^i contains the following contributions :

$$K^i = G^i - M_{vol}^i - M_{\Sigma}^i(jk_0) \tag{3.18}$$

where the main matrices are such that

$$G^i = \int\limits_{\Omega^i} \frac{1}{\mu_r} \nabla\Psi^i \cdot (\nabla\Psi^i)^T \, d\Omega \tag{3.19}$$

$$M_{vol}^i = \int\limits_{\Omega^i} k_0^2 \varepsilon_r \; \Psi^i \cdot \Psi^{i^T} \, d\Omega \tag{3.20}$$

$$M_{\Sigma}^i(jk_0) = \int\limits_{\Sigma^i} jk_0 \; \Psi^i \cdot \Psi^{i^T} \, dS \tag{3.21}$$

$$f^i = -\int\limits_{\Omega_i} jk_0 Z_0 \; J_h^i \cdot \Psi^i \, d\Omega \tag{3.22}$$

As we can note, the last term of the system of linear equations presented in Eq. (3.17) is generally indefinite. This term represents the contribution of the unknown boundary condition imposed to the internal boundary Γ^i.

In order to "glue" the subdomains together, we build a Lagrangian formulation with the addition of new unknown Lagrange multipliers. In this work we will consider two types of methods. The first one employs a single Lagrange multiplier field to glue the local solutions at the subdomain interface boundaries. The second type of methods employs two Lagrange multiplier fields for that purpose. Physically, the Lagrange multipliers represent the continuity of the tangential component of the field (and the continuity of its derivatives). The available constraints imposed between subdomains are summarized below :

Type of continuity	Number of Lagrange Multipliers	Name of the method	References
$\mathbf{E}$	1	FETI or FETI-DP	[38, 39, 100, 101]
$\partial\mathbf{E}/\partial n$	1	Primal Schur complement	[102]
$\mathbf{E}$ & $\partial\mathbf{E}/\partial n$	1	FETI1, FETI-DP1	[36, 103]
$\mathbf{E}$ & $\partial\mathbf{E}/\partial n$	2	FETI2,FETI-DPEM2	[36, 43]
$\mathbf{E}$ & $\partial\mathbf{E}/\partial n$	-	DDM	[45, 104]
$\mathbf{E}$ & $\partial\mathbf{E}/\partial n$	-	Cement DDM	[50, 51]

Table 3.1 – Summary of the different DD-based methods and related types of constraints.

When we are using one Lagrange multiplier per interface and if we are not using a regular partition of the domains, the weak form used in each subdomain and equiped with a Neumann type boundary condition at the interfaces yields spurious solutions of the Helmholtz equation which are not physically acceptable in the global solution [36].

When we pass from the Neumann- to the Robin-type boundary condition by introducing the augmented Lagrangian, it stabilizes the solution, but there is a problem in the choice of the normals [36] which has to be done very carefully for the case of irregular arbitrary mesh partitioning. Indeed, for arbitrary mesh partitions, the specific treatment of the subdomain interfaces described in [37], which is designed for ensuring a constant sign of the regularizing matrix related to the Robin-type condition, results from the fact that the FETI method considers only one normal and therefore employs only one Lagrange multiplier per mesh partition interface. Thus, in order to regularize the subdomain problem independently of the pattern of the mesh decomposition, another approach of the FETI method based on introducing two Lagrange multipliers was proposed in [36]. Such a strategy, which can also be viewed as introducing two different normals at each interface (one on each side), in the domain decomposition literature is often referred as a three-field method [105].

All the methods presented here must thus be carefully preconditioned in order to be numerically scalable with respect to different parameters :

- Number of subdomains
- Mesh size in the subdomains
- Number of Lagrange multipliers
- ...

One way of preconditioning the interface problem is to use the corner points of the partitions and to enforce strongly (by modifying the linear system, and not by adding a Lagrange multiplier) the continuity of the field at these corner points in the resolution of the linear system related to the interface parameters [39, 41–43]. In this case, and only in this case, we are dealing with Dual-Primal methods since we are linking the set of dual variables $\{\lambda\}$ to the field at corner points. This method is the one denoted as FETI-DP, or FETI-DPEM method. Finally, it is possible to further regularize the problem by employing not only classical Lagrangian formulations but instead augmented Lagrangian formulations when computing the Lagrange parameters, as these methods are known to provide better regularized solutions (for example [106, 107]).

3.3.2 Different ordering of the linear system

Let us consider each subdomain independently. The boundary conditions as well as the specific handling of the corner points will be detailed later on. We will show that it is possible to reorganize the linear systems when one wants to grap the corner points from various subdomains by using a global numbering for localizing these specific points. The problem under consideration in this section is the matrix form (Eq. 3.17) of the discretized 2D Helmholtz equation (3.1) in the subdomain Ω^i. For each type of ordering, we will introduce as well the set of geometrical matrices required for the Domain Decomposition idea.

3.3.2.1 Ordering without global corner points.

When the corners in this subdomain are not treated in a specific way, we can group the unknown coefficients into two subcategories :

$$E^i = \begin{bmatrix} E_V^i \\ E_I^i \end{bmatrix} \tag{3.23}$$

where in the subdomain Ω^i, the notation V denotes the degrees of freedom ($DOFs$) associated with the internal nodes, while I stands for the interface nodes, corners included. In this case, we can conditionally divide the problem matrix and vector as follows :

$$K^i = \begin{bmatrix} K_{VV}^i & K_{VI}^i \\ K_{IV}^i & K_{II}^i \end{bmatrix} \text{ and } f^i = \begin{bmatrix} f_V^i \\ f_I^i \end{bmatrix} \tag{3.24}$$

3.3.2.2 Ordering with corner points.

When the corners are gathered specifically, we are splitting the unknowns in the following way :

$$E^i = \left[\begin{array}{c} E_V^i \\ E_{I_r}^i \\ \hline E_{I_c}^i \end{array}\right] = \begin{bmatrix} E_r^i \\ E_c^i \end{bmatrix}, \tag{3.25}$$

where the notations E_r^i denote all the internal and interface points belonging to the subdomain Ω^i except for the corner points which are denoted by E_c^i. In this case we get the following matrices and right-hand-side vectors in each subdomain :

$$K^i = \left[\begin{array}{cc|c} K_{VV}^i & K_{VI}^i & K_{Vc} \\ K_{IV}^i & K_{II}^i & K_{Ic} \\ \hline K_{cV}^i & K_{cI}^i & K_{cc}^i \end{array}\right] = \begin{bmatrix} K_{rr}^i & K_{rc}^i \\ K_{cr}^i & K_{cc}^i \end{bmatrix}, f^i = \left[\begin{array}{c} f_V^i \\ f_I^i \\ \hline f_c^i \end{array}\right] = \begin{bmatrix} f_r^i \\ f_c^i \end{bmatrix} \tag{3.26}$$

The index r is thus joining the indexes V, I.

3.4 One Lagrange multiplier with a classical Lagrangian

We are going to deal with FETI methods based on the introduction of an unique set of Lagrange multipliers per interface in order to enforce continuity of the tangential components of the *electric fields* on each interface between subdomains. To start with, we will focus on the methods whose boundary conditions are results of the *classical* Lagrangian formalism. In literature this type of Domain Decomposition techniques is known as one-level FETI method (dual Schur complement method) [40, 108], FETI-DP [39, 109], FETI-DPEM1 method [41], FETI-H method [36, 37], or FETI-DPH method [42]

3.4.1 An unique field continuity condition

The continuity of the tangential component of the electrical fields at the interfaces that we want to impose can be written in 2D as follows :

$$\mathcal{E}|_{\Omega_i}(\vec{r}) = \mathcal{E}|_{\Omega_j}(\vec{r}) \qquad \forall \vec{r} \in \Gamma^{ij} = \Gamma^i \cap \Gamma^j \tag{3.27}$$

In order to imply in the appropriate way the continuity of the subdomain solutions (Eq. 3.27) with the correct sign of the normal on the interface we now introduce a *signed* Boolean matrix $\mathbb{B}^i$ which extracts from a given subdomain vector a signed interface boundary component. To construct the matrix $\mathbb{B} = \{\mathbb{B}^1, ..., \mathbb{B}^{N_s}\}$, we first construct the matrix $\hat{B}$ in the following way

$$\hat{B} = \begin{pmatrix} \hat{B}^1_{\Gamma^{12}} & \hat{B}^2_{\Gamma^{12}} & 0 & \cdots \\ 0 & \hat{B}^2_{\Gamma^{23}} & \hat{B}^3_{\Gamma^{23}} & \cdots \\ \hat{B}^1_{\Gamma^{13}} & 0 & \hat{B}^3_{\Gamma^{13}} & \cdots \\ \cdots & & & \end{pmatrix}$$

Note that this matrix has been constructed for the *general* case. It is an *auxiliary* matrix aiming to construct the matrix $\mathbb{B}$. The number of lines in the matrix $\hat{B}$ corresponds to the number of the interfaces, while the number of columns is equal to the number of subdomains. The local matrices $\hat{B}^i_{\Gamma^{i_1 i_2}}$ are defined as follows :

$$[\hat{B}^i_{\Gamma^{i_1 i_2}}]_{pq} = \begin{cases} 1 & \text{if } p = q \text{ and } i = i_1 \text{ and } i_1 < i_2 \\ -1 & \text{if } p = q \text{ and } i = i_1 \text{ and } i_1 > i_2 \\ 0 & \text{otherwise} \end{cases} \tag{3.28}$$

Each $\hat{B}^i_{\Gamma^{i_1 i_2}}$ matrix will provide indications related to the Dirichlet boundary constraint on the $\Gamma^{i_1 i_2}$ interface. As all the points in the various meshes are not interface points, we extract the matrix $\mathbb{B}$ from $\hat{B}$ by removing all the lines containing only zeros elements. The number of lines in the $\mathbb{B}$ matrix will therefore correspond to the number of effective points which are actually present on all the interfaces. Using this notation, the continuity of the subdomain solutions across the subdomain interfaces can be written as

$$\mathbb{B}\ E = \sum_{i=1}^{N_s} \mathbb{B}^i\ E^i = 0 \tag{3.29}$$

3.4.2 Classical Lagrangian formalism

Solving problem (3.15) is equivalent to finding the stationary point of the following Lagrangian :

$$L(E^1, ..., E^{N_s}, \lambda) = L(\{E^i\}, \lambda) = \sum_{i=1}^{N_s} \left(\frac{1}{2} K^i E^i - f^i \right)^T E^i + \lambda^T \left(\sum_{i=1}^{N_s} \mathbb{B}^i\ E^i \right) \tag{3.30}$$

In the equation above : K^i and f^i are the discretized problem matrix and the rhs vector in the given subdomain Ω^i (Eq. 3.17), E^i is the local solution in this subdomain and, finally, λ is the Lagrange parameter which has a dimension corresponding to the number of effective points which are actually present on all the interfaces. The saddle point which minimizes $\sum_{i=1}^{N_s}(\frac{1}{2}K^iE^i - f^i)^T E^i$ and maximizes $\lambda^T \sum_{i=1}^{N_s} \mathbb{B}^i E^i$ will provide the solution of the optimisation problem we are interested in. To find this saddle point, we need to consider the following set of Karush-Kuhn-Tucker (KKT) [110] conditions :

$$\begin{cases} \nabla_\lambda L(\{E^i\}, \lambda) = 0 \\ \nabla_{E^i} L(\{E^i\}, \lambda) = 0, \quad \forall\, i = 1, ..., N_s \end{cases} \tag{3.31}$$

Using the definition of directional derivatives, for the first equation, we will get

$$\begin{aligned} < \nabla_\lambda L(E^1, ..., E^{N_s}, \lambda), u > &= \\ < \nabla_\lambda L(E, \lambda), u > &= \lim_{h\to 0} \frac{L(E, \lambda + hu) - L(E, \lambda)}{h} \\ &= \lim_{h\to 0} \frac{(\frac{1}{2}KE - f)^T E - (\frac{1}{2}KE - f)^T E + (\lambda + hu)^T \mathbb{B}E - \lambda^T \mathbb{B}E}{h} \\ &= \lim_{h\to 0} \frac{\lambda^T \mathbb{B}E + (hu)^T \mathbb{B}E - \lambda^T \mathbb{B}E}{h} = \lim_{h\to 0} \frac{h(u)^T \mathbb{B}E}{h} = \lim_{h\to 0} (\mathbb{B}E)^T u \\ &= < \mathbb{B}E, u > = < \sum_{i=1}^{N_s} \mathbb{B}^i E^i, u > \end{aligned}$$

Similarly, $\forall i = 1, .., N_s$

$$< \nabla_{E^i} L(\{E^i\}, \lambda), u > \; = \; < K^i E^i - f^i + \mathbb{B}^{i^T} \lambda, u >$$

Thus

$$\begin{cases} \nabla_{E^i} L(\{E^i\}, \lambda) &= K^i E^i - f^i + \mathbb{B}^{i^T} \lambda, \quad \forall i = 1, .., N_s \\ \nabla_\lambda L(\{E^i\}, \lambda) &= \sum_{i=1}^{N_s} \mathbb{B}^i E^i \end{cases} \tag{3.32}$$

The KKT conditions yield the saddle point :

$$\begin{cases} K^i E^i - f^i + \mathbb{B}^{i^T} \lambda &= 0 \\ \sum_{i=1}^{N_s} \mathbb{B}^i E^i &= 0 \end{cases} \tag{3.33}$$

3.4.2.1 Solving the linear system without common corner points

When we do not treat the corner points separately, this corresponds to nothing but the one-level FETI method, known also as a dual Schur complement method [38,40,108]. We start with the saddle point conditions (Eq. 3.33) which can be recasted into

$$K^i E^i = f^i - \mathbb{B}^{i^T} \lambda \tag{3.34}$$

$$\sum_{i=1}^{N_s} \mathbb{B}^i E^i = 0 \tag{3.35}$$

Substituting the value of E^i found with the first equation in the second one we obtain the interface equation, or *interface* problem for the FETI method :

$$F\lambda \;=\; d \tag{3.36}$$

where

$$\begin{aligned} F &= \sum_{i=1}^{N_s} \mathbb{B}^i \left[K^i\right]^{-1} \mathbb{B}^{i^T} && (3.37) \\ d &= \sum_{i=1}^{N_s} \mathbb{B}^i \left[K^i\right]^{-1} f^i && (3.38) \end{aligned}$$

We then first solve for the λ parameter and afterwards for the electrical field everywhere in the various domains.

3.4.2.2 Solving the linear system with common corner points

An other approach is to handle the corner points separately. Various Domain Decomposition techniques are based on this idea. For example, we can refer to the FETI-DP [39,109], FETI-DPH [42], or FETI-DPEM1 [41] methods. The term ”Dual-Primal” (DP) refers to the idea of directly imposing continuity constraints across the corner DOFs between subdomains and indirectly enforcing all other constraints by using dual variables (Lagrange multipliers).

Using the rc-notation of Section (3.3.2.2) and, in particular, the representation of the problem matrix and RHS vector (Eq. 3.26) we start by introducing

— The set of the Lagrange multipliers divided into $\lambda = \{\lambda_r, \lambda_c\}$

— The *signed* Boolean matrix $\mathbb{B}_r^i$ which is an analogue of the matrix $\mathbb{B}^i$ presented previously. The role of the matrix $\mathbb{B}_r^i$ is to extract from a given subdomain Ω^i a signed vector of all the interface boundary components, except the corner DOFs. The number of lines in this matrix is equal to the number of the effective points which are actually present on all the interfaces, but not on the corners. The number of columns corresponds to the number of ”r”-notation in the subdomain Ω^i.

Thus, let us rewrite the saddle point conditions (Eq. 3.33) taking this into account as follows :

$$\begin{bmatrix} K_{rr}^i & K_{rc}^i \\ K_{cr}^i & K_{cc}^i \end{bmatrix} \begin{bmatrix} E_r^i \\ E_c^i \end{bmatrix} = \begin{bmatrix} f_r^i \\ f_c^i \end{bmatrix} - \begin{bmatrix} \mathbb{B}_r^{i^T} \lambda_r \\ \lambda_c^i \end{bmatrix} \qquad (3.39)$$

$$\sum_{i=1}^{N_s} \mathbb{B}_r^i E_r^i = 0 \qquad (3.40)$$

In order to describe the ”Primal-part” of the DP idea we are going to introduce the projection Boolean matrix $\mathbb{Q}_{\mathbf{E}_c}^i$. This Boolean matrix maps the global corner DOFs numbering $\mathbf{E}_c$ to local corner DOFs numbering E_c^i. Mathematically, this can be expressed as

$$E_c^i = \mathbb{Q}_{\mathbf{E}_c}^i \mathbf{E}_c \qquad (3.41)$$

$\mathbb{Q}_{\mathbf{E}_c}^i$ is a matrix of the dimensions $N_c^i \times N_c^g$. N_c^i and N_c^g are respectively the number of corner points in the i-th subdomain and the global number of corner points in Ω.

We can now rewrite the saddle point conditions (Eq. 3.33) using the rc-notation and the presented Boolean matrices as [39, 41] :

$$\begin{aligned} & K_{rr}^i E_r^i + K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i \mathbf{E}_c = f_r^i - \mathbb{B}_r^{i^T} \lambda_r && (3.42) \\ & \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{i^T} K_{rc}^{i^T} E_r^i + \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{i^T} K_{cc}^i \mathbb{Q}_{\mathbf{E}_c}^i \mathbf{E}_c = \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{i^T} f_c^i && (3.43) \\ & \sum_{i=1}^{N_s} \mathbb{B}_r^i E_r^i = 0 && (3.44) \end{aligned}$$

where the term related to the set of Lagrange multipliers λ_c was cancelled out, because the tangential component of the magnetic field is continuous across the interface between the subdomains (assuming that there is no surface electric currents on the interface) [111], in other words :

$$\sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{iT} \lambda_c^i = 0 \tag{3.45}$$

Taking advantage of the diagonal nature of the K_{rr} matrix, and by replacing the value of E_r^i in the two other equations, we obtain a system of equations only in $\mathbf{E}_c$ and λ_r such that

$$\begin{pmatrix} F_{cc} & F_{cr} \\ F_{rc} & F_{rr} \end{pmatrix} \begin{pmatrix} \mathbf{E}_c \\ \lambda_r \end{pmatrix} = \begin{pmatrix} d_c \\ d_r \end{pmatrix} \tag{3.46}$$

with

$$F_{cc} = \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{iT} K_{cc}^i \mathbb{Q}_{\mathbf{E}_c}^i - \sum_{i=1}^{N_s} [K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i]^T K_{rr}^{i^{-1}} [K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i] \tag{3.47}$$

$$F_{cr} = -\sum_{i=1}^{N_s} [K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i]^T K_{rr}^{i^{-1}} \mathbb{B}_r^{iT} \tag{3.48}$$

$$F_{rc} = F_{cr}^T = -\sum_{i=1}^{N_s} \mathbb{B}_r^i K_{rr}^{i^{-1}} K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i \tag{3.49}$$

$$F_{rr} = -\sum_{i=1}^{N_s} \mathbb{B}_r^i K_{rr}^{i^{-1}} \mathbb{B}_r^{iT} \tag{3.50}$$

$$d_c = \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{iT} f_c^i - \sum_{i=1}^{N_s} [K_{rc}^i \mathbb{Q}_{\mathbf{E}_c}^i]^T K_{rr}^{i^{-1}} f_r^i \tag{3.51}$$

$$d_r = \sum_{i=1}^{N_s} \mathbb{B}_r^i K_{rr}^{i^{-1}} f_r^i \tag{3.52}$$

The above problem is a dual-primal one [39] because it relates the dual Lagrange multipliers λ_r to the primal field DOFs $\mathbf{E}_c$. By eliminating $\mathbf{E}_c$, the above system can be transformed into the following symmetric dual interface problem

$$\left(F_{rr} - F_{rc} F_{cc}^{-1} F_{cr}\right) \lambda_r = d_r - F_{rc} F_{cc}^{-1} d_c \tag{3.53}$$

which is closely related to the original FETI interface problem (Eq. 3.36).

The definition of the interface problem (3.53) highly depends on the problem under consideration. For the problems of second-order solid mechanics [39] this problem is *symmetric positive*, although for the electromagnetic problems this system of linear equations is generally *indefinite* [41]. It is worth mentioning that this relation contains a so-called *coarse* problem which serves as a preconditioning technique [39, 41]. We are going to consider this problem later on, in Section (5.3.3.2).

We then solve this system for λ_r and, afterwards for $\mathbf{E}_c$, and replace the obtained values in Eq. (3.39) to deduce E_r^i everywhere in the subdomains.

3.4.3 Interpretation in terms of boundary conditions

As we have already noticed, the system of linear equations (3.17) for the given subdomain Ω^i is an incomplete one, because it lacks the interface condition applied to the internal boundary Γ^i. In this subsection we would like to discuss the type of the boundary conditions associated to the case of Section (3.4).

In the framework of this section, we assume the continuity of the tangential component of the electric field across the subdomain interfaces of the form (3.27) which can be also written for the discretized subdomains as Eq. (3.29). After having constructed the Lagrangian formulation we obtain the saddle point of the form (3.33). If we now consider the first equation which is associated to the given subdomain Ω^i (Eq. 3.33), combined with the incomplete system of linear equations (3.17) than the physical sense of the lagrange multipliers becomes clear. For the sake of clarity we will write these equations one more time :

$$K^i E^i = f^i + \int\limits_{\Gamma^i} \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n} \, \Psi^i \, dS \tag{3.54}$$

$$K^i E^i = f^i - \mathbb{B}^{i^T} \lambda \tag{3.55}$$

Taking advantage from the discussions in Section (2.2.2) we can associate the set of Lagrange multipliers λ with an unknown ***Neumann***-type boundary condition for each local subdomain Ω^i on the internal interface Γ^i. This boundary condition can be denoted as

$$K^i E^i = f^i + \int\limits_{\Gamma^i} \Lambda^i \, \Psi^i \, dS$$

$$\Lambda^i = \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n^i}$$

The constraint that we impose between subdomains is the equality of the electric fields. Also we know that for each Γ^{ij} we have $n^i = -n^j$. This leads us to the following relation between Lagrange multipliers [101] :

$$\Lambda^i = \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n^i} = -\frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^j}{\partial n^j} = -\Lambda^j \tag{3.56}$$

And, finally, as the constraint (Eq. 3.29) is represented by only one Lagrange multiplier $\Lambda^{i\leftrightarrow j}$, we obtain the resulting relation for the unknown boundary condition :

$$\mathbb{B}^{i^T} \lambda = -\int\limits_{\Gamma^i} \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n} \, \Psi^i \, dS = -\int\limits_{\Gamma^i} \Lambda^{i\leftrightarrow j} \, \Psi^i \, dS$$

$$\Lambda^{i\leftrightarrow j} = \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n^i} = -\frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^j}{\partial n^j}$$

3.5 One Lagrange multiplier with an augmented Lagrangian

It is well-known that augmented Lagrangian schemes provide solutions which are better regularized than the classical Lagrangian saddle-point solutions [36,37,43]. In this subsection we are going to deal with FETI methods based on introducing still only one set of Lagrange multipliers per interface in order to enforce continuity of the tangential component of the *electric fields* on each interface between subdomains. But now, we will search for the unknown boundary conditions for the internal interfaces by taking advantage of the *Augmented* Lagrangian formalism. In literature this type of Domain Decomposition techniques is known as FETI-1 method [36,101].

3.5.1 An unique field continuity condition

The continuity of the tangential component of the fields at the internal interfaces in 2D case and for the specific polarization considered here can still be written as follows :

$$\mathcal{E}|_{\Omega_i}(\vec{r}) = \mathcal{E}|_{\Omega_j}(\vec{r}) \qquad \forall \vec{r} \in \Gamma^{ij} = \Gamma^i \cap \Gamma^j$$

Thus, in the finite-element space we can use the same relation (Eq. 3.29) in order to impose the continuity of the subdomain solutions across the subdomain interfaces, with the same matrix $\mathbb{B}$ defined as previously.

3.5.2 The augmented Lagrangian formalism

The augmented Lagrangian functional which has been proposed in [36,37] can be expressed as

$$\mathfrak{L}(\{E^i\},\lambda) = L(\{E^i\},\lambda) + \frac{1}{2}\sum_{i=1}^{N_s} \left(\mathbb{B}^i\ E^i\right)^T \mathcal{M}^i_\Gamma\ \left(\mathbb{B}^i\ E^i\right) \tag{3.57}$$

where $L(\{E^i\},\lambda)$ is a classical Lagrangian functional (Eq. 3.30) discussed before and $\mathcal{M}^i_\Gamma$ is an interface matrix, i.e. a matrix defined on the internal interface Γ^i, which can be constructed as wanted. The saddle-point conditions of this augmented Lagrangian functional are given by

$$K^iE^i - f^i + \mathbb{B}^{i^T}\lambda + (\mathbb{B}^{i^T}\mathcal{M}^i_\Gamma\ \mathbb{B}^i)E^i = 0 \tag{3.58}$$

$$\sum_{i=1}^{N_s}\mathbb{B}^iE^i = 0 \tag{3.59}$$

We can rewrite this system into

$$\tilde{K}^iE^i - f^i + \mathbb{B}^{i^T}\lambda = 0 \tag{3.60}$$

$$\sum_{i=1}^{N_s}\mathbb{B}^iE^i = 0 \tag{3.61}$$

where the new matrix $\tilde{K}^i$ is determined as follows :

$$\tilde{K}^i = K^i + \mathbb{B}^{i^T}\mathcal{M}^i_\Gamma\ \mathbb{B}^i \tag{3.62}$$

The term in $\mathbb{B}^{i^T}\mathcal{M}^i_\Gamma\ \mathbb{B}^i$ only refers to the DOFs which are located on the interface Γ^i of the given subdomain Ω^i. There are full freedom for defining this term. The matrix $\mathcal{M}^i_\Gamma$ is referred to as the augmented matrix in [112]. As we can see, we obtained an almost identical set of equations as for the method considered previously in Section (3.4.2). Thus, in order to construct and then solve the Interface problem for FETI1 and FETI-DP1 methods we follow exactly the same procedure as written in Sections (3.4.2.1) and (3.4.2.2) respectively by just replacing K^i with $\tilde{K}^i$.

3.5.3 Interpretation in terms of boundary conditions

The resulting equation obtained after construction of the augmented Lagrangian functional (Eq. 3.60) combined with the incomplete form of the Helmholtz boundary value problem after discretization (Eq. 3.17) can be written as :

$$\begin{aligned} K^iE^i &= f^i + \int_{\Gamma^i} \frac{1}{\mu_r}\frac{\partial \mathcal{E}^i_h}{\partial n}\,\Psi^i\,dS \\ \left(K^i + \mathbb{B}^{i^T}\mathcal{M}^i_\Gamma\ \mathbb{B}^i\right)E^i &= f^i - \mathbb{B}^{i^T}\lambda \end{aligned} \tag{3.63}$$

Comparing the structure of these equations, the role of the Lagrange multipliers and the matrix $\mathcal{M}_\Gamma^i$ gets clear. For $\mathcal{M}_\Gamma^i = M_\Gamma^i(\alpha)$ (see Eq. 3.21), the FETI method based on the modified Lagrangian formulation (3.57) is equivalent to equipping the exterior Helmholtz boundary value problem (3.17) with the ***Robin***-type boundary condition, and identifying each of the left- and right-hand sides of this condition with the same Lagrange multiplier λ as follows :

$$\mathbb{B}^{i^T}\lambda = \frac{1}{\mu_r}\frac{\partial \mathcal{E}^i}{\partial n^i} + \alpha\mathcal{E}^i = -\frac{1}{\mu_r}\frac{\partial \mathcal{E}^j}{\partial n^j} + \alpha\mathcal{E}^j = \mathbb{B}^{j^T}\lambda \tag{3.64}$$

Note that the relation above takes place only if $\Gamma^{ij} = \Omega^i \cap \Omega^j \neq \emptyset$.

3.6 Two Lagrange multipliers per interface internal node point

For arbitrary mesh partition, the specific treatment of the subdomain interfaces described in Section (3.5), which is designed for ensuring a constant sign of the regularizing matrix M_Γ^i, results from the fact that the FETI1 method considers only one normal and therefore employs only one Lagrange multiplier per mesh partition interface.

Another approach is designed to make the algorithm fully numerically scalable regardless of the electrical size of the subdomains or the frequency of the electromagnetic fields. In order to regularize the subdomain problems independently of the pattern of the mesh decomposition we equip the FETI method with two Lagrange multiplier fields as described in this section.

Such a strategy, which can also be viewed as introducing two diferent normals at each interface (one on each side), is often referred to, in the domain decomposition literature, as a three-field method [105]. The approach of using two Lagrange multipliers per interface helps us to avoid the problems with the domain-numbering and the construction of normals to domains arising in the case of one Lagrange multiplier [36,37]. Moreover, the spurious solutions which arise traditionally are damped by the addition of this extra term.

3.6.1 Two field continuity conditions per interface

The continuity of the tangential component of the fields at the internal interfaces can still be written as follows :

$$\mathcal{E}|_{\Omega_i}(\vec{r}) = \mathcal{E}|_{\Omega_j}(\vec{r}) \qquad \forall \vec{r} \in \Gamma^{ij} = \Gamma^i \cap \Gamma^j \tag{3.65}$$

In order to impose this constraint we present an alternative to the treatment presented in Section (3.4.1) by defining a unique solution $\mathcal{U}^{i\leftrightarrow j}$ on the boundary Γ^{ij} between the subdomains Ω^i and Ω^j and introducing two Lagrange multipliers for each subdomain : the objective of the first Lagrange multiplier is to enforce the compatibility between $\mathcal{E}|_{\Omega_i}$ and $\mathcal{U}^{i\leftrightarrow j}$, and the objective of the second Lagrange multiplier field is to enforce the compatibility between $\mathcal{E}|_{\Omega_j}$ and $\mathcal{U}^{i\leftrightarrow j}$.

To that end, we can rewrite the continuity of the tangential component of the electric fields (3.65) as follows :

$$\mathcal{E}|_{\Omega_i}(\vec{r}) = \mathcal{U}^{i\leftrightarrow j}(\vec{r}) \qquad \forall \vec{r} \in \Gamma^i \tag{3.66}$$

$$\mathcal{E}|_{\Omega_j}(\vec{r}) = \mathcal{U}^{i\leftrightarrow j}(\vec{r}) \qquad \forall \vec{r} \in \Gamma^j \tag{3.67}$$

We partition the interface boundary Γ^i into interface edges Γ_l^i using the following guidelines :

— an interface edge is defined as a collection of connecting interface nodes ;

— each interface node is still not assigned to one and only one subdomain, but belongs at least to two subdomains ;

— in the local sense a corner point is a point which belongs to two interfaces Γ_l^i of the same subdomain Ω^i.

As previously we denote by E_I^i the discretized vector of unknowns $\mathcal{E}|_{\Omega_i}(\vec{r})$, and similarly, $U^{i\leftrightarrow j}$ corresponds to the common unique solution $\mathcal{U}^{i\leftrightarrow j}(\vec{r})$. We now introduce Boolean matrices $\mathbb{D}^i$ which enable to extract only the node points which are on the interface of Ω^i, i.e.,

$$\mathbb{D}^i E^i = E_I^i \tag{3.68}$$

Note that the matrix $\mathbb{D}^i$ is defined differently from the matrix $\mathbb{B}^i$ from Section (3.4.1). Here it denotes a Boolean matrix that simply extracts the unknowns on the interface without sign assignment. This Boolean matrix is of the size $n \times m$, where n is the number of the points located on the Γ^i, and m denotes the number of all the points in Ω^i.

Similarly, we introduce Boolean matrices $\mathbb{T}^{i\to j}$ which aims to extract only the node points which are on the interface edge Γ^{ij} from the interface node points of Ω^i. The Boolean matrix $\mathbb{T}^{i\to j}$ is of the size $n \times m$, where n is the number of the points located on Γ^{ij}, including the corners, and m corresponds to the number of the points located on the Γ^{ij}. This enables to write that

$$\begin{aligned} \mathbb{T}^{i\to j} E_I^i &= E^{i\to j} \\ \mathbb{T}^{i\to j} \lambda^i &= \lambda^{i\to j} \end{aligned} \tag{3.69}$$

We have in particular

$$\lambda^i = \sum_{j \in nbr(\Omega^i)} \mathbb{T}^{i\to j^T} \lambda^{i\to j} \tag{3.70}$$

where the array $nbr(\Omega^i)$ contains all the numbers of subdomains which are neighbors to Ω^i.

3.6.2 Augmented Lagrangian formalism

As we have done previously for the FETI1-formulation, we introduce an augmented Lagrangian formalism combined with a special set of matrices $\mathcal{M}$.

$$\begin{aligned} \mathfrak{L}(\{E^i\}, \{\lambda^{i\to j}\}, \{U^{i\leftrightarrow j}\}) = &\sum_{i=1}^{N_s} \left(\frac{1}{2} K^i E^i - f^i\right)^T E^i + \sum_{i=1}^{N_s} \sum_{j \in nbr(\Omega^i)} \lambda^{i\to j^T} \left(E^{i\to j} - U^{i\leftrightarrow j}\right) + \\ &\sum_{i=1}^{N_s} \sum_{j \in nbr(\Omega^i)} \frac{1}{2} \left(E^{i\to j^T} \mathcal{M}^{i\to j}\, E^{i\to j} - U^{i\leftrightarrow j^T} \mathcal{M}^{i\to j}\, U^{i\leftrightarrow j}\right) \end{aligned} \tag{3.71}$$

Note that this equation is not "correct" in terms of unknowns. Indeed, Eq. (3.71) contains a term $E^{i\to j}$ which does not take a part of the optimization variables of $\mathfrak{L}(\{E^i\}, \{\lambda^{i\to j}\}, \{U^{i\leftrightarrow j}\})$. Thus, with the aid of Eq. (3.69) we present the vector of unknowns on the local interface Γ^{ij} as follows :

$$E^{i\to j} = \mathbb{T}^{i\to j} \mathbb{D}^i E^i \tag{3.72}$$

Then using this relation, the Lagrangian (3.71) can be rewritten as :

$$\begin{aligned} &\mathfrak{L}(\{E^i\}, \{\lambda^{i\to j}\}, \{U^{i\leftrightarrow j}\}) = \\ &\quad \sum_{i=1}^{N_s} \left(\frac{1}{2} K^i E^i - f^i\right)^T E^i + \sum_{i=1}^{N_s} \sum_{j \in nbr(\Omega^i)} \lambda^{i\to j^T} \left([\mathbb{T}^{i\to j} \mathbb{D}^i E^i] - U^{i\leftrightarrow j}\right) + \\ &\qquad \sum_{i=1}^{N_s} \sum_{j \in nbr(\Omega^i)} \frac{1}{2} \left([\mathbb{T}^{i\to j} \mathbb{D}^i E^i]^T \mathcal{M}^{i\to j}\, [\mathbb{T}^{i\to j} \mathbb{D}^i E^i] - U^{i\leftrightarrow j^T} \mathcal{M}^{i\to j}\, U^{i\leftrightarrow j}\right) \end{aligned} \tag{3.73}$$

Unlike the case described in Section (3.4) and (3.5), the Lagrange parameter λ corresponds now to the the double number of effective points which are actually present on all the interfaces. The saddle point will provide the solution of the optimization problem we are interested in. To find it, we consider the set of KKT conditions :

$$\begin{cases} \nabla_{E^i} \quad \mathfrak{L}(\{E^i\},\{\lambda^{i\to j}\},\{U^{i\leftrightarrow j}\}) = 0, \quad i=1,..,N_s \\ \nabla_{\{\lambda^{i\to j}\}}\mathfrak{L}(\{E^i\},\{\lambda^{i\to j}\},\{U^{i\leftrightarrow j}\}) = 0 \\ \nabla_{\{U^{i\leftrightarrow j}\}}\mathfrak{L}(\{E^i\},\{\lambda^{i\to j}\},\{U^{i\leftrightarrow j}\}) = 0 \end{cases} \tag{3.74}$$

After computation the derivatives of the Lagrangian with respect to the various parameters we obtain

$$\nabla_{E^i}\mathfrak{L} = \tilde{K}^i E^i - f^i + \mathbb{D}^{i^T} \sum_{j\in nbr(\Omega^i)} \mathbb{T}^{i\to j^T} \lambda^{i\to j}, \quad i=1,..,N_s \tag{3.75}$$

$$\nabla_{\{\lambda^{i\to j}\}}\mathfrak{L} = [\mathbb{T}^{i\to j}\mathbb{D}^i E^i] - U^{i\leftrightarrow j}, \quad \forall i,j \in nbr(\Omega^i) \tag{3.76}$$

$$\nabla_{\{U^{i\leftrightarrow j}\}}\mathfrak{L} = -\lambda^{i\to j} - \lambda^{j\to i} - (\mathcal{M}^{i\to j} + \mathcal{M}^{j\to i})U^{i\leftrightarrow j} \tag{3.77}$$

where the local problem-matrix $\tilde{K}^i$ is similar to the one from Eq. (3.60), but constructed in the form

$$\tilde{K}^i = K^i + \sum_{j\in nbr(\Omega^i)} [\mathbb{T}^{i\to j}\mathbb{D}^i]^T \mathcal{M}^{i\to j} \, [\mathbb{T}^{i\to j}\mathbb{D}^i] \tag{3.78}$$

Note that the system of equations (3.75-3.77) is absolutely identical to the Euler equations associated to the Lagrangian proposed in [36], if we take into account the fact that the interface unique solution $U^{i\leftrightarrow j}$ can have only two contributions (which is a logical conclusion), these come from the subdomains Ω^i and Ω^j.

The first of the KKT conditions ($\nabla_{\{E^i\}}\mathfrak{L} = 0$) yields the general FEM equation for the subdomain Ω^i

$$\tilde{K}^i E^i = f^i - \mathbb{D}^{i^T} \lambda^i \tag{3.79}$$

We then eliminate the common interface solution $U^{i\leftrightarrow j}$ in order to reduce the system. Thus, the two last KKT conditions give

$$\lambda^{i\to j} + \lambda^{j\to i} = -(\mathcal{M}^{i\to j} + \mathcal{M}^{j\to i})E^{j\to i} \tag{3.80}$$

with

$$E^{j\to i} = E^{i\to j} \tag{3.81}$$

3.6.3 Separating interface and corner points

In order to consider the FETI-DPEM2 classical and full methods we partition the vector of unknowns according to the "rc"-notation described in Section (3.3.2.2). Thus, some geometrical matrices should be presented.

First of all, we introduce an analogue of the matrix $\mathbb{D}^i$ - a Boolean matrix $\mathbb{D}^i_r$ which extracts the interface DOFs of the subdomain Ω^i according to the "rc"-notation. Mathematically it can be expressed as :

$$\mathbb{D}^i_r E^i_r = E^i_I \tag{3.82}$$

Let us denote the size of matrix $\mathbb{D}^i_r$ as $n \times m$. Then, n corresponds to the number of the points located on the internal interface Γ^i without the corner nodes, while m is the total number of the points inside Ω^i except for the corners.

Secondly, we will use the Boolean matrix $\mathbb{Q}^i_{\mathbf{E}_c}$ (Eq. 3.41) which enables here to map the global corner DOFs numbering $\mathbf{E}_c$ with the local numbering E^i_c.

Then, we divide the Boolean matrix $\mathbb{T}$ (Eq. 3.69) into two matrices $\mathbb{T}_r$ and $\mathbb{T}_c$. The first of them extracts the interface nodes related to the given interface Γ^{ij} from all the interfaces node points of Ω^i, which can be mathematically expressed as :

$$\begin{aligned} \mathbb{T}_r^{i\to j} E_I^i &= E_r^{i\to j} \\ \mathbb{T}_r^{i\to j} \lambda_r^i &= \lambda_r^{i\to j} \end{aligned} \tag{3.83}$$

The second Boolean matrix $\mathbb{T}_c$ enables to select all the corner DOFs related to the given internal interface Γ^{ij} from all the corners of the subdomain Ω^i as

$$\mathbb{T}_c^{i\to j} E_c^i = E_c^{i\to j} \tag{3.84}$$

The idea of the construction of these matrices is presented in Figure (3.3). Note that the local indexes of the selected nodes in this figure correspond to the non-null positions in this matrices, while the number of selected points represents the vertical size of the matrices $\mathbb{T}_r^{i\to j}$ and $\mathbb{T}_c^{i\to j}$.

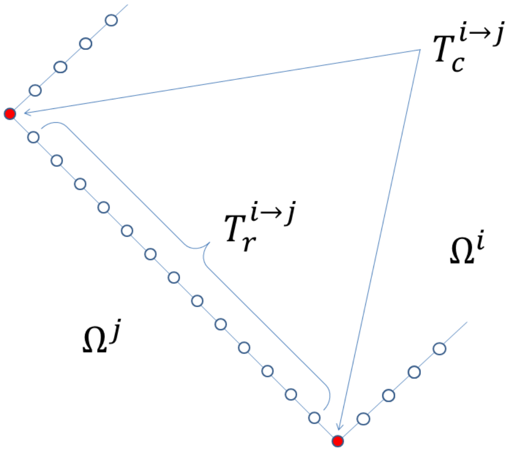


FIGURE 3.3 – Schematic presentation of the unknowns belonging to the subdomain Ω^i selected by matrices $\mathbb{T}_r^{i\to j}$ and $\mathbb{T}_c^{i\to j}$ to the interface Γ^{ij}.

As you can see, the matrix $\mathbb{T}_r^{i\to j}$ selects all the points related to the interface Γ^{ij}, while the another matrix $\mathbb{T}_c^{i\to j}$ extracts only two points in this particular case. These are the corner points related to the interface Γ^{ij}. It is simple to show, that the size of the matrix $\mathbb{T}_r^{i\to j}$ is equal to the number of the interface points related to the given interface × the number of all the interface points belonging to the given subdomain, while the size of the matrix $\mathbb{T}_c^{i\to j}$ is equal to the number of corner points related to this interface × the number of the corner points belonging to the given subdomain.

Note that we ***can not*** write $\mathbb{T}_c^{i\to j}\lambda_c^i = \lambda_c^{i\to j}$ as it has been done for the previous matrix, because the number of the Lagrange multipliers λ_c^i is ***twice bigger*** than the number of the corner points E_c^i belonging to the internal interface Γ^i of the subdomain Ω^i. The special treatment of the Lagrange multipliers related to the corner points is going to be considered later on, in Section (3.6.4.2).

Finally, another Boolean $\mathbb{Q}_{\lambda_r}^i$ is introduced to extract the Lagrange multipliers λ_r^i related to the given subdomain Ω^i from all the list of Lagrange multipliers λ_r. This can be mathematically expressed as :

$$\mathbb{Q}_{\lambda_r}^i \lambda_r = \lambda_r^i \tag{3.85}$$

If we denote the size of this matrix as $n \times m$ then n will be the number of points located on the internal interface Γ^i except for the corners, whereas the term m will correspond to the total number of the dual-vector λ_r. Similarly to Eq. (3.70), the following relations can be derived :

$$\begin{aligned} \lambda_r^i &= \textstyle\sum_{j\in nbr(\Omega^i)} \mathbb{T}_r^{i\to j^T} \lambda_r^{i\to j} \\ \lambda &= \textstyle\sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{iT} \lambda^i \end{aligned} \tag{3.86}$$

These matrices are transposed [41, 43] because we assemble the local contributions into more global ones.

3.6.4 Finite-element analysis of the subdomain problem

In order to construct the interface problem corresponding to the FETI2 method we will use an approach which is a little bit different from the one of [36]. It was introduced in [43]. Using this approach we can explain the main difference between the FETI-DPEM2 classical method and the one, proposed in this PhD work.

Applying the "rc"-notation to Eq. (3.79) we obtain an important relation for the vectors of unknowns E_r^i and E_c^i for the local subdomain Ω^i

$$\begin{bmatrix} \tilde{K}_{rr}^i & \tilde{K}_{rc}^i \\ \tilde{K}_{cr}^i & \tilde{K}_{cc}^i \end{bmatrix} \begin{bmatrix} E_r^i \\ E_c^i \end{bmatrix} = \begin{bmatrix} f_r^i \\ f_c^i \end{bmatrix} - \begin{bmatrix} \mathbb{D}_r^{i\,T} \lambda_r^i \\ \lambda_c^i \end{bmatrix} \tag{3.87}$$

From Eq. (3.87) we can find that the subdomain system matrices for different subdomains become decoupled. Indeed, this system of linear equations is written only for the subdomain Ω^i, but not for its neighbors, while all the interactions with the neighboring subdomains are included in the mixed boundary condition at the interfaces given by Eq. (3.80). By using the first equation in system (3.87), the electric field can be found as

$$E_r^i = \tilde{K}_{rr}^{i^{-1}} \left(f_r^i - \mathbb{D}_r^{iT} \lambda_r^i - \tilde{K}_{rc}^i E_c^i \right) \tag{3.88}$$

Then we derive the subdomain level corner DOFs related system by eliminating E_r^i as

$$(\tilde{K}_{cc}^i - \tilde{K}_{cr}^i \tilde{K}_{rr}^{i^{-1}} \tilde{K}_{rc}^i) \mathbb{Q}_{\mathbf{E}_c}^i \mathbf{E}_c = f_c^i - \lambda_c^i - \tilde{K}_{cr}^i \tilde{K}_{rr}^{i^{-1}} f_r^i + \tilde{K}_{cr}^i \tilde{K}_{rr}^{i^{-1}} \mathbb{D}_r^{iT} \lambda_r^i \tag{3.89}$$

This equation can be considered as the subdomain contribution to a global corner DOFs related finite element system. After assembling all the subdomain contributions we obtain the global corner DOFs related finite element system, which can be written as follows :

$$\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}\, \mathbf{E}_c - \mathrm{F}_{\mathbf{E}_c\lambda_r}\, \lambda_r + \mathrm{F}_{\mathbf{E}_c\lambda_c}\, \lambda_c = \mathrm{d}_{\mathbf{E}_c} \tag{3.90}$$

with

$$\begin{aligned} \mathrm{F}_{\mathbf{E}_c\mathbf{E}_c} &= \mathcal{S}_{\mathbf{E}_c} \left(K_{cc}^i - K_{cr}^i K_{rr}^{i^{-1}} K_{rc}^i \right) \mathbb{Q}_{\mathbf{E}_c}^i \\ \mathrm{F}_{\mathbf{E}_c\lambda_r} &= \mathcal{S}_{\mathbf{E}_c} \left(K_{cr}^i K_{rr}^{i^{-1}} \mathbb{D}_r^i \right) \mathbb{Q}_{\lambda_r}^i \\ \mathrm{d}_{\mathbf{E}_c} &= \mathcal{S}_{\mathbf{E}_c} \left(f_c^i - K_{cr}^i K_{rr}^{i^{-1}} f_r^i \right) \end{aligned}$$

where the operator $\mathcal{S}_{\mathbf{E}_c}$ is defined by

$$\mathcal{S}_{\mathbf{E}_c} u \;=\; \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{iT} u^i \tag{3.91}$$

This operator assembles local contributions from all the subdomain corners in order to construct a part of the Interface Problem related to the $\mathbf{E}_c$ vector.

Note that we do not introduce the matrix $\mathrm{F}_{\mathbf{E}_c\lambda_c}$, because this is the key difference between the FETI-DPEM2 classical and full methods. So, it is going to be presented later.

Let us start the inter-subdomain analysis with the main transmission condition written in the form of Eq. (3.80). In order to explain the principal difference between the classical FETI-DPEM2 method and the method proposed in [55], we apply the "rc"-notation to Eq. (3.80). This yields :

$$\begin{bmatrix} \lambda_r^{i\to j} \\ \lambda_c^{i\to j} \end{bmatrix} + \begin{bmatrix} \lambda_r^{j\to i} \\ \lambda_c^{j\to i} \end{bmatrix} = -\begin{bmatrix} \mathcal{W}_{rr}^{i\leftrightarrow j} & \mathcal{W}_{rc}^{i\leftrightarrow j} \\ \mathcal{W}_{cr}^{i\leftrightarrow j} & \mathcal{W}_{cc}^{i\leftrightarrow j} \end{bmatrix} \begin{bmatrix} E_r^{j\to i} \\ E_c^{j\to i} \end{bmatrix} \quad \forall\, \Gamma^{ij} \tag{3.92}$$

where $\mathcal{W}_{*-}^{i\leftrightarrow j} = \mathcal{M}_{*-}^{i\to j} + \mathcal{M}_{*-}^{j\to i}$.

3.6.4.1 FETI-DPEM2 classical method

In various works related to the FETI-DPEM2 method (for ex. [43, 48, 96, 111, 113]), we can note that there is no contribution of the corner DOFs to the set of Lagrange multipliers related to the interface DOFs and the size of Eq. (3.80) is equal to the number of interface points related to the interface Γ^{ij}. Thus, we can conclude that in the classical FETI-DPEM2 method, the matrices $\mathcal{W}_{rc}^{i\leftrightarrow j} = \mathcal{W}_{cr}^{i\leftrightarrow j} = \mathcal{W}_{cc}^{i\leftrightarrow j}$ are all set to 0. Let us first focus on the first row of Eq. (3.92), that is the one related to λ_r. Having applied the notations of Eq. (3.92), in the FETI-DPEM2 classical method, we have :

$$\lambda_r^{i\to j} + \lambda_r^{j\to i} = -\mathcal{W}_{rr}^{i\leftrightarrow j} E_r^{j\to i} \tag{3.93}$$

Note that this equation is identical to the one from [43, 111] in another notations. Indeed, the size of the unknown Lagrange multipliers and the definition of the matrices are absolutely the same.

With the aid of the projection matrices presented in Section (3.6.3), we transform Eq. (3.93) as follows :

$$\lambda_r^{i\to j} + \lambda_r^{j\to i} = -\mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} \mathbb{D}_r^j E^j$$

Then, using Eq. (3.88) we can simplify the equation above by eliminating E_r^j as :

$$\lambda_r^{i\to j} + \lambda_r^{j\to i} = -\mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} \mathbb{D}_r^j {K_{rr}^j}^{-1} \left(f_r^j - {\mathbb{D}_r^j}^T \lambda_r^j - K_{rc}^j E_c^j \right)$$

If we summate it over j with the help of Eq. (3.86) we will obtain :

$$\lambda_r^i + \sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \left[\mathbb{T}_r^{j\to i} - \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rr}^i \right] \lambda_r^j =$$

$$\sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rc}^j \mathbb{Q}_{\mathbf{E}_c}^j \mathbf{E}_c - \sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} d_r^j$$

Then, using the same idea for the summation over i we obtain :

$$\lambda_r + \sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{i^T} \sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \left[\mathbb{T}_r^{j\to i} - \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rr}^i \right] \lambda_r^j =$$

$$\sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{i^T} \sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rc}^j \mathbb{Q}_{\mathbf{E}_c}^j \mathbf{E}_c - \sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{i^T} \sum_{j\in nbr(\Omega^i)} {\mathbb{T}_r^{i\to j}}^T \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} d_r^j$$

or, simply

$$\lambda_r + \mathcal{S}_{\lambda_r} \left[\mathbb{T}_r^{j\to i} - \mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rr}^i \right] \mathbb{Q}_{\lambda_r}^j \lambda_r =$$

$$\mathcal{S}_{\lambda_r} \left[\mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} F_{rc}^j \mathbb{Q}_{\mathbf{E}_c}^j \right] \mathbf{E}_c - \mathcal{S}_{\lambda_r} \left[\mathcal{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i} d_r^j \right]$$

Here the operator $\mathcal{S}_{\lambda_r}$ is defined by

$$\mathcal{S}_{\lambda_r} u \;=\; \sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{iT} \sum_{j\in nbr(i)} \mathbb{T}_r^{i\to j^T} u^{j\to i} \tag{3.94}$$

This operator can assemble the contributions from all the interfaces and, then, from all the subdomains, in order to construct a part of the Interface Problem related to the λ_r vector.

This is the first line of the interface problem of the FETI-DPEM2 classical method [43,111], which can be written as

$$\mathrm{F}_{\lambda_r\lambda_r}\lambda_r - \mathrm{F}_{\lambda_r\mathbf{E}_c}\mathbf{E}_c = -\mathrm{d}_{\lambda_r} \tag{3.95}$$

where the main matrices are defined as

$$\begin{aligned}
\mathrm{F}_{\lambda_r\lambda_r} &= \mathcal{S}_{\lambda_r}\left[\mathbb{T}_r^{j\to i} - \mathcal{W}_{rr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}F_{rr}^{i}\right]\mathbb{Q}_{\lambda_r}^{j} + \mathbf{I} && (3.96)\\
\mathrm{F}_{\lambda_r\mathbf{E}_c} &= \mathcal{S}_{\lambda_r}\left[\mathcal{W}_{rr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}F_{rc}^{j}\right]\mathbb{Q}_{\mathbf{E}_c}^{j} && (3.97)\\
\mathrm{d}_{\lambda_r} &= \mathcal{S}_{\lambda_r}\left[\mathcal{W}_{rr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}d_r^{j}\right] && (3.98)\\
F_{rr}^{i} &= \mathbb{D}_r^{i}K_{rr}^{i^{-1}}\mathbb{D}_r^{iT} && (3.99)\\
F_{rc}^{i} &= \mathbb{D}_r^{i}K_{rr}^{i^{-1}}K_{rc}^{i} && (3.100)\\
d_r^{i} &= \mathbb{D}_r^{i}K_{rr}^{i^{-1}}f_r^{i} && (3.101)
\end{aligned}$$

Let us now focus on the second row of Equation (3.92), that is on the Lagrange multipliers λ_c related to the corner points. The specificity of the FETI-DPEM2 classical method is to set

$$\lambda_c^{i\to j} + \lambda_c^{j\to i} = 0 \quad \forall\,\Gamma^{ij} \tag{3.102}$$

This equation means that the summation of all the λ_c over each global corner point is equal to 0 [48], i.e.

$$\mathrm{F}_{\mathbf{E}_c\lambda_c} = \sum_{i=1}^{N_s} \mathbb{Q}_{\mathbf{E}_c}^{i}{}^{T}\lambda_c^{i} = 0 \tag{3.103}$$

Taking into account this term, the global corner DOFs related system of equations (Eq. 3.90) can be written as follows :

$$\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}\;\mathbf{E}_c - \mathrm{F}_{\mathbf{E}_c\lambda_r}\;\lambda_r = \mathrm{d}_{\mathbf{E}_c} \tag{3.104}$$

After combining Eqs. (3.95) and (3.104) we obtain

$$\begin{bmatrix} \mathrm{F}_{\lambda_r\lambda_r} & -\mathrm{F}_{\lambda_r\mathbf{E}_c} \\ -\mathrm{F}_{\mathbf{E}_c\lambda_r} & \mathrm{F}_{\mathbf{E}_c\mathbf{E}_c} \end{bmatrix} \begin{bmatrix} \lambda_r \\ \mathbf{E}_c \end{bmatrix} = \begin{bmatrix} -\mathrm{d}_{\lambda_r} \\ \mathrm{d}_{\mathbf{E}_c} \end{bmatrix} \tag{3.105}$$

Following the idea presented in [43,48], $\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}^{-1}$, which is referred to as the coarse problem and has to be solved at each iteration step, couples all the subdomains by propagating the residual error globally and increases the convergence rate of the FETI-DPEM2 interface problem. Thus, we then eliminate $\mathbf{E}_c$ and derive the FETI-DPEM2 classical interface equation [43,113] for the dual unknowns

$$\left(\mathrm{F}_{\lambda_r\lambda_r} - \mathrm{F}_{\lambda_r\mathbf{E}_c}\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}^{-1}\mathrm{F}_{\mathbf{E}_c\lambda_r}\right)\lambda_r = -\mathrm{d}_{\lambda_r} - \mathrm{F}_{\lambda_r\mathbf{E}_c}\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}^{-1}\mathrm{d}_{\mathbf{E}_c} \tag{3.106}$$

The interface problem is generally indefinite [43,97,111] and can be solved using a Krylov subspace method. After the resolution of the interface problem, the electric field inside each subdomain can be evaluated individually by using the known mixed boundary condition at the interface for each subdomain with the aid of Eq. (3.88). The main advantage of the DD idea is that by introducing the set of Lagrange multipliers, the original problem is reduced to an interface problem in order to find the unknown boundary condition.

3.6.4.2 FETI-DPEM2-full method

The term "full" in the FETI-DPEM2-full method [55] refers to the idea of considering the full set of relations in Eq. (3.92) both for the λ_r and λ_c Lagrange multipliers. Unlike the classical FETI-DPEM2-full method, all the matrices $\mathcal{W}$ in Eq. (3.92) are different from 0.

To start with, we focus on the first row of the interface equation (3.92), that is the one related to λ_r.

$$\lambda_r^{i\to j} + \lambda_r^{j\to i} = -\mathcal{W}_{rr}^{i\leftrightarrow j} E_r^{j\to i} - \mathcal{W}_{rc}^{i\leftrightarrow j} E_c^{j\to i} \tag{3.107}$$

The size of this equation corresponds also to the number of the interface DOFs excluding the corner ones. But unlike the FETI-DPEM2 classical method, in the full approach we take into account the contribution from the corner points to the nearly located interface points by keeping the term $\mathcal{W}_{rc}^{i\leftrightarrow j} E_c^{j\to i}$. This leads us to a small change in the first line of the interface problem (3.105). More precisely, in the matrix $\mathrm{F}_{\lambda_r \mathbf{E}_c}$ which is equal now to

$$\mathrm{F}_{\lambda_r \mathbf{E}_c} = \mathcal{S}_{\lambda_r}\left[\mathcal{W}_{rr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}F_{rc}^{j} - \mathcal{W}_{rc}^{i\leftrightarrow j}\mathbb{T}_c^{j\to i}\right]\mathbb{Q}_{\mathbf{E}_c}^{j} \tag{3.108}$$

The main difference between the classical and full approaches of the FETI-DPEM2 method consists in the contribution of the second row of the main interface equation (3.92) which for the FETI-DPEM2-full method can be written as

$$\lambda_c^{i\to j} + \lambda_c^{j\to i} = -\mathcal{W}_{cr}^{i\leftrightarrow j} E_r^{j\to i} - \mathcal{W}_{cc}^{i\leftrightarrow j} E_c^{j\to i} \tag{3.109}$$

The treatment of this equation is more tricky. Indeed, as the corner DOFs $\mathbf{E}_c$ are gathered and ordered globally, it is not possible to calculate the full set of Lagrange multipliers both for $\lambda_c^{i\to j}$ and $\lambda_c^{j\to i}$ from each side of the interface Γ^{ij}. Thus, in the framework of the FETI-DPEM2-full approach we are interested in searching a special set of Lagrange multipliers $\boldsymbol{\lambda}_c^{i\leftrightarrow j}$, *one* per each interface Γ^{ij}. Taking this into account we can rewrite Eq. (3.109) as follows :

$$\boldsymbol{\lambda}_c^{i\leftrightarrow j} = -\mathcal{W}_{cr}^{i\leftrightarrow j} E_r^{j\to i} - \mathcal{W}_{cc}^{i\leftrightarrow j} E_c^{j\to i}, \quad \forall\, i \leftrightarrow j = 1 \ldots N_I \tag{3.110}$$

where $i \leftrightarrow j$ is the number of the interface Γ^{ij} among all the N_I interfaces.

A new projection Boolean matrix $\mathbb{Q}_{\boldsymbol{\lambda}_c}^{i\leftrightarrow j}$ is introduced to extract the double Lagrange multipliers $\boldsymbol{\lambda}_c^{i\leftrightarrow j}$ associated to Γ^{ij} from the full list of λ_c, i.e., $\boldsymbol{\lambda}_c^{i\leftrightarrow j} = \mathbb{Q}_{\boldsymbol{\lambda}_c}^{i\leftrightarrow j}\lambda_c$. A new operator $\mathcal{S}_{\boldsymbol{\lambda}_c}$ is also defined such that [36]

$$\mathcal{S}_{\boldsymbol{\lambda}_c} u = \sum_{i\leftrightarrow j=1}^{N_I} \mathbb{Q}_{\boldsymbol{\lambda}_c}^{i\leftrightarrow j^T} u^{i\to j} = \sum_{i\leftrightarrow j=1}^{N_I} \mathbb{Q}_{\boldsymbol{\lambda}_c}^{i\leftrightarrow j^T} u^{j\to i} \tag{3.111}$$

This operator is an analogue of the two operators presented before (Eqs. 3.91 and 3.94). The new operator is here to assemble all the contributions from every interface in order to create the last part of the Interface Problem, that is the one, related to the set of $\boldsymbol{\lambda}_c$. By applying this new operator to (Eq. 3.110), we obtain

$$-\mathrm{F}_{\boldsymbol{\lambda}_c \lambda_r}\lambda_r - \mathrm{F}_{\boldsymbol{\lambda}_c \mathbf{E}_c}\mathbf{E}_c + \mathrm{F}_{\boldsymbol{\lambda}_c \boldsymbol{\lambda}_c}\boldsymbol{\lambda}_c = \mathrm{d}_{\boldsymbol{\lambda}_c} \tag{3.112}$$

where

$$\begin{aligned}
\mathrm{F}_{\boldsymbol{\lambda}_c \lambda_r} &= \mathcal{S}_{\boldsymbol{\lambda}_c}\left[\mathcal{W}_{cr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}F_{rr}^{j}\right]\mathbb{Q}_{\lambda_r}^{j}\\
\mathrm{F}_{\boldsymbol{\lambda}_c \mathbf{E}_c} &= \mathcal{S}_{\boldsymbol{\lambda}_c}\left[\mathcal{W}_{cr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}F_{rc}^{j} - \mathcal{W}_{cc}^{i\leftrightarrow j}\mathbb{T}_c^{j\to i}\right]\mathbb{Q}_{\mathbf{E}_c}^{j}\\
\mathrm{F}_{\boldsymbol{\lambda}_c \boldsymbol{\lambda}_c} &= \mathcal{S}_{\boldsymbol{\lambda}_c}\\
\mathrm{d}_{\boldsymbol{\lambda}_c} &= \mathcal{S}_{\boldsymbol{\lambda}_c}\left[\mathcal{W}_{cr}^{i\leftrightarrow j}\mathbb{T}_r^{j\to i}d_r^{j}\right]
\end{aligned}$$

We will obtain as well a small change in the second row of the interface problem (3.105) which is related to the fact that we associate the Lagrange multipliers λ_c to the Robin-type boundary condition. Thus, unlike the classical approach, in the FETI-DPEM2-full method, we have

$$\sum_{i=1}^{N_s} \mathbb{Q}^{i}_{\mathbf{E}_c}{}^{T} \lambda^i_c \neq 0 \tag{3.113}$$

By assembling and summing all the subdomain contributions, we obtain the global corner points related system of equations which is similar to the one from Eq. (3.90) :

$$\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}\,\mathbf{E}_c - \mathrm{F}_{\mathbf{E}_c\lambda_r}\,\lambda_r + \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\,\boldsymbol{\lambda}_c = \mathrm{d}_{\mathbf{E}_c} \tag{3.114}$$

The new matrix $\mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}$ can not be introduced through the operator $\mathcal{S}_{\mathbf{E}_c}$. It is a Boolean projection matrix which puts in correspondence all the global corner Lagrange multipliers $\boldsymbol{\lambda}_c^{i\leftrightarrow j}$ related to the global corner $\mathbf{E}_c^k$ for every $k = 1...N_{\mathbf{E}_c}$. The size of this matrix is equal to $N_{\mathbf{E}_c} \times N_{\boldsymbol{\lambda}_c}$ and its construction is quite simple. In each line (that corresponds to the global index of the corner edge) we put 1 if the given global corner Lagrange multiplier belongs to this corner and 0 if not.

Combining Eq. (3.95) with a new matrix $\mathrm{F}_{\lambda_r\mathbf{E}_c}$ (Eq. 3.108), Eqs. (3.112) and (3.114) we arrive at the Full interface problem of the FETI-DPEM2-full method

$$\begin{bmatrix} \mathrm{F}_{\lambda_r\lambda_r} & -\mathrm{F}_{\lambda_r\mathbf{E}_c} & 0 \\ -\mathrm{F}_{\mathbf{E}_c\lambda_r} & \mathrm{F}_{\mathbf{E}_c\mathbf{E}_c} & \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c} \\ -\mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r} & -\mathrm{F}_{\boldsymbol{\lambda}_c\mathbf{E}_c} & \mathrm{F}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c} \end{bmatrix} \begin{bmatrix} \lambda_r \\ \mathbf{E}_c \\ \boldsymbol{\lambda}_c \end{bmatrix} = \begin{bmatrix} -\mathrm{d}_{\lambda_r} \\ \mathrm{d}_{\mathbf{E}_c} \\ -\mathrm{d}_{\boldsymbol{\lambda}_c} \end{bmatrix} \tag{3.115}$$

But in practice, the Full interface problem is not directly solved but replaced by a linear system with smaller dimension and better conditioning number [43,48,97]. Moreover, we do not need to know the Lagrange multipliers $\boldsymbol{\lambda}_c$ because we are looking directly for the solution $\mathbf{E}_c$. Once these Lagrange multipliers are dropped out after few mathematical transformations, we arrive at the Reduced Interface Problem

$$\boxed{\left(\mathrm{F}_{\lambda_r\lambda_r} + \mathrm{F}_{\lambda_r\mathbf{E}_c}\;\hat{\mathrm{F}}^{-1}_{\mathbf{E}_c\mathbf{E}_c}\hat{\mathrm{F}}_{\mathbf{E}_c\lambda_r}\right)\lambda_r = \hat{\mathrm{d}}_{\lambda_r}} \tag{3.116}$$

where

$$\begin{aligned} \hat{\mathrm{F}}_{\mathbf{E}_c\lambda_r} &= -\mathrm{F}_{\mathbf{E}_c\lambda_r} + \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\;\mathrm{F}^{-1}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c}\;\mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r} \\ \hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c} &= \mathrm{F}_{\mathbf{E}_c\mathbf{E}_c} + \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\;\mathrm{F}^{-1}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c}\;\mathrm{F}_{\boldsymbol{\lambda}_c\mathbf{E}_c} \\ \hat{\mathrm{d}}_{\mathbf{E}_c} &= \mathrm{d}_{\mathbf{E}_c} + \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\;\mathrm{F}^{-1}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c}\;\mathrm{d}_{\boldsymbol{\lambda}_c} \\ \hat{\mathrm{d}}_{\lambda_r} &= -\mathrm{d}_{\lambda_r} - \mathrm{F}_{\lambda_r\mathbf{E}_c}\;\hat{\mathrm{F}}^{-1}_{\mathbf{E}_c\mathbf{E}_c}\hat{\mathrm{d}}_{\mathbf{E}_c} \end{aligned}$$

In the following, unless specified otherwise, the Interface Problem will correspond to this Reduced Interface Problem. We can now easily compute $\mathbf{E}_c$ by solving the following system of equations

$$\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}\mathbf{E}_c = \hat{\mathrm{d}}_{\mathbf{E}_c} - \hat{\mathrm{F}}_{\mathbf{E}_c\lambda_r}\lambda_r \tag{3.117}$$

Finally, replacing $\mathbf{E}_c$ and λ_r in (Eq. 3.88) gives access to E_r^i in each subdomain Ω^i.

3.6.5 Interpretation in terms of boundaries conditions

The difference between the FETI-DPEM2 classic and full methods can also be explained in terms of the boundary conditions [96,111]. In this work we stick to the approach given by Farhat [36,39,42] which consists in :

1. Firstly, constructing a Lagrange functional, as it has been done in Eqs. (3.57) and (3.71) using the desired type of constraints between subdomains.
2. Secondly, looking for the saddle point of the set of KKT conditions
3. And, finally, determining the type of boundary conditions related to each local subdomain.

In order to determine the type of the boundary conditions, we will write the indefinite system of linear equations related to the local subdomain Ω^i coupled with the previously obtained saddle point :

$$\begin{aligned} K^i E^i &= f^i + \int\limits_{\Gamma^i} \frac{1}{\mu_r} \frac{\partial \mathcal{E}_h^i}{\partial n}\, \Psi^i \, dS \\ \tilde{K}^i E^i &= f^i - \mathbb{B}^{i^T} \lambda \end{aligned} \tag{3.118}$$

where $\tilde{K}^i$ is determined as in Eq. (3.78). By comparing these two relations from Eq. (3.118) the role of the Lagrange multipliers as well as the matrix $\mathcal{M}$ becomes clear. For the FETI-DPEM2 full and classical methods, the unknown boundary condition corresponds to the ***Robin-type***. But the difference between these methods should be explained using the "rc"-notation. For both methods we assume at the internal interfaces for all the interface points that the field satisfies a Robin-type boundary condition, given by

$$\frac{1}{\mu_r} \frac{\partial \mathcal{E}^i}{\partial n} + \alpha^i \mathcal{E}^i = \Lambda_r^i, \text{ on } \Gamma_r^i \tag{3.119}$$

$$\{\lambda_r^i\} = \int_{\Gamma^i \cap \Gamma_r^i} (\Lambda_r^i \; \Psi^i - \alpha^i \Psi^i \; \Psi^i) \, dS \tag{3.120}$$

The term $\Gamma^i \cap \Gamma_r^i$ denotes the space located at interface Γ^i only related to the interface points.

The main difference between the classical and full FETI-DPEM2 methods can be explained on the boundary conditions which are associated to the corners. For the FETI-DPEM2 *classical* method, according to [111], the Lagrange multipliers related to the corner points can be associated to the unknown *Neumann-type* boundary conditions, which mathematically can be expressed as follows :

$$\frac{1}{\mu_r} \frac{\partial \mathcal{E}^i}{\partial n} = \Lambda_c^i, \text{ on } \Gamma^i \cap \Gamma_c^i \tag{3.121}$$

$$\{\lambda_c^i\} = \int_{\Gamma^i \cap \Gamma_c^i} (\Lambda_c^i \; \Psi^i) \, dS \tag{3.122}$$

The term $\Gamma^i \cap \Gamma_c^i$ denotes the space located at interface Γ^i only related to the corner points.

Unlike the classical FETI-DPEM2 method, the one proposed in this work is based on the idea of setting the *Robin-type* boundary conditions for all the points, including the corners. Therefore, instead of assuming (Eqs. 3.121 and 3.122) at the corner points we assume

$$\frac{1}{\mu_r} \frac{\partial \mathcal{E}^i}{\partial n} + \alpha^i \mathcal{E}^i = \Lambda_c^i, \text{ on } \Gamma^i \cap \Gamma_c^i \tag{3.123}$$

$$\{\lambda^i\} = \int_{\Gamma^i \cap \Gamma_c^i} (\Lambda_c^i \; \Psi^i - \alpha^i \Psi^i \; \Psi^i) \, dS \tag{3.124}$$

Then, we determinate the common unknown $\Lambda^i = \{\Lambda_r^i, \Lambda_c^i\}$, given as :

$$\frac{1}{\mu_r} \frac{\partial \mathcal{E}^i}{\partial n} + \alpha^i \mathcal{E}^i = \Lambda^i, \text{ on } \Gamma^i \tag{3.125}$$

The matrix $\mathcal{M}^{i\leftrightarrow j}$ and, therefore, $\mathcal{W}^{i\leftrightarrow j}$ which are participating in the Interface Problem (3.115) is constructed according to the weak form of the Robin-type boundary condition presented above for the subdomain Ω^i, as follows :

$$\mathcal{M}^{i\leftrightarrow j} = \mathbf{M}_{\Gamma^{ij}}(\alpha) = \int_{\Gamma^{ij}} \alpha \; \Psi^i \; {\Psi^i}^T \; dS \tag{3.126}$$

By writing the similar equation for the subdomain Ω^j we obtain :

$$\frac{1}{\mu_r}\frac{\partial \mathcal{E}^j}{\partial n} + \alpha^j \mathcal{E}^j = \Lambda^j, \text{ on } \Gamma^j \tag{3.127}$$

By adding then Eq. (3.125) to Eq. (3.127) and applying the continuity of the electric and magnetic field, we obtain :

$$\Lambda^i_j + \Lambda^j_i = (\alpha^i + \alpha^j)\mathcal{E}^i_j \tag{3.128}$$

$$\Lambda^i_j + \Lambda^j_i = (\alpha^i + \alpha^j)\mathcal{E}^j_i \tag{3.129}$$

According to [111] for the classical FETI-DPEM2 method this relation takes place only for the interface points, but not for the corners. We proposed in this work to extend this relation even for the corner points. This is the key of the FETI-DPEM2-full method.

3.7 Domain decomposition in 3D

3.7.1 Finite Element system

Like in the two-dimensional configuration, we start with the following 3D Helmholtz equation in a bounded domain Ω which can contain inhomogeneities :

$$\nabla \times \left(\frac{1}{\mu_r}\nabla \times \vec{\mathcal{E}}\right) - k_0^2 \varepsilon_r \vec{\mathcal{E}} = jk_0 Z_0 \vec{J} \text{ in } \Omega \tag{3.130}$$

where $\vec{\mathcal{E}} = (\mathcal{E}_x(\vec{r}), \mathcal{E}_y(\vec{r}), \mathcal{E}_z(\vec{r}))^T$ is the total or the incident field with a time dependency in $\exp(-j\omega t)$ originated in the current distribution $\vec{J}$. This incomplete form, from the mathematical point of view, is complemented then by the radiation boundary condition (Eq. 3.131) on the *external* boundary Σ, in order to provide the unicity of the solution in the bounded domain Ω.

$$\vec{n} \times \left(\frac{1}{\mu_r}\nabla \times \vec{\mathcal{E}}\right) + jk_0 \, \vec{n} \times \vec{n} \times \vec{\mathcal{E}} = 0 \quad \text{on } \Sigma \tag{3.131}$$

As we have already done for the 2D case (Section 3.2), we then represent the discretized field $\vec{\mathcal{E}}_h$ as a linear combination of basis functions in the finite-dimensional subspace $V^h(\Omega) \subset H^1_\theta(\Omega)$. After finite-element discretization we obtain the integral equation in weak form for the 3D case :

$$\begin{aligned}\sum_{j=1}^{n^j} \left[\int_\Omega \frac{1}{\mu_r}(\nabla \times \vec{\psi}_j)\cdot(\nabla \times \vec{\psi}_i) \; d\Omega - \int_\Omega k_0^2 \varepsilon_r \; \vec{\psi}_j \cdot \vec{\psi}_i \; d\Omega \; + jk_0 \int_\Sigma (\vec{n} \times \vec{\psi}_j)\cdot(\vec{n}\times\vec{\psi}_i) dS \right] q_j = \\ - \int_\Omega jk_0 Z_0 \; \vec{J} \cdot \vec{\psi}_i \; d\Omega, \qquad i = 1, ..., n_i\end{aligned} \tag{3.132}$$

where the set of vectors $\vec{\psi} = \vec{\psi}(\vec{r})$ are basis functions which belong to $V^h(\Omega)$. The integral equation above can then be written in a matrix form, which is similar to Eq. (3.15), as follows :

$$\mathbf{KE} = \mathbf{f} \tag{3.133}$$

where the FEM problem matrix and vector are defined as :

$$\mathbf{K} = \mathbf{G} - \mathbf{M}_{vol} + \mathbf{M}_{\Sigma}(jk_0) \tag{3.134}$$

$$\mathbf{G} = \int_{\Omega} \frac{1}{\mu_r} (\nabla \times \Psi) \cdot (\nabla \times \Psi)^T \, d\Omega \tag{3.135}$$

$$\mathbf{M}_{vol} = \int_{\Omega} k_0^2 \varepsilon_r \; \Psi \cdot \Psi^T \, d\Omega \tag{3.136}$$

$$\mathbf{M}_{\Sigma}(jk_0) = \int_{\Sigma} jk_0 \varepsilon_r \; (\vec{n} \times \Psi) \cdot (\vec{n} \times \Psi)^T \, dS \tag{3.137}$$

$$\mathbf{f} = -\int_{\Omega} jk_0 Z_0 \; J_h \cdot \Psi \, d\Omega \tag{3.138}$$

The term Ψ denotes a column vector containing the first-order linear vector basis functions $\vec{\psi}_i,\ i = 1, ..., n_i$.

As we have a full liberty for choosing the finite-elements, we can in particular select the tetrahedras (Figure 3.4) as generic mesh elements. In this case every element of volume contains :

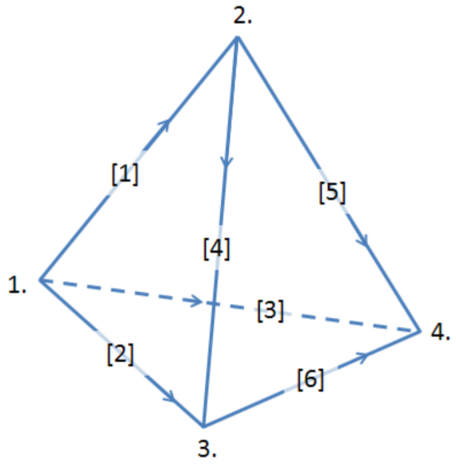


Figure 3.4 – The finite-element cell.

- 6 numbered edges (i = [1], [2], [3], [4], [5], [6])
- 4 numbered points (j = 1., 2., 3., 4.)
- 4 triangle-faces (Table 3.2)

N	p_1	p_2	p_3	e_1	e_2	e_3
1	1	2	3	1	2	4
2	1	2	4	1	3	5
3	1	3	4	2	3	6
4	2	3	4	4	5	6

Table 3.2 – The numbering (p - by points and e - by edges) of 4 face-triangles in each tetrahedra.

Using the scalar finite elements, sometimes we observe the occurrence of non-physical solutions as well as the inconvenience of imposing boundary conditions at material interfaces and conducting surfaces [16]. For these reasons, more than 50 years ago Whitney [114] and later on Nedelec [115] proposed to use an other type of elements which are called vector basis or vector elements. For the 3D case in this work we are going to use the vector basis elements of the 1st (linear) order. In this case the size of the problem corresponds to the global number of edges in the domain Ω.

3.7.2 FETI-DPEM2-full method in 3D

As for the 2D case (Section 3.6.4.2), we assume that the domain Ω is divided into a set of N_s non-overlapping domains :

$$\Omega = \bigcup_{i=1}^{N_s} \Omega^i, \qquad \Omega^i \cap \Omega^j = \emptyset, \, \forall i \neq j \tag{3.139}$$

where the index i denotes the subdomain number. In terms of the formulas there is no difference between the 2D and the 3D cases. But now all the unknown Lagrange multipliers λ represent the Robin-type boundary condition for the 3D case (Eq. 3.140).

$$\vec{n} \times \left(\frac{1}{\mu_r} \nabla \times \mathcal{E}^i \right) + \alpha^i \, \vec{n} \times \vec{n} \times \mathcal{E}^i = \Lambda^i, \text{ on } \partial\Gamma^i \tag{3.140}$$

The choice of the mathematical parameter α^i for the 3D case will be discussed later on, in Section (5.3.5.1). Taking into account the aspects of imposing the Robin-type boundary condition previously described in Section (3.2.1), we obtain the system of linear equations associated with the local electromagnetic problem in the subdomain Ω^i :

$$K^i E^i = f^i + \int_{\Gamma^i} \Lambda^i \, \Psi^i \, dS \tag{3.141}$$

In order to make this section self-sufficient, we would like to notice, that after having applied the "rc"-notation to the equation above and to the matrix form of the transmission conditions (Eq. 3.92), we obtain absolutely the same full interface problem (Eq. 3.115) and the reduced one (Eq. 3.116) as in the 2D case. Thus, we will not detail further more the mathematics associated to the implementation of the FETI-DPEM2-full method in a 3D configuration as there are fully similar to the 2D case.

3.8 How to test the FETI-methods

The FETI-DPEM2 method becomes more complex in terms of geometry when we are dealing with arbitrary partition. That is why we would like to dedicate this chapter to explain how it is possible to guarantee almost all the steps of the method. In this work we have used two main ways to test the FETI-DPEM2-full method :

- Tests based on the use of analytical solutions
- Tests based on the use of previously computed numerical solutions

3.8.1 Algorithm flowchart

Here we will present the main steps of the algorithm :

```
 1. Read input data
 2. Creation of the mesh discretization
 3. Mesh partition
 4. Creation of the list of the interface-triangles
 5. Creation of the list of $\lambda_r, \mathbf{E}_c$ and $\lambda_c$
 6. Printing the mesh
 7. Creation of the geometry matrices (3.6.3)
 8. Creation of the global rhs
 9. Division of the global rhs
10. do {number of partitions}
        Creation of $K$ in "rc"-notations
        Inversion of $K_{rr}$
        Make all the transformations (3.6.4.2)
        Make a contribution into Eq. (3.115)
    end
11. do {number of sources}
        solve the Interface Problem (3.115)
        do {number of partitions}
            creation of $K$ in "rc"-notations
            Solution for $E_r$ (3.87)
        end
    end
12. Store field
```

3.8.2 Test with analytical solutions

Two approaches presented in Section (3.8) help us to test and to debug the method. We would like to explain the idea of these two approaches for the 2D case, because for 2D problems we have a clear vision of the ongoing process and geometry. Let us imagine that we are dealing with an elliptic problem (3.142-3.143) in the same bounded domain Ω, as in Figure (3.1).

$$\mathrm{div}\,(\lambda \mathrm{grad}\, u) + \gamma\, u = -f, \text{ in } \Omega \tag{3.142}$$

$$\lambda \frac{\partial u}{\partial n} + \alpha\, u = \Lambda, \text{ on } \Sigma \tag{3.143}$$

This boundary-problem can describe the stationary distribution of the temperature u in the domain Ω, if λ is the thermal conductivity and $f - \gamma u$ is the density of the heat sources. But in the frame of this test case, we consider this equation as a mathematical equation without any physical sense. In particular, we can "search for" the *exact* solution

$$u = 2x^3 + 8y^2 \sin(x) \tag{3.144}$$

By setting the coefficients λ, γ and α we can find the rest of the unknown parameters for this problem - f and Λ :

$$f = -\left((8y^2 - 16)\sin(x) - 12x\right) + jk_0\left(2x^3 + 8y^2\sin(x)\right), \quad \text{with } \lambda = 1 \text{ and } \gamma = jk_0$$
$$\Lambda = \frac{\partial u}{\partial n} + jk_0 u, \quad \text{with } \alpha = jk_0$$

By setting these parameters we guarantee the unicity of the solution (3.144) in the bounded domain Ω. To start with, we set the size of Ω (Figure 3.1) equal to $[-4, 4] \times [-4, 4]$ and then we calculate the elliptic boundary-problem (3.142-3.143) inside the given domain with the classical FEM method. The obtained solution u is presented in Figure (3.5).

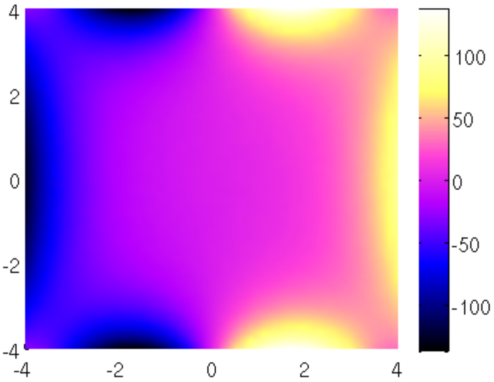


Figure 3.5 – The exact solution u in a bounded domain Ω

Now, following the idea of the FETI-DPEM2 method we divide our domain Ω into 4 subdomains with respect to the center-point $(0, 0)$. In other words, we represent the field u as a union of the local fields in each subdomain $u = \cap\, u^i,\ i = 1..4$. Then we calculate the field u^i as a solution of the local bounded problem (3.145-3.146) in each subdomain Ω^i using still the classical FEM method.

$$\operatorname{div}\left(\lambda \operatorname{grad} u^i\right) + \gamma\, u^i = -f^i, \text{ in } \Omega \tag{3.145}$$

$$\begin{aligned} \lambda\frac{\partial u^i}{\partial n} + \alpha\, u^i &= \Lambda^i, \text{ on } \Sigma^i \\ \lambda\frac{\partial u^i}{\partial n} + \alpha\, u^i &= \Lambda^i, \text{ on } \Gamma^i \end{aligned} \tag{3.146}$$

The results of the classical FEM method are presented in Figure (3.6). Note that *all* the values Λ^i on the external (Σ^i) and internal (Γ^i) interfaces can be calculated analytically. It leads us to the main idea of this approach, that is we can use the values $\lambda^{i\to j}$ from the Boundary conditions (3.143) to test the transmission condition of the FETI-DPEM2-full method in the matrix form :

$$\begin{bmatrix} \lambda_r^{i\to j} \\ \lambda_c^{i\to j} \end{bmatrix} + \begin{bmatrix} \lambda_r^{j\to i} \\ \lambda_c^{j\to i} \end{bmatrix} = -\begin{bmatrix} \mathrm{W}_{rr}^{i\leftrightarrow j} & \mathrm{W}_{rc}^{i\leftrightarrow j} \\ \mathrm{W}_{cr}^{i\leftrightarrow j} & \mathrm{W}_{cc}^{i\leftrightarrow j} \end{bmatrix} \begin{bmatrix} E_r^{i\leftrightarrow j} \\ E_c^{i\leftrightarrow j} \end{bmatrix}$$

As it was noticed before, we know the *exact* solution (3.144) for this problem. That means that we also know the values of $E_r^{i\leftrightarrow j}$ and $E_c^{i\leftrightarrow j}$ everywhere in the domain. And now if we calculate the relative

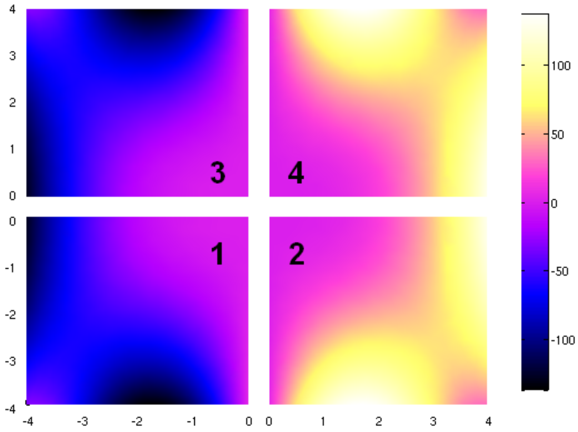


FIGURE 3.6 – The solutions u_i obtained in each subdomain with classical FEM method

error between two parts of this equation (left and right) for *any* interface between every Ω^i and Ω^j we will obtain an error of the order of 10^{-14} to 10^{-13}. This idea gives us an opportunity to control and to trace the error on λ.

But the main advantage of this approach is that from the beginning until the end of the method we can control errors of constructions of *all* the geometrical(projection) matrices. In our work we paid main attention on the construction of the Full Interface Problem (3.115). It is evident that if we know the values of λ_r, $\boldsymbol{\lambda}_c$ and $\mathbf{E}_c$ we can check step by step all the equations.

1. $\mathrm{F}_{\lambda_r\lambda_r} \cdot \lambda_r + \mathrm{F}_{\lambda_r\mathbf{E}_c} \cdot \mathbf{E}_c$ by comparing with d_{λ_r}
2. $\mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r} \cdot \lambda_r + \mathrm{F}_{\boldsymbol{\lambda}_c\mathbf{E}_c} \cdot \mathbf{E}_c + \mathrm{F}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c} \cdot \boldsymbol{\lambda}_c$ by comparing with $\mathrm{d}_{\boldsymbol{\lambda}_c}$
3. $\mathrm{F}_{\mathbf{E}_c\lambda_r} \cdot \lambda_r + \mathrm{F}_{\mathbf{E}_c\mathbf{E}_c} \cdot \mathbf{E}_c + \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c} \cdot \boldsymbol{\lambda}_c$ by comparing with $\mathrm{d}_{\mathbf{E}_c}$

and if the comparisons do not provide an error of the order 10^{-14} to 10^{-13} then there is an error of construction in the projection matrices. It is an easy way to go deeper and to check each of the *matrices* F.

3.8.3 Test with numerical solutions

The second approach is similar. But instead of the *exact* solution we take the FEM-one. In our work we used this approach to test and to debug both the two- and three-dimensional electromagnetic problems. For example we have a numerical solution $\mathbf{E}$ of a physical problem obtained with the classical FEM method (Figure 3.7). It is obvious, that we can not calculate any more the electric field in each subdomain, because of the lack of internal boundary conditions. Indeed, we do not know the values Λ^i for each subdomain Ω^i, unlike the previous mathematical approach (Section 3.8.2) where these values could be calculated analytically.

In the framework of this test approach, we construct matrices $M^{i\to j}$ of the Robin-type boundary conditions applied to the internal interfaces $\Gamma^{i\to j}$ with the help of the electric field obtained with the FEM method which serves us as an *exact* solution. For example, if we use the same domain decomposition in 4 subdomains, as previously, then we can get the values λ on each internal interface as shown in Figure (3.7) with red color lines.

All the following steps are the same as for the previous approach detailed in Section (3.8.2). As we know λ, the rhs and the solution obtained with FEM, we also know all the needed information to test the method as we did previously.

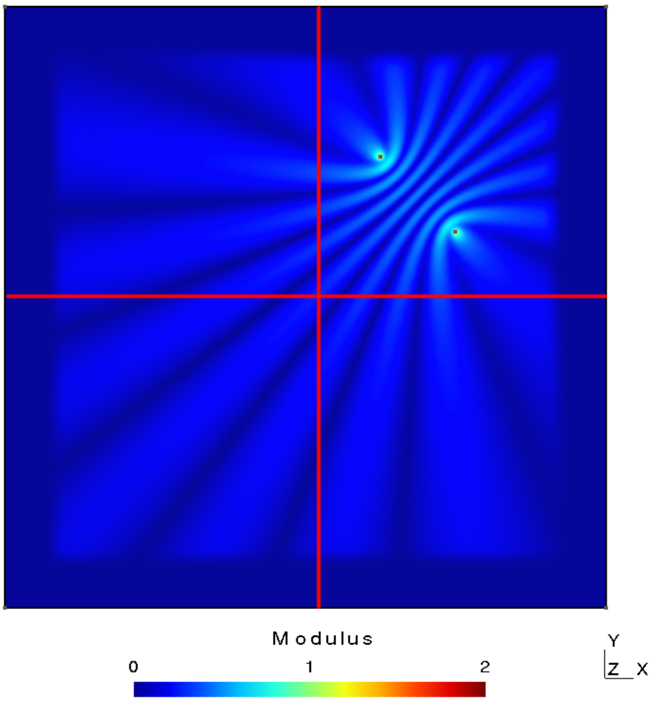


Figure 3.7 – Modulus of an electric field in the bounded domain Ω, obtained with the FEM method. Red lines denote the domain decomposition, where we can calculate the values of Lagrange multipliers λ

3.9 Source term implementation

In this section we would like to explain the way for constructing the right-hand side vector. We will consider two specific test-cases when the source point is located on

- a boundary between the subdomains
- a corner point

The first test case is schematically presented in Figure (3.8).

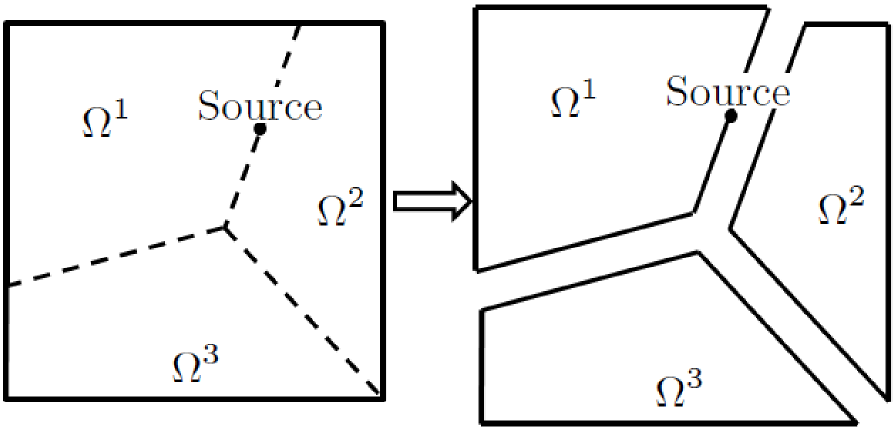


Figure 3.8 – One source in the domain and the domain decomposition in this case

The main rule in this case is that the summation of the values of the local RHS-vectors in Ω^1 and Ω^2 must be equal to the value of the global FEM-RHS vector in this point. In our approach we place the full value to one domain (for example to the first one), but the value of the second local RHS vector in this point is set to be equal to 0. In this case we do not need to calculate the column of K_{rr}^{-1} in this point.

We use the same idea when the source point is placed on a corner point. In our approach we set to non-zero only one out of all the local corner points. Schematically :

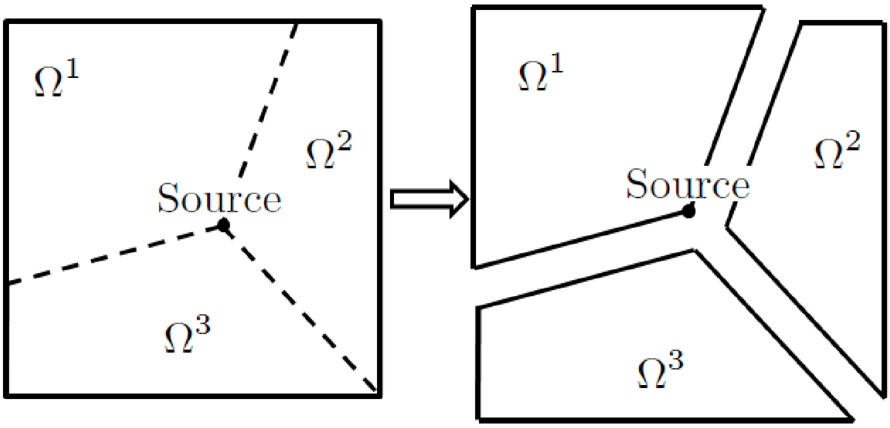


Figure 3.9 – Source located in a corner point, domain decomposition and the structure of the local RHS-vectors in this case.

3.10 Conclusion

Starting from the classical Finite Element formulation we introduced the principal Domain Decomposition idea as well as some of the Domain Decomposition techniques. The principal methods based on the Lagrangian theory have been presented here in order to introduce the FETI-DPEM2 method which distinguishes itself through its excellent convergence results, scalability and the parallel implementation. In this chapter we introduced a modification of this method, so-called, FETI-DPEM2-full method, and showed the main difference between these two methods. We considered both the two- and three- dimensional formulations and have discussed at the end some implementation issues.

As a next step, we are going to study the efficiency of the proposed method and compare the results obtained with the FEM and the FETI classical methods.

Chapitre 4

2D Forward Problems

Contents

4.1 Introduction

In the previous chapters, we have considered the mathematical and physical formulations of the different electromagnetic problems, as well as the numerical algorithms which provide a solution in an efficient way. In particular, Finite Element Tearing and Interconnecting method seems very robust when one is dealing with arbitrary mesh partitions. The general principle of FETI methods is to divide the entire computational domain into non-overlapping subdomains, in order to construct the Interface problem, and to calculate the solution in each subdomain independently. In the previous chapter, we proposed an extension of the dual-primal techniques by enforcing a Robin-type boundary condition (with two Lagrange multipliers) not only on the edges related to the internal interfaces but also to the ones related to the corner nodes. Indeed, in the methods already proposed (for example FETI-DPEM2 [43]), only one Lagrange multiplier was applied to these corner nodes, yielding a local Neumann-type boundary condition.

We will demonstrate in the following chapter that the proposed extension more efficiently simulates the scattering from two-dimensional objects made of either isotropic or anisotropic materials. We have focus our attention so far on two-dimensional configurations as it allows easiest comparisons between different methods. We will test the FETI-DPEM2-full method in order to study :

- Its computational limits
- The influence of the various physical parameters
- The influence of the transmission conditions at the various interfaces
- The influence of Anisotropic media [116] and Perfectly matched layer (PML) [111]

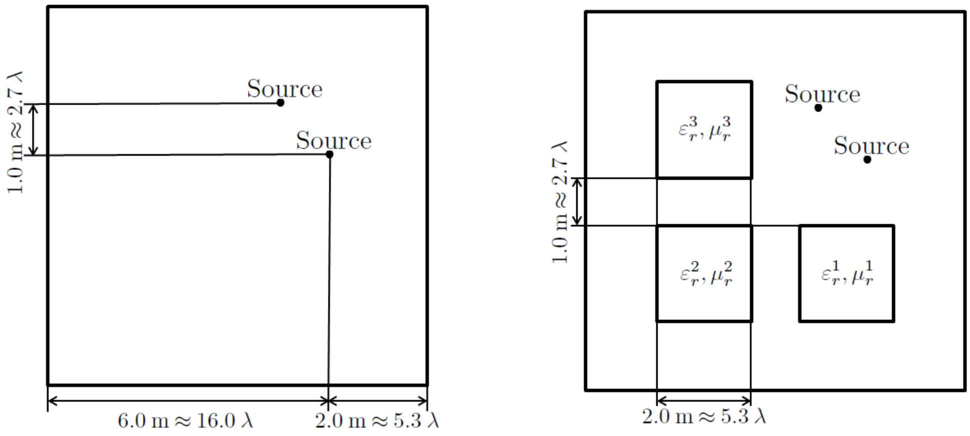
In order to verify the obtained results, during all the calculations, we will compare them with the results obtained with the classical FEM and FETI-DPEM2 methods with a criteria based on the L^2-norm discrepancy.

We do not pursue the goal of comparing these two methods in terms of convergence results, nor in time/memory efficiency. Thus, during all of the numerical experiments in the two-dimensional case, we use the frontal solver Mumps [117] based on the **LU decomposition**. We also set the classical parameter α^i from the Robin-type boundary condition to be equal to jk_0 for all the calculations, except the ones aiming to study the influence of this parameter.

For the two-dimensional configuration, the domain is discretized thanks to a free unstructured mesh generator GMSH [118]. This global mesh is inserted into a home-made mesh partitioner which provides, for each subdomain, the structure of the local mesh as well as the boundaries and the corner nodes lists between subdomains. This partitioner intensively uses subroutines provided by METIS [54]. The factorization of each sub-matrices is performed thanks to the direct sparse solver MUMPS [117, 119] and stored during the resolution of the interface problem.

4.2 Computational limits of the FETI-DPEM2-full method

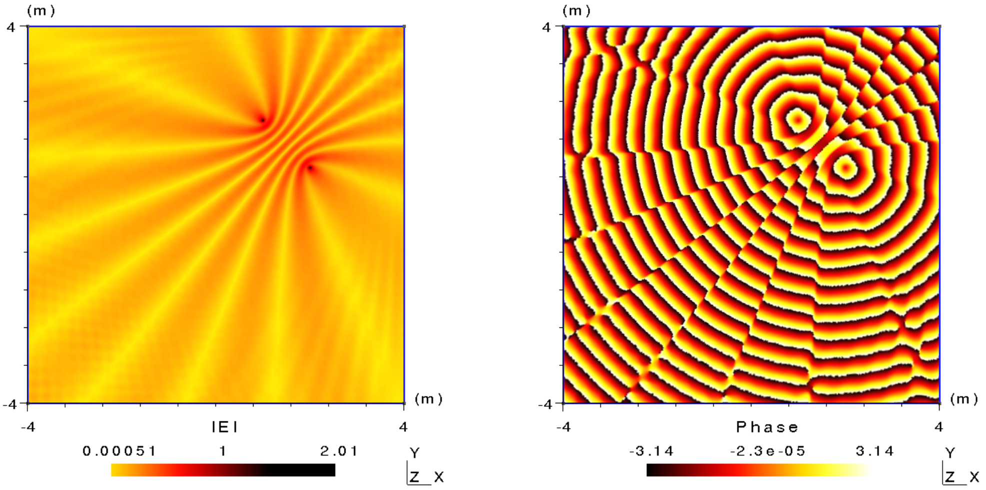
In a domain filled with air ($\varepsilon_0 = 8.85 \cdot 10^{-12}\mathrm{F} \cdot \mathrm{m}^{-1}, \mu_0 = 4\pi \cdot 10^{-7}\mathrm{H} \cdot \mathrm{m}^{-1}$) whose size is $[8 \times 8]\mathrm{m}^2$ two sources are radiating at the same frequency $f_1 = f_2 = 800\mathrm{MHz}$ (Figure 4.1a).



(a) Configuration for the incident field
(b) Configuration for the total field

FIGURE 4.1 – (a) Positions of the sources in the area filled with air and (b) 3 scatterers inside the domain with different relative permittivities.

As the wavelength is $\lambda \approx 0.37\mathrm{m}$, the size of the domain is of $\approx 21\lambda \times 21\lambda$. For the reasons, previously discussed in Section (2.2.2) we set the radiation boundary conditions on the external boundary Σ. The modulus and the phase of the electric field obtained with the FEM classical method are presented in Figure (4.2).



(a) Modulus
(b) Phase (rad)

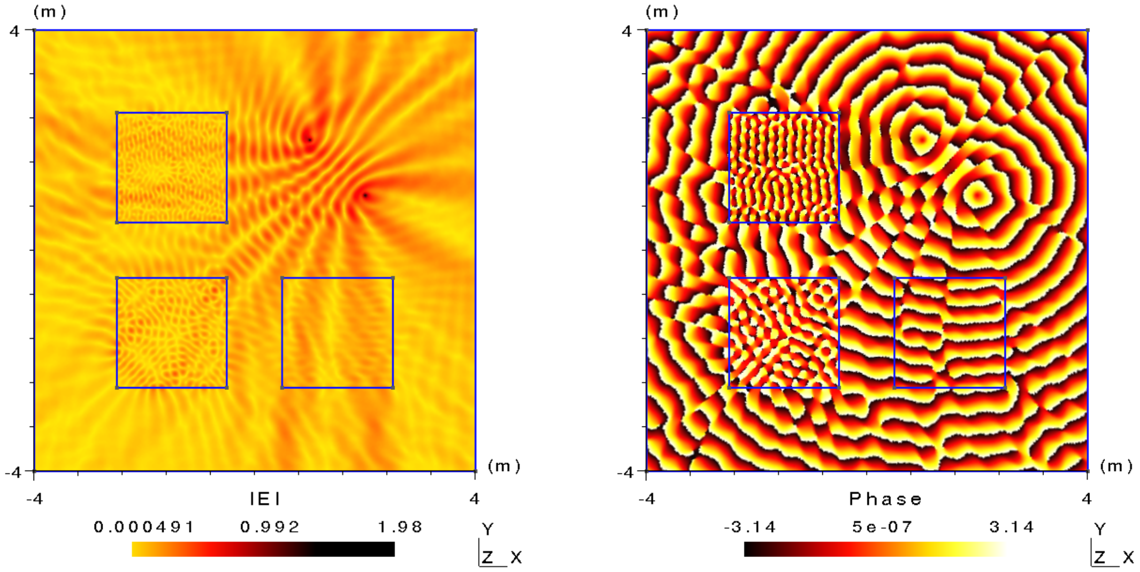
FIGURE 4.2 – Maps of the modulus and the phase of the electric field in free space, obtained with the classical FEM method, for the configuration described in Figure (4.1a).

As a next step, we introduce 3 scatterers with different relative permittivities $\varepsilon_r^1 = 1.5, \varepsilon_r^2 = 3.0$ and $\varepsilon_r^3 = 5.0$ (Figure 4.1b). To obtain physically accurate results we have to discretize the domain with the step l given by Eq. (4.1)

$$l = \lambda / \left(\kappa \sqrt{\varepsilon_r}\right) \tag{4.1}$$

where λ corresponds to the wavelength and the numerical parameter κ represents the number of points per wavelength which is normally varying from 5 to 20.

Previously the relative permittivity ε_r was set equal to 1 everywhere. But since we put scatterers inside the domain we now need to refine the mesh with respect to the relative permittivity of the scatterers. Taking this into account we refine the mesh from the previous case and calculate the given problem with the classical FEM method. The modulus and the phase of the electric field corresponding to the configuration of Figure (4.1b) and obtained with the FEM method are presented in Figure (4.3).



(a) Modulus

(b) Phase (rad)

FIGURE 4.3 – Maps of the modulus and the phase of the electric field with three scatterers inside the domain, computed with the classical FEM method, for the configuration described in Figure (4.1b).

The same configuration has also been computed with the FETI-DPEM2-full method using a partitioning into 5 subdomains.

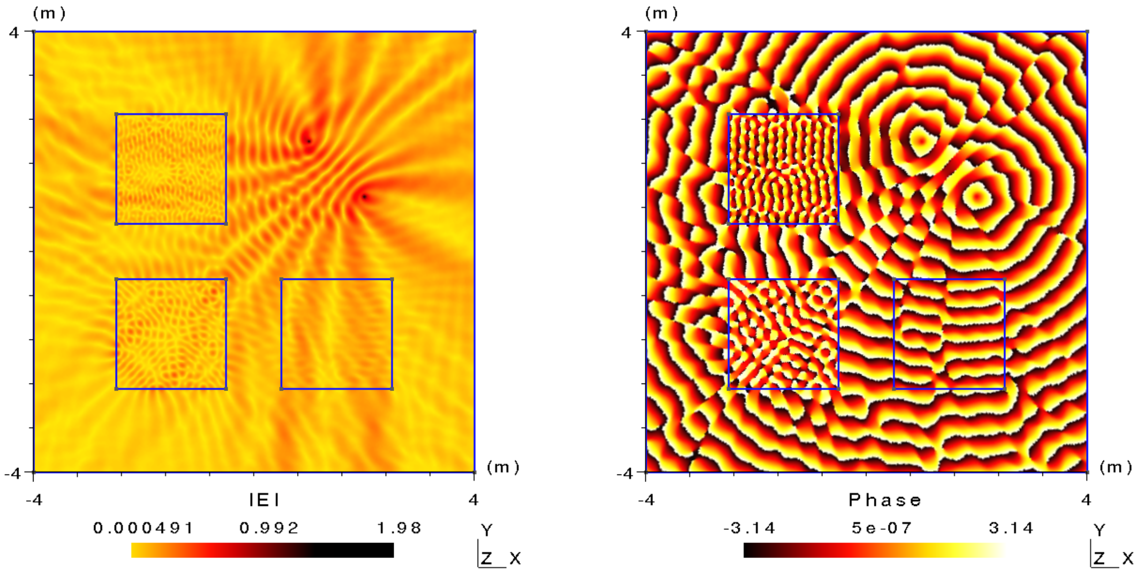
Comparing these two sets of figures, we can not visually see the difference between the results of the FEM-classic and the FETI-DPEM2-full methods. In order to compare the obtained results not only visually, but also in a quantitative way, the following criteria (Eq. 4.2) has been introduced.

$$L^2 - \text{error} = \frac{||\mathbf{E}_1 - \mathbf{E}_2||^2}{||\mathbf{E}_1||^2} \tag{4.2}$$

where $\mathbf{E}_1$ is the value of the electric field calculated with the FEM-classic method and $\mathbf{E}_2$ is the value of the electric field calculated with the FETI-DPEM2 classic or the FETI-DPEM2-full methods.

In Table (4.1) the results of comparison with the different partitioning are presented. We can see a huge difference between results obtained with the FETI-DPEM2 classical and the FETI-DPEM2-full methods in terms of relative L^2-error. We also can conclude from this table, that the results of the FETI-DPEM2-full method in terms of relative error **do not** depend on the number of subdomains or the way the partitioning is performed.Indeed, the partitioning, done with METIS is arbitrary (Figure 4.5) and does not follow explicitly the geometry of the scatterers.

The logical question now would be : How can we explain such valuable difference in Table (4.1) between the FETI methods and where does this misfit come from ? Previously, in Section (3.6), we discussed the principal modifications that we proposed in order to modify the classical FETI method. It has been shown that the main difference consists in a special treatment of the corner DOFs.



(a) Modulus

(b) Phase (rad)

Figure 4.4 – Maps of the modulus and the phase of the electric field with three scatterers inside, computed with the FETI-DPEM-full method, for the configuration described in Figure (4.1b).

N_s	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2-full
5	8.2052E-003	2.6031E-012
10	9.3415E-003	2.4195E-012
15	2.0310E-002	5.6008E-012
20	5.0769E-002	7.8799E-012
25	4.6738E-002	1.3348E-011
30	4.8722E-002	1.3182E-011
35	5.4464E-002	1.3785E-011
40	6.6617E-002	2.1932E-011
45	9.7185E-002	1.2690E-011
70	0.1185	4.7477E-011
80	0.1539	1.9958E-011

Table 4.1 – The relative L^2-error of the classical and full FETI-DPEM2 methods applied to the configuration presented in Figure (4.1a) with different partitionings.

In order to answer the two questions above, let us consider the worst case in terms of relative error, i.e. the partitioning into 80 subdomains (Figure 4.5). The number of the corner points is around 140 inside the chosen domain decomposition. Thus, we plot at various corner points the values of the electric field obtained with the FEM classical, FETI-DPEM2 classical and FETI-DPEM2-full methods (Figure 4.6). From this figure we can see the relatively large difference between the values obtained with the FETI-DPEM2 classical method as compared to the FEM method, while we almost do not have any difference between the FETI-DPEM2-full method and the FEM classical method. This difference leads to errors in the computed Lagrange multipliers λ_r, and errors on the estimated field E_r. This explains the discrepancy observed between the two methods in Table (4.1).

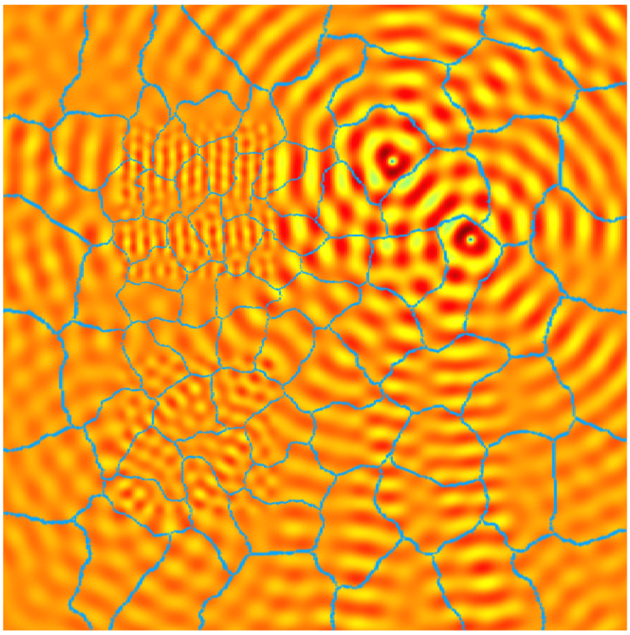


Figure 4.5 – Map of the imaginary part of the electric field, showing the way the partitioning has been performed into $N_s = 80$ subdomains, for the configuration described in Figure (4.1b).

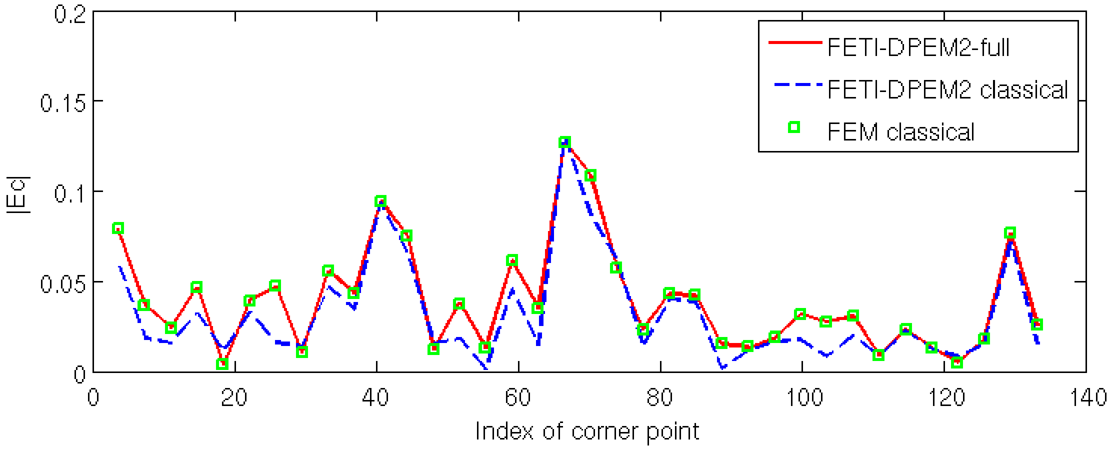
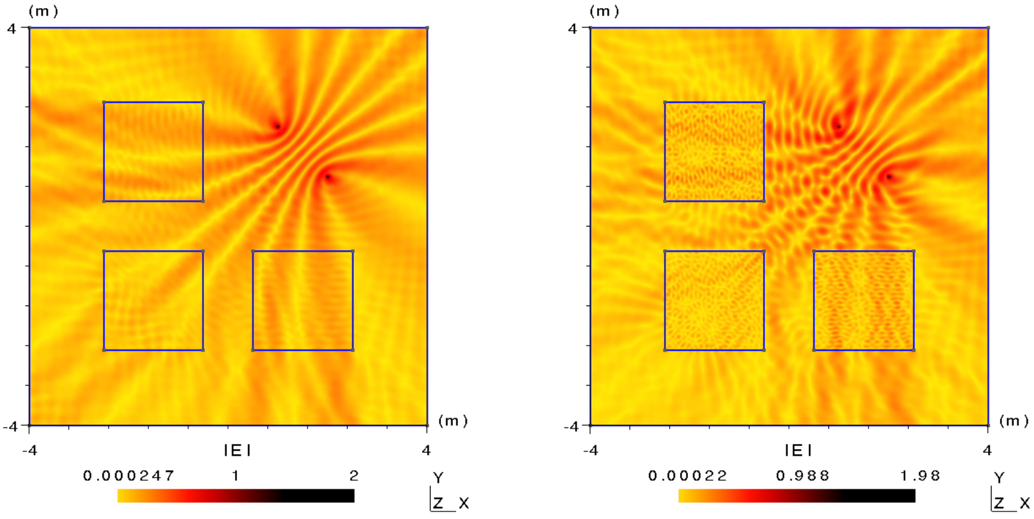


Figure 4.6 – Value of the electrical field extracted at the various corner points obtained with the FEM classical, FETI-DPEM2 classical and FETI-DPEM2-full methods.

4.3 Influence of the physical parameters

The purpose of the second test is to study the influence of the different physical parameters. For example, from Figure (4.7) we can see the variations of the electric field when various relative permittivities of the scatterers are used, in particular the difference between electric fields obtained with the FETI-DPEM2-full method applied for the partitioning into 80 subdomains, when all the scatterers inside have the permittivity 1.5 and 5.0.



(a) $\varepsilon_r = 1.5$ (b) $\varepsilon_r = 5$

Figure 4.7 – Maps of the modulus of the electric field obtained with the FETI-DPEM2-full method. The physical properties of the scatterers are set to (a) $\varepsilon_r^1 = \varepsilon_r^2 = \varepsilon_r^3 = 1.5$ and (b) $\varepsilon_r^1 = \varepsilon_r^2 = \varepsilon_r^3 = 5.0$.

Following the same idea as in Section (4.2), we play with the relative permittivity of the scatterers (Table 4.3) and compare the relative error of the FETI-DPEM2 classical method with the one of the FETI-DPEM2-full method.

ε_r	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2-full
1.5	2.3461E-002	1.9562E-012
2.0	2.9300E-002	2.3079E-012
2.5	2.6335E-002	6.2068E-012
3.0	2.6232E-002	1.0380E-011
3.5	3.4079E-002	8.5070E-011
4.0	3.4153E-002	1.4048E-011
4.5	2.5168E-002	3.9277E-011
5.0	2.5871E-002	1.1491E-011

Table 4.2 – The relative error with various relative permittivities of the scatterers.

From Table (4.3) we can draw two conclusions :

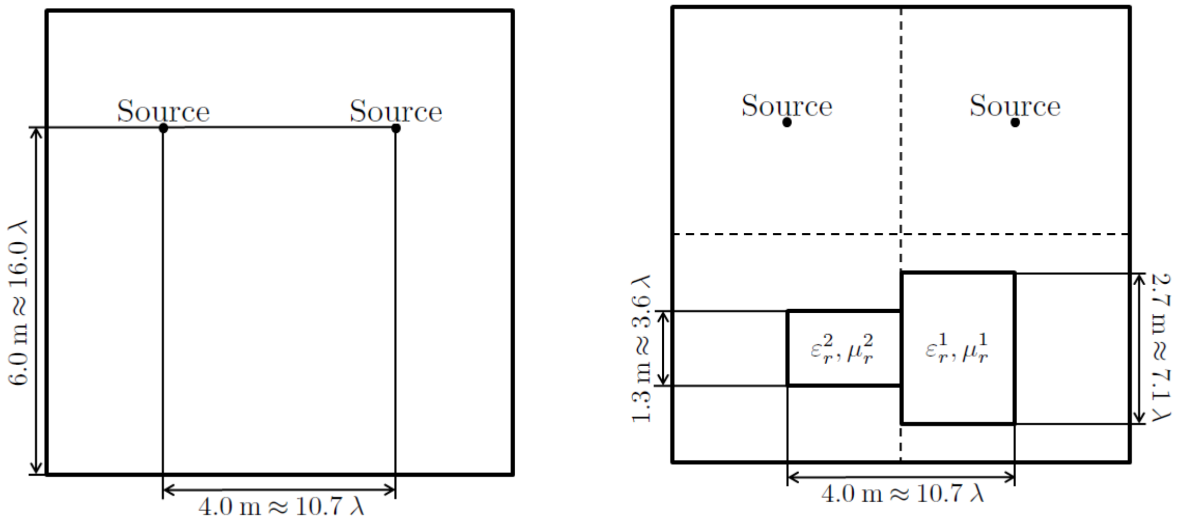
— The results of the FETI-DPEM2 methods do not depend on the choice of the physical parameters.

— Like in the previous case presented in Section (4.2), the FETI-DPEM2-full method provides more precise results in terms of relative L^2-error than its classical approach.

The fact that the results of the FETI-DPEM2-full method do not depend on the choice of the physical parameters means that this method is able to treat a large range of physical applications with great confidence.

4.4 Influence of the transmission conditions

The purpose of the third test is to analyse the influence of the mathematical parameter α^i which arises in the Robin-type boundary condition given in Eq. (3.125). For this, we will consider the following problem : inside the area filled with air as previously shown, there are 2 sources (Figure 4.8a) which radiate at the same frequency $f = 800$ MHz.



(a) Configuration for the incident field

(b) Configuration for the total field

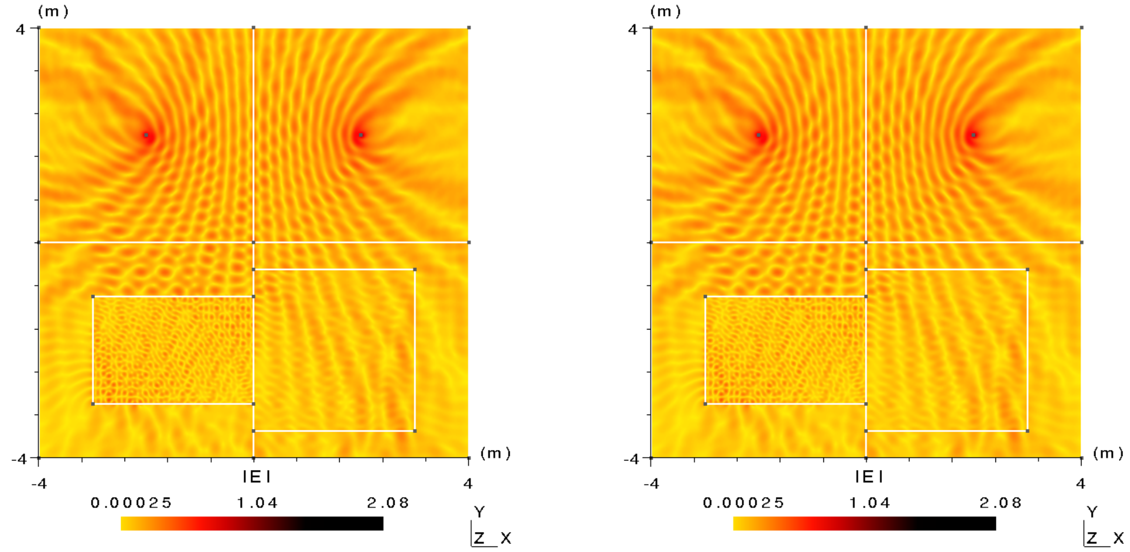
FIGURE 4.8 – (a) Positions of the sources in the area filled with air and (b) 2 scatterers inside the domain with different relative permittivities, as well as a special partitioning of the domain Ω.

As a next step we place 2 scatterers with different values of relative permittivity $\varepsilon_r^1 = 1.5, \varepsilon_r^2 = 5$ and divide this domain into 4 subdomains such that the boundary between the 2 subdomains corresponds at the same time to the boundary between two regions with different physical properties. Such partitioning is presented in Figure (4.8b). The principal geometrical parameters of such domain decomposition are given in Table (4.4).

	Ω^1	Ω^2	Ω^3	Ω^4
$N_\triangle$	191 474	102 752	66 553	65 795
$N_\bullet$	95 962	51 601	33 436	33 057

TABLE 4.3 – Number of triangles $N_\triangle$, number of nodes $N_\bullet$, i.e. number of unknowns in every local subdomain given by the decomposition in Figure (4.8b).

Using such domain decomposition into 4 subdomains (Figure 4.8b) we calculate the problem with the aid of the FEM classical and FETI-DPEM2-full methods. The results obtained with these two methods are presented in Figure (4.9). Then, in order to compare the obtained results not only visually, but also in a quantitative way, we compute the L^2 discrepancy (Eq. 4.2) presented in Table (4.4). When analysing this table, we can say that the special treatment of the corner DOFs in the framework



(a) Results of FEM

(b) Results of FETI

FIGURE 4.9 – Maps of the modulus of the electric field calculated with (a) the classical FEM and (b) the FETI-DPEM2-full methods, for the configuration of Figure (4.8b).

	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2-full
1	5.0637E-003	8.1753E-012
2	6.5133E-003	3.6063E-012
3	2.7721E-003	1.7938E-013
4	3.0508E-003	7.9388E-013

TABLE 4.4 – Relative L^2-errors of the FETI-DPEM2 classic and FETI-DPEM2-full methods in each subdomain for the decomposition into 4 subdomains given in Figure (4.8b).

of the FETI-DPEM2-full method, presented in Section (3.6.4.2), provides us better results in terms of relative L^2-error, as compared to the classical approach.

We have then played with the coefficient α^i arising in the Robin-type boundary conditions. The results are presented in Table (4.5). In this table the value *diff* corresponds to $\alpha^i = \frac{\sqrt{\mu_r^i \varepsilon_r^i} + \sqrt{\mu_r^k \varepsilon_r^k}}{2} jk_0$ suggested in [120]. From Table (4.5) we can see that the relative error does not depend on the value of the mathematical parameter α^i selected for the Robin-type boundary condition.

We also would like to pay attention on the importance of the arbitrary domain decomposition (Table 4.4). As you can see, even in such simple two-dimensional case, the equal partitioning in terms of geometry yields a big difference in terms of local number of unknowns. Note that the number of unknowns in Ω^1 is almost three times bigger than the one in Ω^4. This is because a strong scatterer is located inside Ω^1. It means that the LU-decomposition of the problem matrix, related to Ω^1, takes much more time than the one of Ω^4, which is not efficient in terms of computational time, nor required memory.

α^i	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2-full
jk_0	4.3500E-002	3.1887E-012
$2jk_0$	5.7462E-002	3.1887E-012
$3jk_0$	5.9161E-002	1.9781E-012
$4jk_0$	6.0614E-002	5.4414E-012
$5jk_0$	8.6018E-002	8.1677E-012
$\sqrt{\varepsilon_r}jk_0$	7.9300E-002	8.3769E-012
diff	5.6018E-002	8.5804E-012

TABLE 4.5 – Relative L^2-error of the FETI-DPEM2 full and classical methods while using different mathematical parameters α^i.

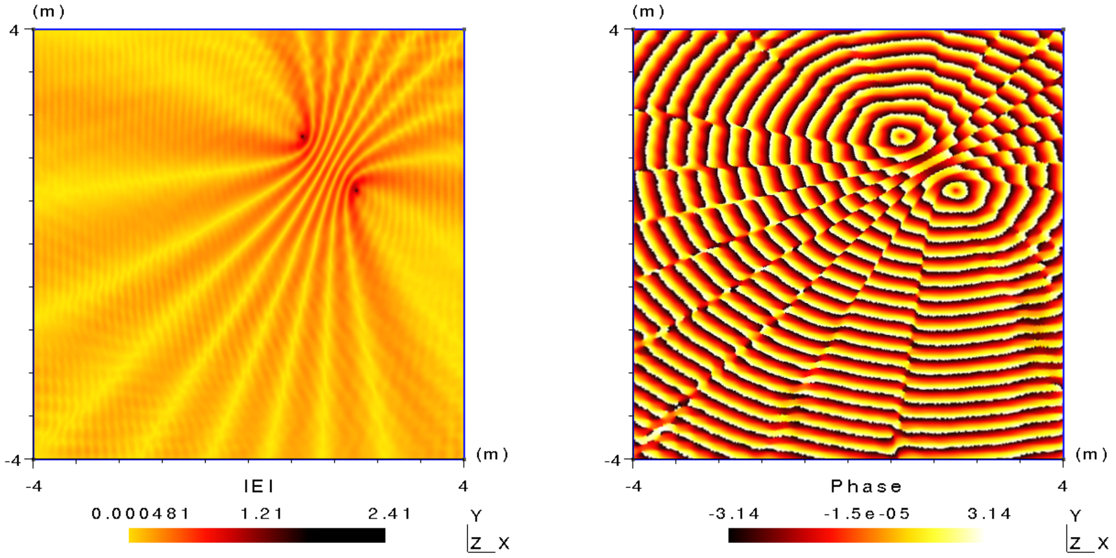
4.5 Anisotropic media

In this Section we are going to analyse the influence of anisotropic media. The aspects of the anisotropy were considered in Section (2.5) when we talked about PML. In this section we will consider a more general case, which in 2D E// configuration is given as :

$$\bar{\bar{\varepsilon}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \varepsilon_{zz} \end{bmatrix}, \ \bar{\bar{\mu}} = \begin{bmatrix} \mu_{xx} & \mu_{xy} & 0 \\ \mu_{yx} & \mu_{yy} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4.3}$$

Considering still the configuration of two sources radiating in free space at the working frequency of 800 GHz presented in Figure (4.1a), we will suppose now that the area is filled with a material which has the following properties : $\varepsilon^r_{zz} = 1.0, \mu^r_{xx} = 1.0, \mu^r_{yy} = 2.0$ and $\mu^r_{xy} = \mu^r_{yx} = 0$.

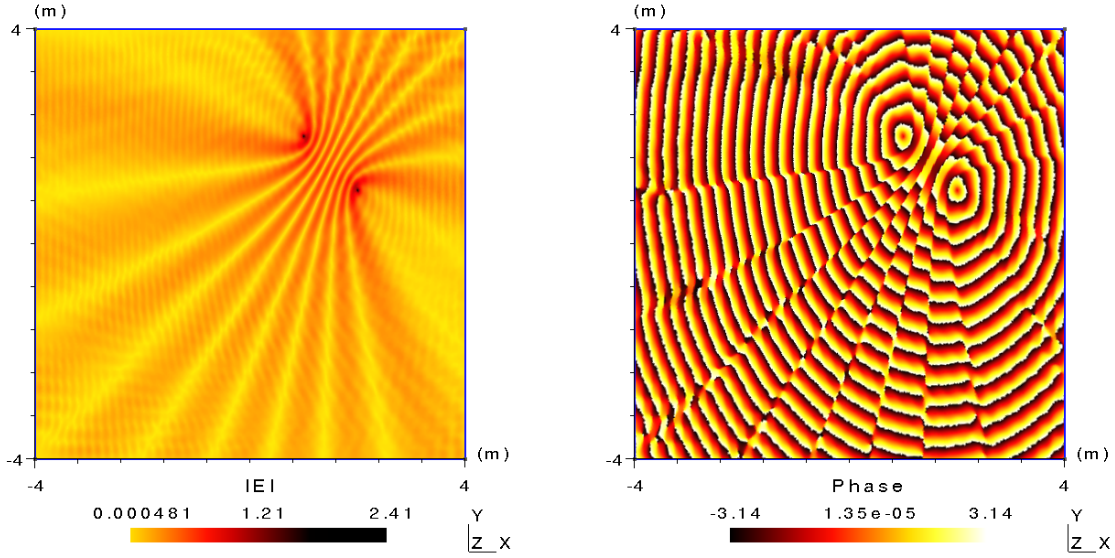
Following the idea of the FETI-DPEM2-full method, the domain has been partitioned into 20 subdomains. The coefficient α^i of the Robin-type boundary condition is set again to be equal to jk_0. The simulation performed with the proposed method is presented in Figure (4.10).



(a) Modulus (b) Phase

FIGURE 4.10 – Maps of the modulus and the phase of the electric field in an anisotropic media with $\varepsilon^r_{zz} = 1.0, \mu^r_{xx} = 1.0, \mu^r_{yy} = 2.0$ and $\mu^r_{xy} = \mu^r_{yx} = 0$, obtained with the FETI-DPEM2-full method.

Using the same domain decomposition into 20 subdomains we will consider the same problem, but now with $\mu^r_{xx} = 2.0, \mu^r_{yy} = 1.0$ and $\mu^r_{xy} = \mu^r_{yx} = 0$ for comparison. The results obtained with the FETI-DPEM2-full method are presented in Figure (4.11).



(a) Modulus

(b) Phase

Figure 4.11 – Maps of the modulus and the phase of the electric field in an anisotropic media with $\varepsilon^r_{zz} = 1.0, \mu^r_{xx} = 2.0, \mu^r_{yy} = 1.0$ and $\mu^r_{xy} = \mu^r_{yx} = 0$, obtained with the FETI-DPEM2-full method.

To prove the accuracy of these results, we can consider the quantitative comparison (Eq. 4.2) of the FETI-DPEM2 classic and full methods, as shown in Table (4.6).

	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2 full
Figure (4.10)	3.1836E-002	1.7933E-012
Figure (4.11)	5.5128E-002	5.5661E-012

Table 4.6 – Quality comparison of the FETI-DPEM2 classic and full methods in terms of relative error L^2.

As we can see from this table, the anisotropic media does not have any influence on the FETI results in terms of relative L^2-error when the Interface Problem is solved with a direct method.

4.6 PML influence

In this section we are going to analyse the influence of a Perfectly Matched Layer (PML) on the results of the FETI-DPEM2-full method. For this, let us add the PML to the problem considered in Figure (4.1a). We can schematically present the resulting configuration as shown in Figure (4.12).

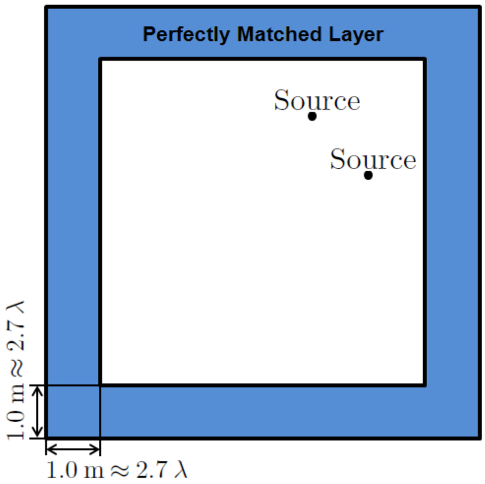
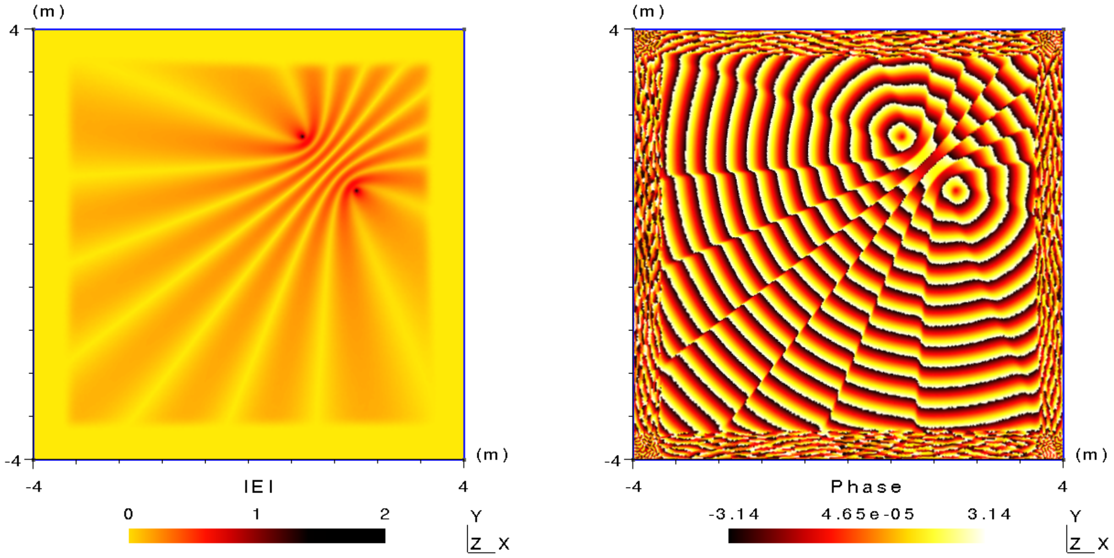


FIGURE 4.12 – Positions of the sources in the area filled with air and bounded with PML

Following the DD-idea, the discretized domain corresponding to the configuration of Figure (4.12) has been divided into 20 subdomains. The results obtained with the FETI-DPEM2-full method are presented in Figure (4.13).

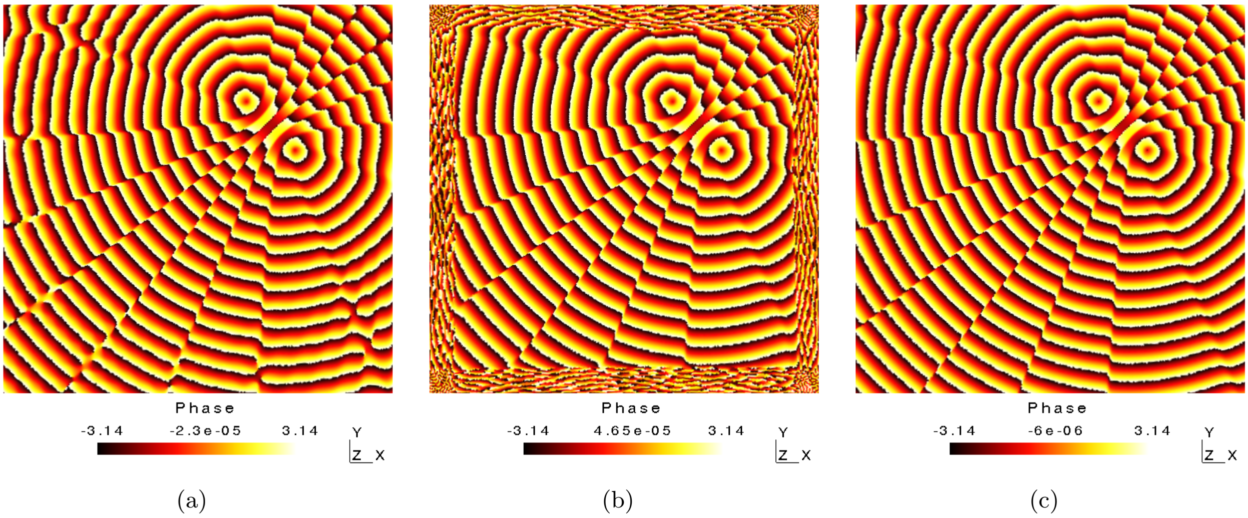


(a)

(b)

FIGURE 4.13 – Maps of the modulus and the phase of the electric field in a free space bounded with PML.

In order to understand the advantage of using PML, the reader should visually compare Figures (4.2) and (4.13). There exist some differences between the electric fields obtained with and without the PML. For the electric field computed with the radiation boundary conditions (Figure 4.2), close to the boundaries, there are some tenuous reflections, while these reflections disappear with the aid of PML (Figure 4.13). We then calculate the exact solution with the help of the Hankel function $\mathrm{H}_0^+(k_0 r)$ and compare the 3 cases in terms of the phase of the electric field (Figure 4.14).



(a) (b) (c)

Figure 4.14 – Phases of the electric field obtained with the FETI-DPEM2-full method in the domain bounded with (a) the radiation boundary condition, (b) the PML layer, compared to (c) the exact solution obtained with the aid of the Hankel function $\mathrm{H}_0^+(k_0 r)$.

From this figure we can clearly see that the solution obtained in the domain bounded with PML is closer to the exact one, than the solution obtained in the domain bounded with the radiation boundary condition.

In order to compare the obtained results not only visually, but in a quantitative sense as well, we calculate this problem with the FETI-DPEM2 classical method and perform the comparison of Eq. (4.2) with the method FEM.

PML	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2 full
Without	3.7902E-003	2.1513E-012
With	2.2056E-002	3.6919E-010

Table 4.7 – Quality comparison of the FETI-DPEM2 classic and full methods in terms of relative error L^2.

As we can see from Table (4.7), we almost do not have any influence of the PML area to the numerical results of FETI method when the interface problem is solved with a direct method.

4.7 Scattered field computation

Considering still the problem described previously in Section (4.6), we now introduce 3 scatterers described in the configuration of Figure (4.1b) in the domain filled with air and bounded with the PML. Schematically this configuration is presented in Figure (4.15).

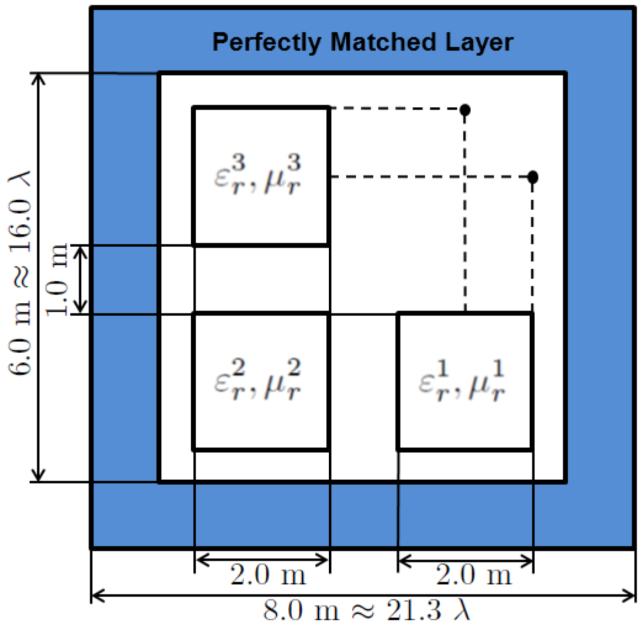


FIGURE 4.15 – Schematic configuration of the domain which is bounded with PML and containing two sources with 3 scatterers inside.

Following the principal Domain Decomposition idea, we divide the domain Ω into 16 subdomains without any geometrical conditions on the internal boundaries between sudbomains, as shown in Figure (4.16).

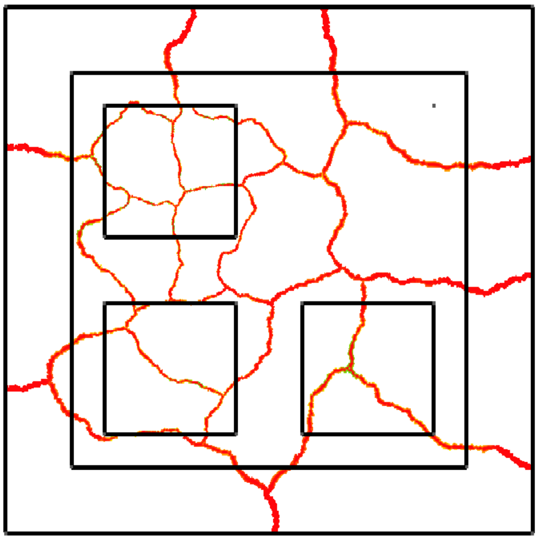
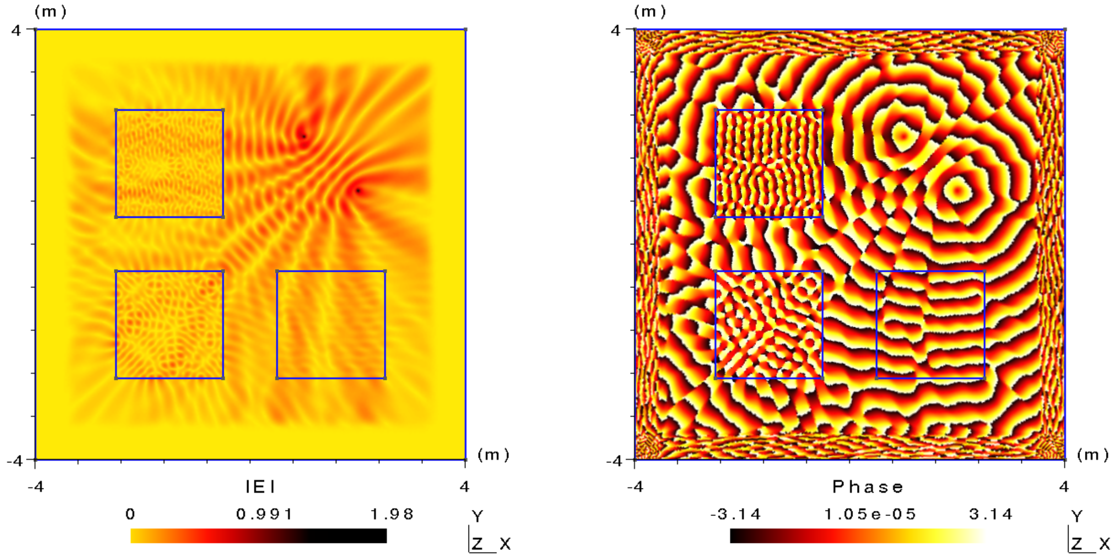


FIGURE 4.16 – Map of the schematic domain, showing the way the partitioning has been performed.

According to the scattered field formulation described in Section (2.2.1), the field $\mathcal{E}^{\text{sc}}$ can be calculated in two ways. To start with, we can find the scattered field as a subtraction of the total field with the incident field. Note that the latter has been calculated previously in Section (4.6), while the former can be calculated with the FETI-DPEM2-full method, using the domain decomposition presented above in Figure (4.16). The modulus and the phase of $\mathcal{E}^{\text{tot}}$ obtained with the proposed method are presented in Figure (4.17).

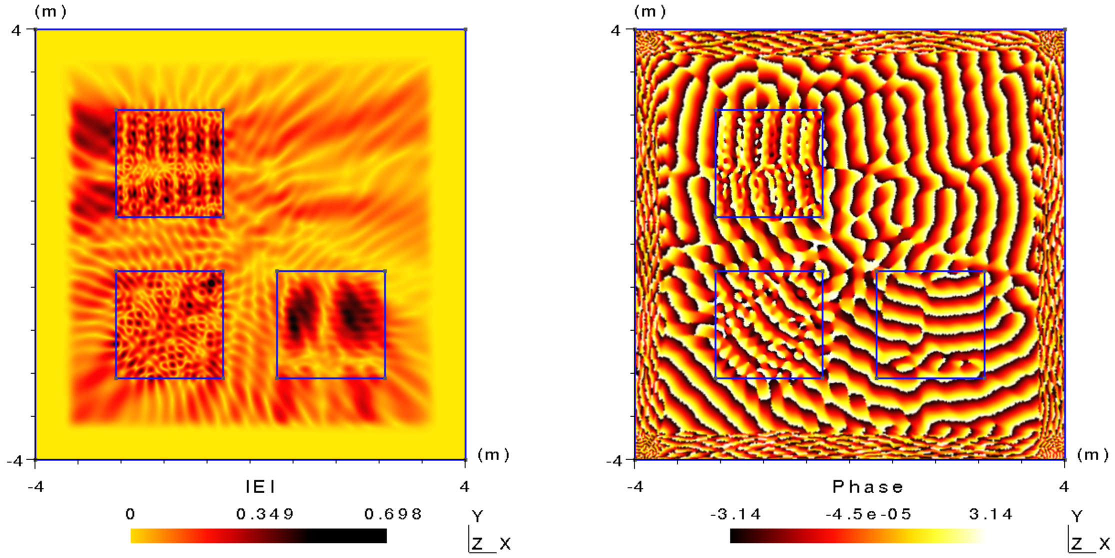


(a) Modulus

(b) Phase

FIGURE 4.17 – Maps of the modulus and the phase of the total field $\mathcal{E}^{\text{tot}}$ calculated with the help of the FETI-DPEM2-full method using a domain decomposition into 16 subdomains.

By subtracting the incident field $\mathcal{E}^{\text{inc}}$ from the total field $\mathcal{E}^{\text{tot}}$ we obtain the scattered field $\mathcal{E}^{\text{sc}}$ which is presented in Figure (4.18) in terms of modulus and phase.



(a) Modulus

(b) Phase

FIGURE 4.18 – Maps of the modulus and the phase of the scattered field obtained after subtracting the incident field from the total field.

In order to verify the obtained results we provide the quantitative comparison of the FEM and FETI methods (Table 4.8) as in Eq. (4.2).

	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2 full
$\mathcal{E}^{\text{inc}}$	2.6464E-002	1.0621E-012
$\mathcal{E}^{\text{tot}}$	6.2947E-002	4.4222E-011
$\mathcal{E}^{\text{sc}}$	4.4706E-002	2.2642E-011

TABLE 4.8 – Quantitative comparison of the FETI-DPEM2 classical and full methods in terms of relative L^2-error for the scattered field computation.

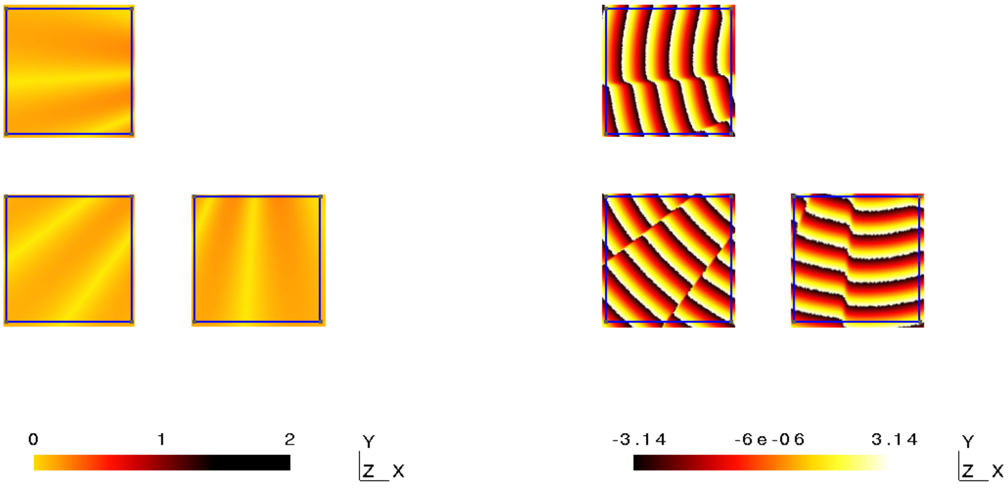
According to the second approach presented in Section (2.2.1), we have to know *a-priori* the incident field $\mathcal{E}^{\text{inc}}$ in order to solve Eq. (2.5) whose rhs J^{sc} contains this term. We can do it in two ways. The incident field $\mathcal{E}^{\text{inc}}$ can be found as :

1/ The exact solution based on the Hankel function $\mathrm{H}_0^+(k_0 r)$

2/ The solution of the Helmholtz equation (2.2)

Note that we have to know the incident field not everywhere in the domain Ω, but only in the places where the scatterers are located (Section 2.2.1).

To start with, we obtain the incident field as the value of the Hankel function $\mathrm{H}_0^+(k_0 r)$ applied for the configuration of Figure (4.12) inside the scatterers. The modulus and the phase of such incident field are given in Figure (4.19).



(a) Modulus

(b) Phase

FIGURE 4.19 – Maps of the modulus and the phase of the incident field in a free space obtained with the help of Hankel function $\mathrm{H}_0^+(k_0 r)$ applied for the configuration of Figure (4.12).

Using this incident field, we then calculate the Scattering problem (2.5) with the help of the FETI-DPEM2-full method with the domain decomposition into 16 subdomains presented above in Figure (4.16). The obtained results are almost identical and visually are indistinguishable as compared to the results obtained with the first approach (Figure 4.18).

We then use a previous solution $\mathcal{E}^{\text{inc}}$ (Figure 4.13) obtained with the FETI-DPEM2-full method in order to complete the rhs J^{sc} of Eq. (2.5). As previously, the obtained results visually are also very close to those of Figure (4.18) and are almost indistinguishable. The qualitative comparison of both approaches is given in Table (4.9).

	L^2-error FETI-DPEM2 classic	L^2-error FETI-DPEM2 full
Hankel	4.0017E-003	7.8773E-013
Helmholtz	9.0057E-002	8.7216E-012

Table 4.9 – Quality comparison of the FETI-DPEM2 classical and full methods in terms of relative L^2 error.

We then compare the two solutions of the proposed method obtained with Hankel and Helmholtz approaches in terms of relative L^2-error which is equal to 7.0289E-002. Such huge error can be explained by the same-order error existing between the $\mathrm{H}_0^+(k_0 r)$ solution and the regular incident field FEM solution, which is equal to 6.0384E-002. As the two incident fields are different, it is thus normal that the two resulting scattered fields are also different.

4.8 Conclusion

The two-dimensional numerical results presented here have shown that the new treatment of the corner points enables to provide more accurate results with respect to the classical FETI-DPEM2 method. We have been able to partition and handle internal interfaces wich are not necessarily straight lines without major drawback in terms of computational accuracy.

The Interface Problem corresponding to both the FETI-DPEM2 classical and full methods was solved with the direct method based on the LU-decomposition. We have seen in this chapter that there is no restriction for taking into account scatterers which are made of heterogeneous structures or even anisotropic materials. Showing the comparison of the results obtained for different values of the mathematical parameter α^i, one can be sure that by playing with this parameter we will not degrade the final solution of the method. This is an important conclusion, since we are going to apply iterative methods in the following chapter in order to solve the Interface Problem.

Chapitre 5

3D Forward Problems

Contents

5.1 Introduction

Finally, we arrive at the last chapter related to the Direct part, which is dedicated to solving three-dimensional electromagnetic problems. This type of problem is strongly dimension dependent and if in 2D the use of direct solvers like [117, 119, 121] is obvious (a typical scattering problem with 10^6 unknowns is solved in few dozen of seconds on a classical PC), the resolution of the linear system arising from the discretization of 3D configurations with a direct solver is much more tricky, time and memory consuming. The proposed FETI-DPEM2-full method is applied here in order to make this process faster and less memory dependent.

In the previous chapter, a series of numerical tests have shown the efficiency of the proposed method applied to two-dimensional electromagnetic simulations. In this chapter, we follow the same ideas applied this time to a 3D configuration and for the vectorial Helmholtz equation. We analyse a 3D free-space scattering problem, which leads us to implement a scattered field formulation within a domain truncated by Perfectly Matched Layers (PML) [87, 111]. We do not thoroughly explore the numerical efficiency of the FETI-DPEM2-full method as we have previously done in 2D, but prefer to focus our attention on the convergence behaviour of the iterative algorithm which is used to solve the interface problem. Indeed, iterative procedures appear unavoidable when dealing with large size problems. We show in this chapter that the convergence of the iterative algorithm is strongly affected by the presence of anisotropic materials. We also show how the method can be strongly accelerated by following the Evanescent Modes Damping Algorithm (EMDA) [106] which originally was proposed as a more accurate approximation of the Dirichlet-to-Neumann (DtN) operator for the non-overlapping Schwarz DDM [106], instead of classical Sommerfield boundary conditions between adjacent domains.

After discussing the efficiency of the FETI-DPEM2-full method, we will provide various numerical tests where we will present the obtained electrical fields visually (for example in terms of its imaginary part). This will allow us to perform primary tests in order to verify the obtained results. We will also compare the results of simulations with scattered fields measurements [80] in order to verify the effectiveness of the proposed method in practical situations.

5.2 Benchmark configuration

In this section, we compare the field computed thanks to the various approaches in the framework of a 3D scattering configuration. We assume that a domain Ω is filled with air (Figure 5.1a). The operating frequency is 1 GHz. The domain is a cube whose size is $[0.6 \times 0.6 \times 0.6]\text{m}^3$ that corresponds to the size $2\lambda \times 2\lambda \times 2\lambda$. A dipole is located at the position $[0.05, 0.05, 0.0]$m and oriented as $[1.0, 1.0, 0.0]$. The domain is surrounded by a cartesian uni-axial Perfectly Matched Layer (PML) whose width is $0.1\text{m} \approx 0.33\lambda$. The mesh size corresponds to 20 points per wavelength.

In this work we are going to describe the three-dimensional mesh obtained after finite-element discretisation by the following set of parameters (Table 5.1)

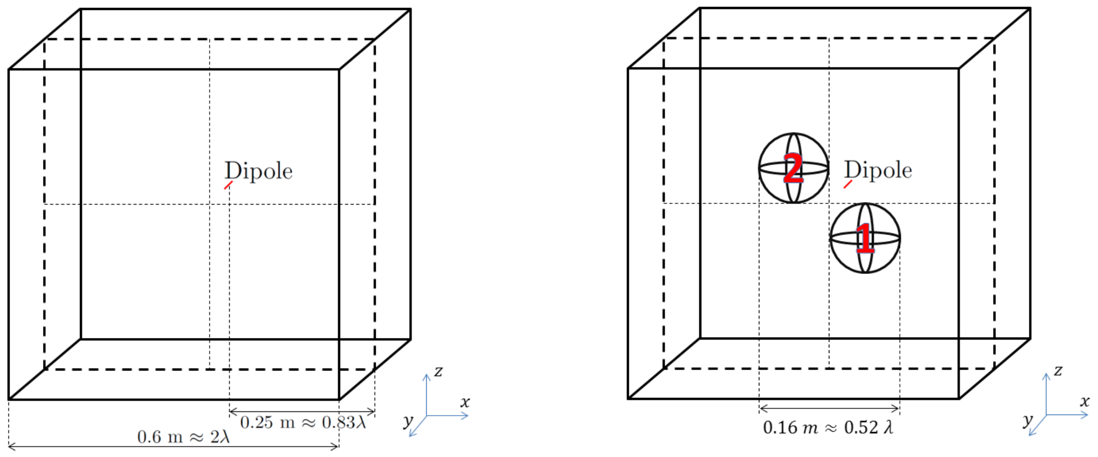
κ	N_p	N_t	N_e
20	43 189	245 174	298 771

Table 5.1 – Mesh obtained after finite-element discretisation of the problem presented in Figure (5.1a) consisting of N_p points, N_t tetrahedras and N_e edges, which corresponds to κ (Eq. 4.1) points per wavelength.

Next, we position two spherical scatterers with radiuses $R_1 = R_2 = 0.04$ m $= 0.13$ λ inside the domain Ω. The centers of the spheres are positioned at the points $(-R_1,\ -R_1,\ 0)$ and $(R_2,\ R_2,\ 0)$. A schematic map of the configuration is given in Figure (5.1b) All the scatterers are assumed to be non-magnetic ($\bar{\bar{\mu}}_r = I$), and they present a relative permittivity respectively equal to $\bar{\bar{\varepsilon}}_r^1 = 2.85 \cdot I$ and $\bar{\bar{\varepsilon}}_r^2 = 5.0 \cdot I$. After refining the mesh using Eq. (4.1) we obtain the following domain discretisation (Table 5.2) of the problem presented in Figure (5.1b)

κ	N_p	N_t	N_e
20	44 308	249 031	302 561

Table 5.2 – Mesh obtained after finite-element discretisation of the problem whose configuration is given in Figure (5.1b).



(a) Incident field configuration

(b) Total field configuration

Figure 5.1 – Position of (a) the dipole and (b) two spherical scatterers in the area filled with air. This configuration corresponds to (a) the incident and (b) the total field.

5.2.1 Fields distributions

Let us calculate at first the incident field, i.e. let us solve the problem presented in Figure (5.1a). Following the idea of the FETI-DPEM2-full method, we divide the domain Ω into 10 non-overlapping subdomains without any geometrical constraints and conditions on the boundaries between them (Figure 5.2) using a greedly-like algorithm provided by an automatic meshing partition software METIS [54].

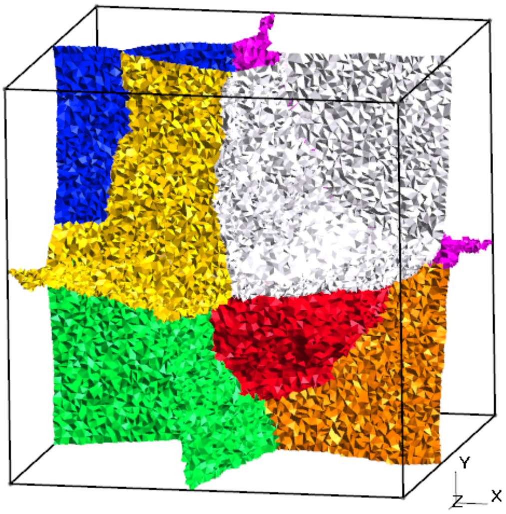


Figure 5.2 – Schematic map of the domain, showing the way the partitioning has been performed for $N_s = 10$.

Throughout all the three-dimensional computations we are going to describe the resulting domain decomposition by the following set of parameters :

N_s	N_{λ_r}	$N_{\mathbf{E}_c}$	$N_{\boldsymbol{\lambda}_c}$
10	33 739	463	1 379

Table 5.3 – Number of Lagrange multipliers λ_r associated to the Interface-edges (N_{λ_r}), number of corner edges ($N_{\mathbf{E}_c}$) and number of Lagrange multipliers $\boldsymbol{\lambda}_c$ associated to the corner edges ($N_{\boldsymbol{\lambda}_c}$) obtained after a domain partitioning of the problem presented in Figure (5.1a) into N_s subdomains.

As in the 2D case, we now use the frontal solver MUMPS based on the LU-decomposition [117] to solve the full Interface Problem (3.115). Following this idea we can control geometrical errors and guarantee the correct solution by comparing it with the solution of the classical FEM method. The imaginary parts of the incident field obtained with the classical FEM method and the FETI-DPEM2-full method are presented in Figures (5.3a) and (5.4a) respectively.

We then play with the number of subdomains ($N_s = 5, 10, 15, ...$) and present the results obtained with the FETI-DPEM2-full method in terms of the relative L^2-error (Table 5.4).
As we can see from Table (5.4) the relative difference is very small, so there is no reason to show the fields obtained with the different domain decomposition, because we will not be able to distinguish the results visually.

We now calculate the electromagnetic problem corresponding to the configuration of Figure (5.1b) in order to obtain the total field. First of all we use the classical FEM method for this aim. Then,

N_s	5	10	15
L^2-error	7.1134E-011	3.9787E-011	4.4427E-011

Table 5.4 – Relative L^2-error of FETI-DPEM2-full method for the incident field when different partitioning ($N_s = 5, 10, 15$) are used.

following the Domain Decomposition idea, we divide the given domain into 10 subdomains by using a similar partitioning as for the incident field (Figure 5.2). The main parameters of the domain decomposition for this problem are presented in Table (5.5)

N_s	N_{λ_r}	$N_{\mathbf{E}_c}$	$N_{\boldsymbol{\lambda}_c}$
10	35 586	495	1 418

Table 5.5 – Principal domain decomposition parameters of the problem corresponding to the total field presented in Section (5.2) given for a partitioning into $N_s = 10$ subdomains.

First of all, we compare the results obtained with the FETI-DPEM2-full and FEM classical methods visually, in terms of the imaginary part of the total field (Figures 5.3b and 5.4b). Like previously, we can not visually see the difference between these two results. Then, in order to give some quantitative criteria, we play with the number of subdomains and calculate the relative L^2-error between FETI and FEM methods. The results are listed in Table (5.6).

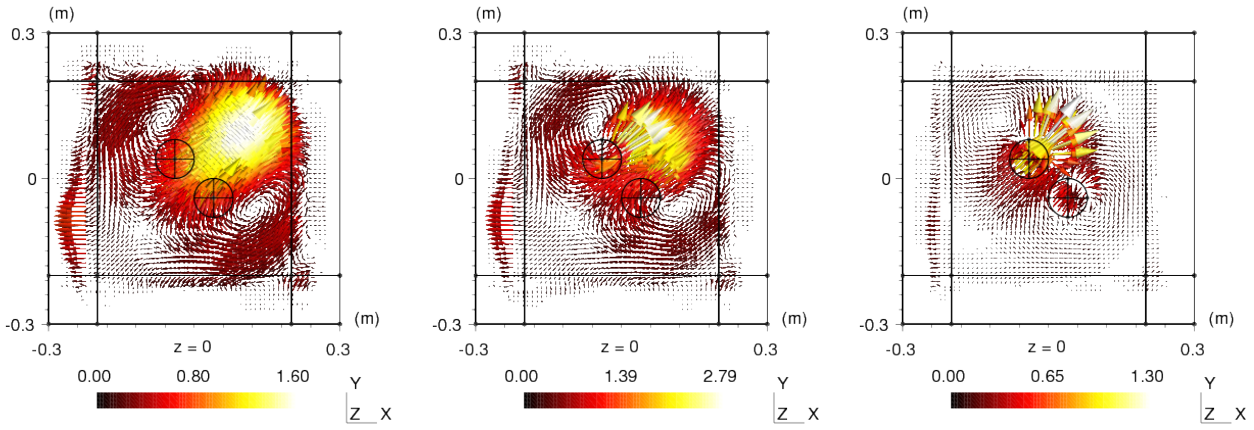
N_s	5	10	15
L^2-error	8.1993E-012	2.9217E-011	9.5612E-012

Table 5.6 – Relative L^2-error of the FETI-DPEM2-full method for the total field when different partitioning ($N_s = 5, 10, 15$) are used.

Finally, we present the results of the scattered-field formulation. Using the same finite-element discretisation (Table 5.2) and domain decomposition into 10 subdomains (Table 5.5) as for the total field configuration, we solve Eq. (2.13) using the classical FEM method and the method FETI-DPEM2-full. It must be noticed here that we have used the scattered field formulation even if the source is enclosed in the computational domain. The results of the imaginary part of the scattered field obtained with the FEM classical and FETI-DPEM2-full methods are presented in Figures (5.3c) and (5.4c) respectively. From the comparison of these figures, we conclude that the results are in good agreement. This also can be proved by comparing these results in terms of the relative L^2-error between FETI and FEM methods (Table 5.7).

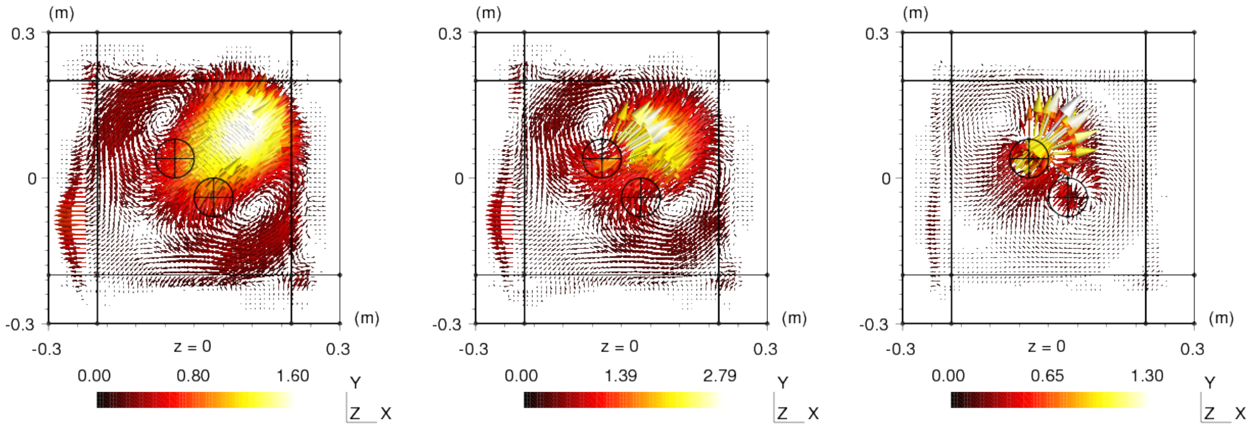
N_s	5	10	15
L^2-error	1.7938E-012	9.9886E-011	3.6063E-011

Table 5.7 – Relative L^2-error of FETI-DPEM2-full method for the scattered field when different partitioning ($N_s = 5, 10, 15$) are used.



(a) Incident field $\mathcal{E}^{\text{inc}}$ (b) Total field $\mathcal{E}^{\text{tot}}$ (c) Scattered field $\mathcal{E}^{\text{sc}}$

FIGURE 5.3 – Imaginary part in the plane $z = 0$ of the (a) incident, (b) total and (c) scattered fields calculated with the FEM classical method.



(a) incident field $\mathcal{E}^{\text{inc}}$ (b) Total field $\mathcal{E}^{\text{tot}}$ (c) Scattered field $\mathcal{E}^{\text{sc}}$

FIGURE 5.4 – Imaginary part in the plane $z = 0$ of the (a) incident, (b) total and (c) scattered fields calculated with the FETI-DPEM2-full method. The investigation domain has been divided into 10 subdomains.

5.2.2 Implementation difficulties

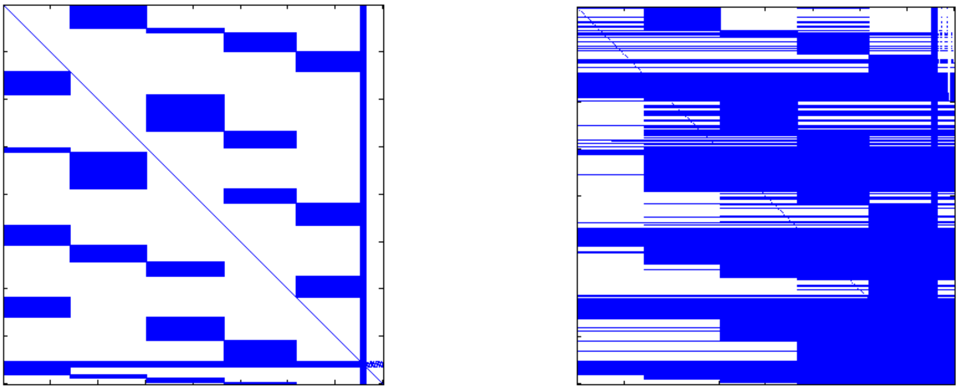
All the FETI methods with arbitrary mesh partition become very complex in terms of 3D geometry. In particular, it is very difficult to control the construction of the geometrical matrices presented in Section (3.6.4). Indeed, there are a huge number of geometrical quantities that must be linked properly and/or extracted by means of the matrices $\mathbb{B}$, $\mathbb{D}$, $\mathbb{T}$, ... Moreover, the normals at the interfaces as well as the connection with the neighbouring domain must be handled with care. A lot of effort has thus been spent in order to verify the construction of such matrices and to control step by step the implementation of the FETI method. In order to verify the obtained results, we have used three approaches :

- Tests based on the use of analytical functions (Section 3.8.2)
- Tests based on the resolution of physical problems (Section 3.8.3)
- Comparisons with the classical FEM method

5.3 Iterative methods

As it was already noticed, all the results in Section (5.2.1) are obtained with the help of the frontal solver MUMPS based on the LU-decomposition. Unfortunately, the direct methods for the Interface Problem become very expensive in terms of required memory and computational time, especially for the large-dimension systems.

As an example, we can consider the profile of the Full Interface Problem (Figure 5.5a) which corresponds to the scattered field configuration from Section (5.2), where the domain Ω has been divided into 5 subdomains. In order to solve the Interface Problem we perform the LU-decomposition of this matrix. The profile of L and U matrices are given in Figure (5.5b)



(a) Before LU-decomposition (b) Before LU-decomposition

FIGURE 5.5 – Profile of (a) the Full Interface Problem (Eq. 3.115) and (b) profile of its LU-decomposition presented as one entire matrix. Colored points denote non-zero elements.

It is obvious that these 2 matrices are definitely not sparse, requiring therefore a large amount of memory storage. To avoid these problems we have to use iterative methods.

5.3.1 Krylov-subspace methods

In our work we have considered two iterative methods such as the Generalized Minimal Residual Method (GMRES) and the Biconjugate Gradient Stabilized Method (BICGSTAB). These two methods are based on Krylov subspaces [23], i.e. subspaces of the form

$$\mathcal{K}_m(A, v) \equiv \text{span}\{v, Av, A^2v, ..., A^{m-1}v\} \tag{5.1}$$

where A is a matrix arising in the system of linear equations $Ax = b$ to solve, v is a vector of the space of rational numbers. The dimension of the subspace of approximants m increases by one at each step of the approximation process.

The GMRES method is derived from Arnoldi process [122] for constructing an L_2-orthogonal basis of Krylov subspaces. In [18] it was shown that the method is theoretically equivalent to the Generalized Conjugate Residual (GCR) method and to ORTHODIR [23]. The implementation of this method is similar to the Full Orthogonalization Method.

In [48, 96, 97] the Interface Problem arising from the classical FETI-DPEM2 method is solved with the help of the BICGSTAB method. This algorithm is a variation of the conjugate gradient squared (CGS) method, which, in cases of irregular convergence, may lead to substantial build-up of rounding errors, or possible overflow. BICGSTAB was developed to remedy this difficulty. This method produces iterations whose residual vectors r'_j are of the form

$$r'_j = \varrho_j(A)\varsigma_j(A)r_0 \tag{5.2}$$

in which r_0 is a initial residual, ς_j is the residual polynomial associated with the BCG algorithm and ϱ_j is a new polynomial defined recursively at each step with the goal of stabilizing the convergence behavior of the algorithm.

Looking ahead, we have to say that from a numerical point of view, the BICGSTAB method is almost not suitable to solve the electromagnetic problems with PML area. In all the test cases that we have performed it never converged and always broke down. That is why in this work we decided to use the GMRES method, because, as it was shown in [18], it can not break down and therefore is more stable.

Considering still the system of linear equations in the general form $Ax = b$, we recall that a matrix A is called ***positive definite*** if its symmetric part $(A + A^*)/2$ is Symmetric Positive Definite matrix. This is equivalent to the property that $\langle Ax, x\rangle > 0$ for all nonzero real vectors x. Also we would like to remind that all the eigenvalues of every symmetric matrix are always *real*.

The convergence of the iterative methods based on the Krylov-subspace highly depends on the definition of the matrix A. Indeed, the methods based on this subspace stagnate or may not converge (it depends on the method) if the symmetric part of the matrix is not positive definite [18, 23]. This property is important for us, as we are going to apply the GMRES method which is parts of the Krylov-subspace methods.

5.3.2 Applying the Iterative method to the Interface Problem

In this section we are going to use the GMRES method to solve the full Interface Problem arising from the scattered problem described in Section (5.2). Let us now set some parameters for the GMRES method (Eq. 5.3)

$$\begin{aligned} m &= 10 \\ \eta &= 5 \cdot 10^{-3} \end{aligned} \tag{5.3}$$

where m is the size of the Krylov subspace and η is the stopping criterion. These parameters of the GMRES method will be called *normal* and will be used in the future during *all the calculations*. For the domain decomposition into $N_s = 10$ subdomains we plot the convergence results of the full Interface Problem (Figure 5.6).

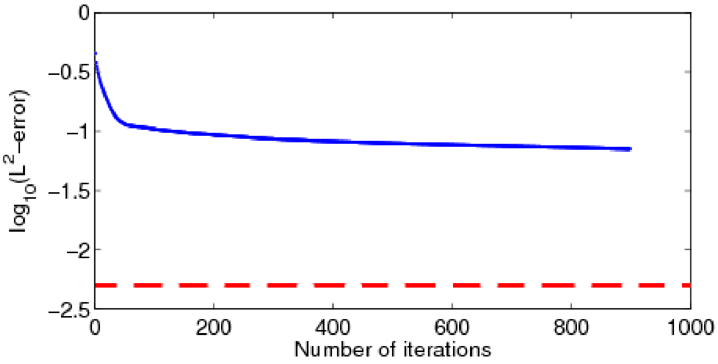


FIGURE 5.6 – (—) Convergence of the GMRES method without preconditioning for solving the full Interface Problem ($N_s = 10$) and (- -) the required relative error of $5 \cdot 10^{-3}$.

From Figure (5.6) we can conclude that the GMRES method for the full Interface Problem stagnates at some level and will hardly achieve the required relative error (in our case 5.E-003). It is worth mentioning that for the different domain decomposition into $N_s = 5..15$ subdomains, we have not been able to calculate this Scattering problem when the full Interface Problem was solved, i.e. the GMRES method has always stagnated. In the following sections we will consider the possible sources of problems and how to overcome these issues.

5.3.3 Spectrum behavior

5.3.3.1 Bad conditioning issue

This problem happens very rarely, but anyway deserves to be considered and discussed. It may happen that the local matrices K^i are very badly conditioned. In this case, if the relative error after inverting the matrix K^i is 10^{-5} (for example), at the end of the resolution of the Interface Problem (even with a direct method), we will not obtain a relative error for the FETI-method with respect to the method FEM better than 10^{-5}. It is totally normal and, again, it does not happen very often.

5.3.3.2 Eigenvalues issue

As a rule, the system of linear equations which we call the Interface Problem is indefinite [43, 97]. Sometimes it leads to problems, due to the reasons described at the end of Section (5.3.1).

In this section we will study the properties of the Interface Problem. Note, that some research work has been performed in order to study the definition of the matrix corresponding to the Interface Problem [44, 45, 96]. In this PhD thesis we stick with the classical approach given by Y. Saad, M.H. Schultz in 1986 [18, 23]. According to this approach, we need to calculate the eigenvalues of the *symmetric part* of the matrix corresponding to the Interface Problem. It is the only way that we are going to use in order to study the definition of the Interface Problem in this PhD thesis.

For the Scattering problem described in Section (5.2) with the domain decomposition into $N_s = 3,\ 5,\ 10$ subdomains we calculate at first the vector of eigenvalues for the full Interface Problem (3.115). The results are presented in Figure (5.7).

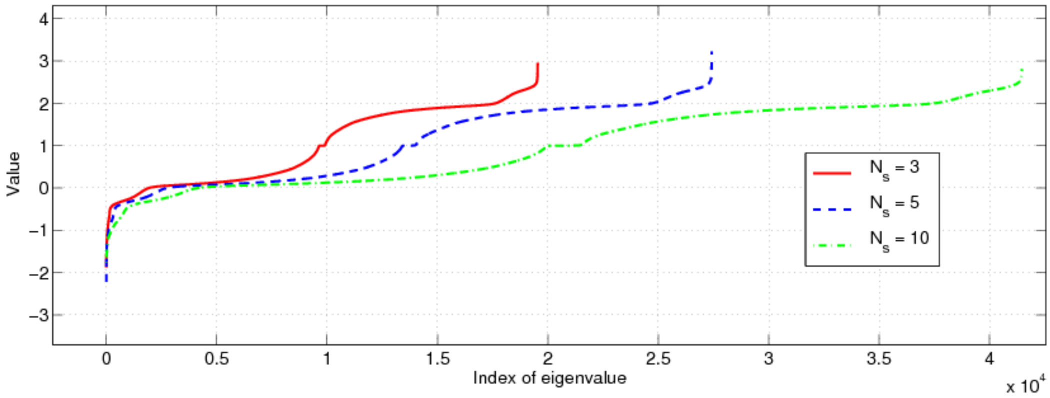


FIGURE 5.7 – Eigenvalues of the symmetric part of the full Interface Problem with PML when different partitioning ($N_s = 3,\ 5,\ 10$) are used.

From Figure (5.7) we can conclude that the eigenvalues of the full Interface Problem have a regular, almost anti-symmetrical structure. As you can see, most of the eigenvalues are positive, but unfortunately not all of them.

Considering still the Scattering problem from Section (5.2), we then have studied the eigenvalues corresponding to the reduced Interface Problem (3.116). The eigenvalues behaviour is presented in Figure (5.8). From this figure we can see, that the number of negative eigenvalues is almost the same

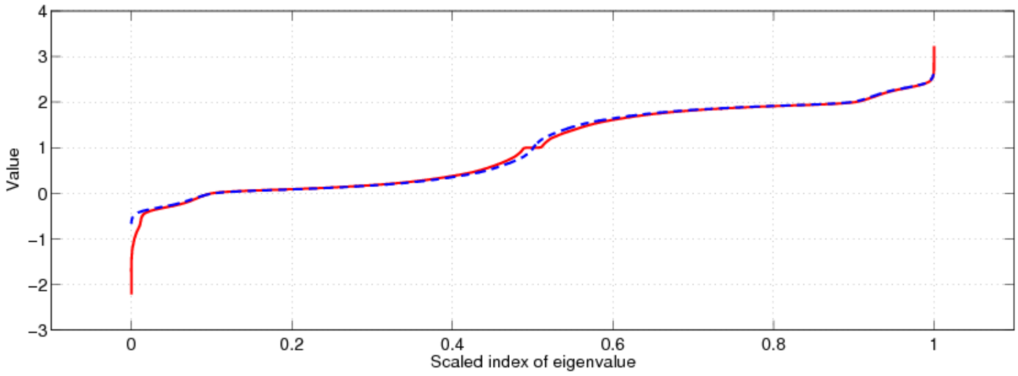


Figure 5.8 – The eigenvalues of the symmetrical part of (—) the full (Eq. 3.115) and (- - -) the reduced Interface Problem (Eq. 3.116) on the common scaled axis X = [0,1].

for the full and reduced Interface problems. Nevertheless, the reduced problem is better conditioned than the full one, which correlates with [39,41,43,113]. In consequence of such results we would like to note that, from this moment, ***we are going to consider only the reduced Interface Problem.*** In order to study the efficiency of the reduced IP, we now play with different domain decompositions. The results obtained with the FETI-DPEM2-full method applied to $N_s = 5, 10, 15$ subdomains are given in Table (5.8).

N_s	L^2-error FETI-DPEM2-full + MUMPS	L^2-error FETI-DPEM2-full + GMRES	Number of iterations
5	1.7938E-012	2.0807E-002	48
10	9.9886E-011	1.9119E-002	308
15	3.6063E-011	3.2492E-002	899

Table 5.8 – Results of calculations using the FETI-DPEM2-full method with different partitioning ($N_s = 5, 10, 15$) when the reduced Interface Problem is solved.

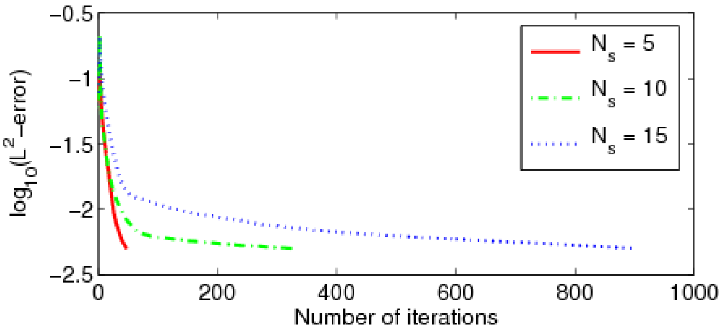


Figure 5.9 – Convergence of the GMRES method without preconditioning for solving the reduced Interface Problem when different partitioning ($N_s = 5,\ 10,\ 15$) are used.

Note that, by moving from the full Interface Problem to the reduced one, we have been able to calculate the Scattering Problem described in Section (5.2) with the FETI-DPEM2-full method thanks to the GMRES method. Nevertheless, as you can see from Table (5.8) when using this iterative method, the number of iterations depends highly on the number of subdomains. The convergence behavior is presented in Figure (5.9).

From Figure (5.9) we can conclude that for the problem at hand, the GMRES method can stagnate and will never reach the level even of 1.E-4 in the case of $N_s \geqslant 10$. A way to overcome this problem will be proposed later on, in Section (5.3.5.1).

5.3.4 PML-layer issue

After performing various numerical tests, it appeared that the PML area had some influence on the convergence process of the GMRES method. This issue was not encountered in 2D as we only employed there a direct solver.

5.3.5 PML influence

To understand the influence of the PML, we will consider the Scattering problem described in Section (5.2) with a partitioning into $N_s = 50$ subdomains. The main domain decomposition parameters corresponding to such partitioning are given in Table (5.9)

N_s	N_{λ_r}	$N_{\mathbf{E}_c}$	$N_{\boldsymbol{\lambda}_c}$
50	81 956	2 760	7 838

TABLE 5.9 – Principal domain decomposition parameters of the Scattering problem presented in Section (5.2) for a partitioning into $N_s = 50$ subdomains.

To solve the reduced Interface Problem we use the GMRES method with normal parameters given by Eq. (5.3). The main results related to this problem are presented in Table (5.10).

N_s	L^2-error FETI-DPEM2f + MuMPs	L^2-error FETI-DPEM2f + GMRES	Number of iterations
50	5.5514E-010	9.1813E-002	1275

TABLE 5.10 – Results of calculations with the FETI-DPEM2-full method using a partitioning into $N_s = 50$ subdomains.

For this problem we calculated also the relative L^2-error in *each* subdomain. Some of them are presented in Table (5.11).

Ω^i	27	28	29	31	41
L^2-error	0.5085	0.2435	0.3509	0.5687	0.2995

TABLE 5.11 – Relative L^2-error within some subdomains.

As you can see from Table (5.11), sometimes this error is really huge. Such results seem a little bit surprising for us, because the relative error between the FEM-solution and FETI-one is 9.1813E-002 (Table 5.10). Moreover, visually we hardly can see the difference between these two solutions. Then we studied more precisely these 5 ”problem-subdomains”, whose numbers are given in Table (5.11), and it turned out that all of them are subdomains *containing a PML area*.

To explain such behaviour we are going to obtain the *exact* values of the vectors λ_r, $\mathbf{E}_c$ and $\boldsymbol{\lambda}_c$. By *exact* values, we understand the solution of the same Scattering problem with the FETI-DPEM2-full method, whose Interface Problem is solved with a direct method (LU-decomposition of MUMPS). Now we select the subdomain with the largest error (the subdomain Ω^{31}, see Table 5.11) and trace its Lagrange multipliers λ_r (Figure 5.10).

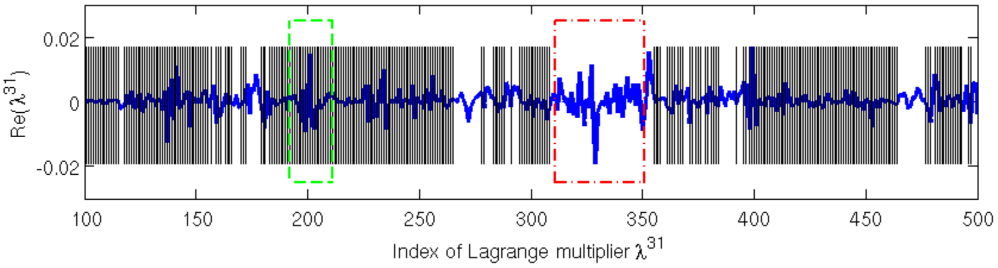


FIGURE 5.10 – Real part of the Lagrange multipliers λ in the domain Ω^{31}. The black lines denote the Lagrange multipliers which correspond to a PML region. The regions denoted with (–·) and (– –) are going to be considered more precisely in Figures (5.11) and (5.12) respectively.

The black vertical lines denote the Lagrange multipliers which are in the zone of the PML. As you can see, almost all this subdomain is positioned into the PML. Now we are going to perform few GMRES iterations and to look precisely the convergence behaviour of the Lagrange multipliers which are located in the area *with* PML and *without* PML.

To start with, we plot in Figure (5.11) the Lagrange multipliers during the iteration process in the region *without* PML (which corresponds to the square (–·) in Figure 5.10).

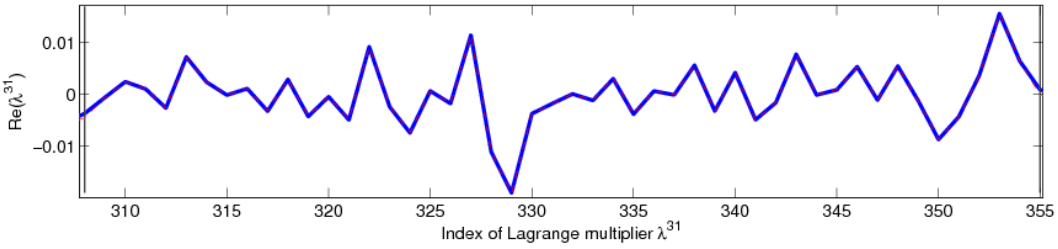


FIGURE 5.11 – Real part of some Lagrange multipliers λ in Ω^{31}, which are not inside a PML. (—) Exact values, (- - -) Computed after 10 GMRES iterations.

From this figure we can conclude that even after 10 iterations we obtain a perfect convergence in the area *without* PML, as the iterative solution in this figure totally corresponds to the exact one.

Unfortunately, we can not say the same about the Lagrange multipliers which are *inside* the zone of the PML. Even after 100 iterations (Figure 5.12) for this Scattering problem we will not obtain a convergence as good as the one we obtained for the previous case without PML.

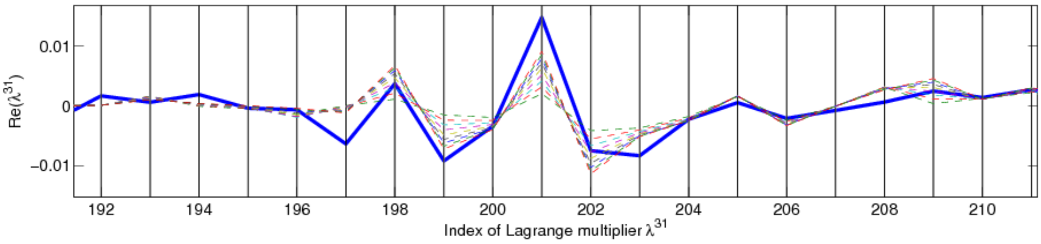


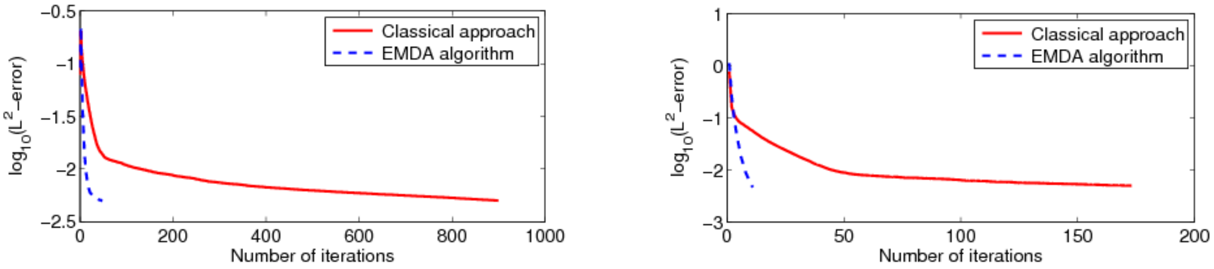
FIGURE 5.12 – Real part of some Lagrange multipliers λ in Ω^{31}, which are inside a PML. (—) Exact values, (- - -) Computed after 10 to 100 GMRES iterations.

As we can see from Figure (5.12), at some points, there is even no convergence at all. That is the reason why the iterative method can stagnate at some level [18, 23].

5.3.5.1 The EMDA-algorithm

Considering still the Scattering Problem described previously, we will now try to decrease the number of iterations obtained in Sections (5.3.3.2) and (5.3.5). To do so, we need to reconsider the internal boundary conditions, as it is known to impact the convergence of the domain decomposition algorithms. When α is simply set to jk_0 [123], the approximation of the Dirichlet-to-Neumann operator in the Robin-type boundary conditions given by Eq. (3.140) does not treat efficiently the evanescent modes [45]. The *Evanescent Modes Damping algorithm* (EMDA) proposed in [106] enables to extend the transmission boundary conditions to these evanescent modes. We only consider here the EMDA case where $\alpha = jk_0(1+j\chi)$, χ being a real-valued positive coefficient. The optimal value for χ depends in particular from the mean curvature on the interface [46]. As the partitioning is performed in an irregular fashion, we simply set $\chi = 0.5$ in the following computations as it is favorized in [46].

The electromagnetic problems corresponding to the total and scattered field (Section 5.2) are then computed for the domain decomposition into $N_s = 15$ subdomains using the EMDA approach. The convergence results for these problems are presented in Figure (5.13).



(a) Scattered field (b) Total field

FIGURE 5.13 – Number of iterations required for achieving a relative error of 5.E-3 when computing the scattered and the total fields using (—) the classic approach and (- - -) the EMDA approach for the reduced Interface Problem.

From this figure we observe a very good progress in terms of the number of iterations required to reach the desired stopping criterion for both the total and the scattered field computation.

The same Scattering problem is computed for different partitioning ($N_s = 15$ to 50 subdomains) and the associated L^2-norm error is presented in Table (5.12). When comparing this table with Tables (5.8) and (5.10), it is obvious that the number of iterations has been drastically reduced and no stagnation effect is now visible.

N_s	L^2-error FETI-DPEM2-full + MUMPS	L^2-error FETI-DPEM2-full + GMRES	Number of iterations
15	3.6063E-011	7.9494E-002	48
20	1.2913E-010	7.7824E-002	46
30	8.5017E-011	6.7407E-002	38
40	2.7154E-010	8.1373E-002	175
50	5.5514E-010	8.3392E-002	118

Table 5.12 – Results of calculations obtained with the FETI-DPEM2-full method for various partitioning when the reduced Interface Problem is solved with the EMDA approach.

5.3.5.2 Trying without PML

As we have already seen in Section (4.6), the PML helps us to avoid the reflections from the external boundaries which arise when the radiation boundary condition (Eq. 2.17) is set. In our work the quality of the results that we obtain even without PML is fully sufficient when comparing them with measurements. That is why, for the solution of real physical problems (Section 5.5) we set the radiation boundary condition on the external boundary. It is, in our opinion, the most effective tool for obtaining a good convergence when solving a physical problem.

Considering still the Scattering problem described in Section (5.2), we now replace the PML layer with the radiation boundary condition. For this test case, we use both $\alpha = jk_0$ and the algorithm EMDA with $\alpha = jk_0(1 + j\ 0.5)$. We now compare the simulated results of the problem *with* PML (Table 4.4) and *without* (Table 5.13) for different domain decompositions, but for the same GMRES normal parameters (Eq. 5.3).

N_s	L^2-error *with* PML	L^2-error *without* PML	Number of iterations *with* PML	Number of iterations *without* PML jk_0	Number of iterations *without* PML $jk_0(1 + j\ 0.5)$
20	7.7824E-002	5.5654E-003	46	7	10
30	6.7407E-002	7.1044E-003	38	7	9
40	8.1373E-002	9.0133E-003	175	10	11

Table 5.13 – Results obtained *with* and *without* PML when different partitioning ($N_s = 20, 30, 40$) are used.

From Table (5.13) we can clearly see the advantage of replacing the PML region by the radiation boundary condition. Although, we would like to note that the EMDA approach was created in order to damp the evanescent modes. So this method might not improve the results of the propagating modes. That is what we see in Table (5.13), where the classical coefficient jk_0 provided better results than the EMDA. Thus, we apply this EMDA approach only for problems containing anisotropic media and, in particular, PML.

5.4 Efficiency of the FETI-DPEM2-full method

As an edge-element implementation of the FETI-DP method [39] for the electromagnetic analysis, the FETI-DPEM is a method with a strong mathematical and computational background. That is why before applying the FETI-DPEM method to real physical problems we would like to consider its efficient numerical realisation. For this we will detail the algorithm which we use in this work. It will give us an idea of the memory usage and the computational time required for each calculation step.

5.4.1 Numerical algorithm

Following the Domain Decomposition idea, the computational domain Ω is divided into N_s non-overlapping subdomains automatically thanks to an automatic meshing partition software METIS [54], as it has been done previously. In general, the subdomain interfaces resulting from such a decomposition typically have irregular shapes without any conditions on the boundaries between subdomains. Thus, the creation of the Boolean projection matrices presented in Section (3.6.3) must be performed very carefully. To control its creation we already discussed some techniques in Section (3.8).

The implementation of the FETI-DPEM2-full method, as in the classical FETI-DPEM2, consists of three major steps [39, 97]

1. Preprocessing step
2. Solution of the reduced Interface problem (Eq. 3.116)
3. Calculation of the electric field in every subdomain

In the preprocessing step, each subdomain problem is processed, where the subdomain finite element matrices $K^i_{rr}, K^i_{rc}, K^i_{cr}, K^i_{cc}$ and $M^i_{rr}, M^i_{rc}, M^i_{cr}, M^i_{cc}$ are assembled. The global corner DOFs related system $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ is fully assembled over all the subdomains. Such a corner system has the size N_c^2 that makes it relatively small due to the small number of corner DOFs. Hence, a local copy of its factorisation in the sparse format is kept during all the computation [97]. In addition, the sparse matrices K^i_{rr} are factorized with the frontal solver Mumps [117] for all the subdomains, whose factorizations are stored and will be used repeatedly throughout all the computations. Once it is done, we calculate the Right-hand side vector $\hat{\mathrm{d}}_{\lambda_r}$ whose size is N_{λ_r}.

The preprocessing step is followed by the iterative solution of the reduced Interface problem (Eq. 3.116). The idea of the IP solution is similar to the one that was presented in [97] for the classical FETI method. In this work we focus on :

1. The extension of this idea to the proposed FETI-DPEM2-full method
2. The detailed explanation of the Matrix-vector multiplication process throughout the iterative resolution of the reduced Interface problem

Once the Interface problem is solved, as well as the global corner DOFs related system (Eq. 3.117), the fields in the subdomains can be recovered independently by using the computed boundary conditions at the subdomain interfaces with the help of Eq. (3.87).

5.4.1.1 Matrix-vector multiplication

As we are interested in using this method for real physical problems, we are going to talk about the Multi-source formulation at once. In this PhD thesis, we base our work on the extension of the FETI-H method proposed in [124], which was designed for the solution of acoustic scattering problems with multiple right-hand sides. This algorithm is based on the step-by-step solution of the Interface Problem for each rhs vector.

When using an iterative method to solve the Interface problem (Eq. 3.116) the multiplication (5.4) needs to be evaluated at the k_{th} iteration.

$$\delta^k = \left(\mathrm{F}_{\lambda_r\lambda_r} + \mathrm{F}_{\lambda_r\mathbf{E}_c}\ \hat{\mathrm{F}}^{-1}_{\mathbf{E}_c\mathbf{E}_c}\hat{\mathrm{F}}_{\mathbf{E}_c\lambda_r} \right) \lambda_r^k \tag{5.4}$$

In this work we consider the terms λ_r^k and δ^k as vectors of size N_{λ_r}, even if we deal with a Multi-source formulation. It is worth mentioning that we could treat them as matrices, but it would not be effective in terms of memory usage.

Unlike the results presented in [97] we use only *two* cycles over all the subdomains to calculate the matrix-vector multiplication δ^k. In full-form, we can rewrite it as

$$\delta^k = \left[\mathrm{F}_{\lambda_r\lambda_r} + \mathrm{F}_{\lambda_r\mathbf{E}_c}\ \hat{\mathrm{F}}^{-1}_{\mathbf{E}_c\mathbf{E}_c} \left(\mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\ \mathrm{F}^{-1}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c}\ \mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r} - \mathrm{F}_{\mathbf{E}_c\lambda_r} \right) \right] \lambda_r^k \tag{5.5}$$

The multiplication presented in Eq. (5.5) can be accomplished in four steps :

STEP 1. Calculate
$\delta^k = \mathrm{F}_{\lambda_r\lambda_r}\ \lambda_r^k$
$z_1^k = \mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r}\ \lambda_r^k$
$z_2^k = \mathrm{F}_{\mathbf{E}_c\lambda_r}\ \lambda_r^k$
in one cycle over all the subdomains.
And then $z^k = \mathrm{F}_{\mathbf{E}_c\boldsymbol{\lambda}_c}\ \mathrm{F}^{-1}_{\boldsymbol{\lambda}_c\boldsymbol{\lambda}_c} z_1^k - z_2^k$

STEP 2. Solve $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c} y^k = z^k$ for y^k

STEP 3. Calculate $x^k = \mathrm{F}_{\lambda_r\mathbf{E}_c}\ y^k$

STEP 4. Calculate $\delta^k = \delta^k + x^k$

To start with, we will explain how we perform STEP 1 in only one cycle over all the subdomains. For this, we are going to consider properly one more time the matrices $\mathrm{F}_{\lambda_r\lambda_r}, \mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r}$ and $\mathrm{F}_{\mathbf{E}_c\lambda_r}$ which are defined as follows :

$$\mathrm{F}_{\lambda_r\lambda_r} = \mathbf{I} + \sum_{i=1}^{N_s} \mathbb{Q}_{\lambda_r}^{i^T} \sum_{j \in neighbor(i)} \mathbb{T}_r^{i\to j^T} \left(\mathbb{T}_r^{j\to i} - \mathrm{W}_{rr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i}\ \mathbb{D}_r^j K_{rr}^{j^{-1}} \mathbb{D}_r^{j^T} \right) \mathbb{Q}_{\lambda_r}^j \tag{5.6}$$

$$\mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r} = \sum_{n=1}^{N_I} \mathbb{Q}_{\boldsymbol{\lambda}_c}^{n^T} \left(\mathrm{W}_{cr}^{i\leftrightarrow j} \mathbb{T}_r^{j\to i}\ \mathbb{D}_r^j K_{rr}^{j^{-1}} \mathbb{D}_r^{j^T} \right) \mathbb{Q}_{\lambda_r}^j \tag{5.7}$$

$$\mathrm{F}_{\mathbf{E}_c\lambda_r} = \sum_{i=1}^{N_s} \mathbb{Q}_{E_c}^{i^T} \left(K_{cr}^i K_{rr}^{i^{-1}} \mathbb{D}_r^{i^T} \right) \mathbb{Q}_{\lambda_r}^i \tag{5.8}$$

At first, we would like to note that the inversion of the local matrix K_{rr}^i is very expensive in terms of memory and time of computations, even when only some of the columns are required, as for example in $\mathbb{D}_r^i K_{rr}^{i^{-1}} \mathbb{D}_r^{i^T}$. That is why we do not attempt to form the matrix-vector product $x = K_{rr}^{i^{-1}} b$ by computing $K_{rr}^{i^{-1}}$ at first and then perform some multiplications. Instead, we solve the system of linear equations $K_{rr}^i x = b$ for a given x, because it is more efficient and more accurate [125].

Also as we can see from the matrices definition (Eq. 5.6 - 5.8), for every local subdomain Ω^i the term $K_{rr}^{i^{-1}} \mathbb{D}_r^{i^T} \mathbb{Q}_{\lambda_r}^i$ is common. That leads us to the idea of computing this term only one time for every subdomain. Taking this into account we perform the multiplication of the matrices $\mathrm{F}_{\lambda_r\lambda_r}, \mathrm{F}_{\boldsymbol{\lambda}_c\lambda_r}$

and $\mathrm{F}_{\mathbf{E}_c\lambda_r}$ to the matrix λ_r^k in three steps :

```
For $i = 1..N_s$
    Step 1a. Calculate $a^i = \mathbb{D}_r^{i^T} \mathbb{Q}^i_{\lambda_r} \lambda_r^k$
    Step 1b. Solve $K^i_{rr}\ b^i = a^i$ in order to find the term $b^i$
    Step 1c. Make the contribution of $b^i$ into $\delta^k, z_1^k$ and $z_2^k$ using Eq. (5.6 - 5.8)
End
```

Step 3 is performed similarly to Step 1. There is only a difference in the values of the variable a^i at Step 1a.

If we do not take into account the parallel computations, there are two effective ways [97] to implement the solution of the coarse problem (Step 2). In our work, the factorization of the global corner DOFs related system $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ is performed one time on the preprocessing step and it is then kept during all the computations. Such a factorization is used repeatedly in the forward and backward substitutions in the solution of the coarse problem. Note that if we are limited by the memory resources, then the only possibility to calculate a large-scale 3D electromagnetic problem is to use the second way of the coarse problem implementation. Unlike the first approach, in this second way we do not factorize the matrix $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ during the preprocessing step. Moreover, we *do not store* this matrix. This allows us to make the algorithm much more effective in terms of memory requirements, but at the same time it makes the algorithm a bit slower. Indeed it forces us to perform Step 2 in one cycle over all the subdomains. So the reader is highly recommended to use the second approach only in the following two cases :

1. If we can not solve a problem because of the lack of computational resources
2. If we deal with a big number of right-hand sides.

In order to study the gain of these two approaches we will consider the time and memory requirements of the FETI-DPEM2-full method in the following two sections.

5.4.2 Memory requirements

As it was already discussed in Section (5.4.1) we store the factorization of the local matrices K^i_{rr} during all the computations. It is worth mentioning that this step is not mandatory. Indeed, we can factorize all the matrices K^i_{rr} at every iteration one time for Step 1 and one time for Step 3. From one side, this decision would be effective in terms of memory, but from the other - significantly inefficient in terms of computational time. In our realization, we prefer to keep the factorized matrices $L^i_{rr}U^i_{rr} = K^i_{rr}$ in sparse format. Therefore, the memory usage can be estimated as :

$$\text{Memory} \propto M_{\text{Initialization}} + \underbrace{L^i_{rr}U^i_{rr}}_{\forall i\ =\ 1..N_s} + \underbrace{N_c^2}_{\mathrm{F}_{\mathbf{E}_c\mathbf{E}_c}} + \underbrace{N_{\lambda_r}}_{\text{rhs}} \tag{5.9}$$

In the above, $M_{\text{Initialization}}$ represents the memory required to store the solution vectors (optionally), the geometrical matrices, GMRES matrices and temporary arrays. The term $L^i_{rr}U^i_{rr}, \quad \forall i = 1..N_s$ denotes the memory required for storing all the factorized matrices. N_c^2 is the memory required to store the "global" matrix $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$. This number corresponds to the number of DOFs associated to the corners in the power of 2. As we have seen previously in Section (5.4.1.1), this step is an optional one. And finally, the last term represents the memory required to store the right-hand side vector.

5.4.3 Time requirements

In the DDM process, it is obvious that, as the number of subdomains increases, the time associated with the subdomain analysis in the preprocessing step decreases monotonically due to the reduced size of the subdomain problems, whereas the time to solve the interface equation first decreases due to smaller subdomain problems and then increases due to an increasing number of interface and corner unknowns. A practical way for finding an optimal decomposition in terms of computations time is going to be done later on.

Although, the computational time can be estimated as

$$\text{Time} \propto T_{\text{Initialization}} + T_{\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}} + N^{\text{src}} \cdot T_{\text{rhs}} + \sum_{i=1}^{N_s} T^i_{(LU)_{rr}} + T_{\text{GMRES}} \tag{5.10}$$

Here, $T_{\text{Initialization}}$ represents the time spent on the domain decomposition, the creation of the internal interfaces and geometrical matrices. The term $T_{\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}}$ denotes the time spent on the creation and factorization of the matrix $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$. To create the right-hand side vector we spend a time T_{rhs}. This can be explained by applying an algorithm similar to Step 1a - 1c. The term $\sum_{i=1}^{N_s} T^i_{(LU)_{rr}}$ which is equivalent to $N_s \cdot T_{rr}^{\text{average}}$ represents the time spent for the factorization of all the local matrices K^i_{rr}.

The last term takes much more time than all of the first ones. This term denotes the time spent for obtaining the iterative solution of the Interface problem. It can be presented as

$$T_{\text{GMRES}} = N_{\text{iter}} \left[m \left(2 \; N_s \; T_{f\&b} + T^c_{f\&b} \right) \right] \tag{5.11}$$

where, N_{iter} is the number of iterations required to reach the stopping criteria η, m here denotes the restart number of the GMRES method, and the terms $T_{f\&b}$ and $T^c_{f\&b}$ represent the time spent to make one forward and backward substitutions for the Step 1a and 2 respectively.

As it was showed previously, the step corresponding to the term $T_{\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}}$ is optional and can be missed. But in this case the time of GMRES iterations changes as follows :

$$T_{\text{GMRES}} = N_{\text{iter}} \left[m \cdot 3 \; N_s \; T_{f\&b} \right] \tag{5.12}$$

5.5 Comparisons with Fresnel database measurements

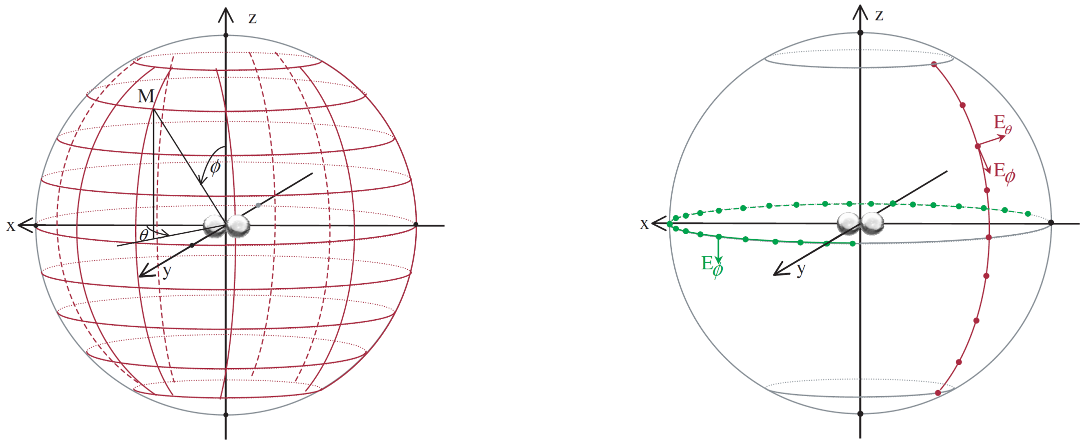
In this section, we are going to compare the results of simulations (FETI-DPEM2-full, FEM-classical) with scattered field measurements [80]. This will allow us to verify the effectiveness of the FETI-DPEM2-full method in practical situations. For each problem, we will provide the size of the domain Ω in terms of the wavelength, all the physical parameters needed for the simulation and the main geometrical criteria both for the classical FEM method and the FETI-DPEM2-full one.

5.5.1 Configuration description

We will not precise the details of the anechoic chamber of the Center Commun de Ressources Micro-Ondes pictured in Figure (5.14) which has been used to perform the measurements, because it has been previously described in several papers [80, 126].

Figure 5.14 – Picture of the anechoic chamber at the Center Commun de Ressources Micro-Ondes (CCRM) which was used to measure the scattered field data of different targets of the Fresnel database.

The sources are located all around the target, at the distance $r = 1.796$ m from the center (Figure 5.15). The azimuthal angle θ_s is ranged from 20° to 360° with a 40° step (i.e. 9 meridians) and the polar angle ϕ_s is ranged from 30° to 150° with a 15° step (i.e. 9 parallels). Due to the design of the system presented above, the receivers are restricted to the azimuthal plane (i.e. $\phi = 90°$) at $r = 1.796$ m from the center and for technical reasons, they can not be closer than 50° from the



(a) Sources

(b) Receivers

FIGURE 5.15 – (a) Positions of the sources and (b) the receivers on the measurement sphere, and definition of the angles and polarizations, with M(r,θ,ϕ) = (1.796, 60, 45).

source meridian. In our simulations, we do not have such problems, thus we calculate the scattered field for an azimuthal angle θ_r ranging from 0° to 360° with a step of 10°. We compare the scattered field calculated with our method only at the angles where the measured fields exist. In this work, the source can be polarized along E_ϕ or E_θ directions, while all the receivers are polarized only along E_ϕ.

As you can see from Figure (5.15) the antennas and receivers are placed far from a target (at $r = 1.796$ m). That is why we only take into account the physical area near the target to reduce the dimension of the problem (Figure 5.16).

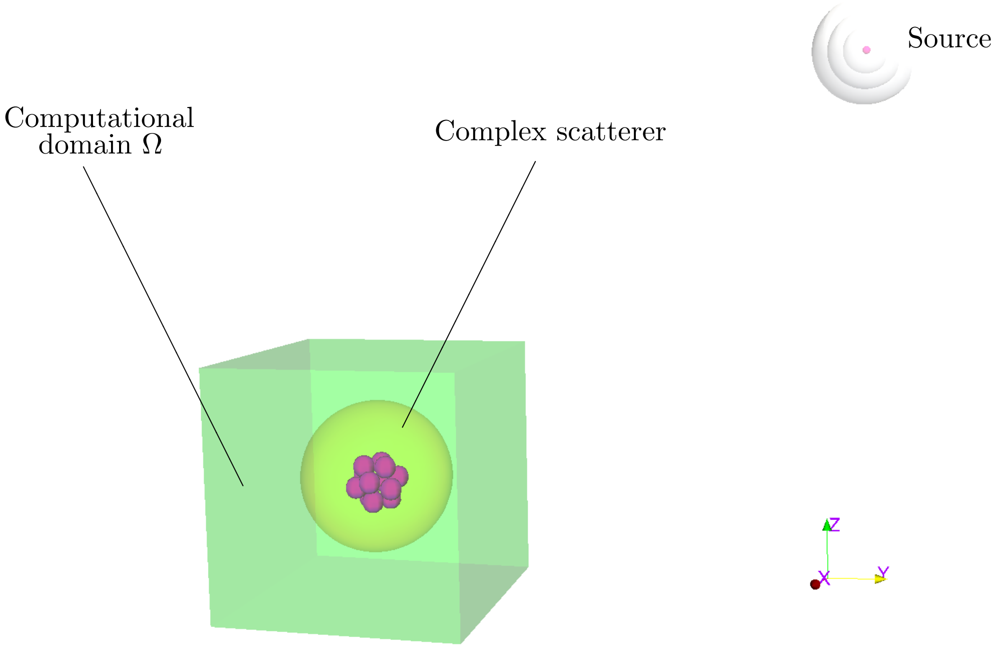


FIGURE 5.16 – Map of the domain Ω under consideration with a target inside.

We assume that the computational domain Ω is filled with air (Figure 5.16). It represents a parallelepiped whose size depends on the operating frequency. At the external boundary, we set the radiation boundary conditions. After finite-element discretization, we obtain a mesh whose element size l depends on the wavelength λ, the parameter κ and the relative permittivity ε_r as Eq. (4.1).

In the next subsections we are going to calculate the scattered field $\mathcal{E}^{\text{sc}}$ created by different targets from the Fresnel database inside the domain Ω. Then we are going to obtain $\mathcal{E}^{\text{sc}}$ outside the domain Ω, in some points where the receivers are placed using the Near-to-Far-Field transformation (Section 2.4). Finally, we are going to compare these results with the measured fields.

For each Scattering problem we calculate *one* average relative L^2-error between the FEM-classical method and the FETI-DPEM2-full one. Also we calculate the average number of iterations for all the sources. To check the data we consider the relative error between the measurements and the simulations as follows :

$$\text{Error} = \frac{1}{N_{src}N_{rec}} \frac{\sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} |\mathcal{E}_{s,r}^{\text{mes}} - \mathcal{E}_{s,r}^{\text{sim}}|^2}{\sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} |\mathcal{E}_{s,r}^{\text{mes}}|^2} \tag{5.13}$$

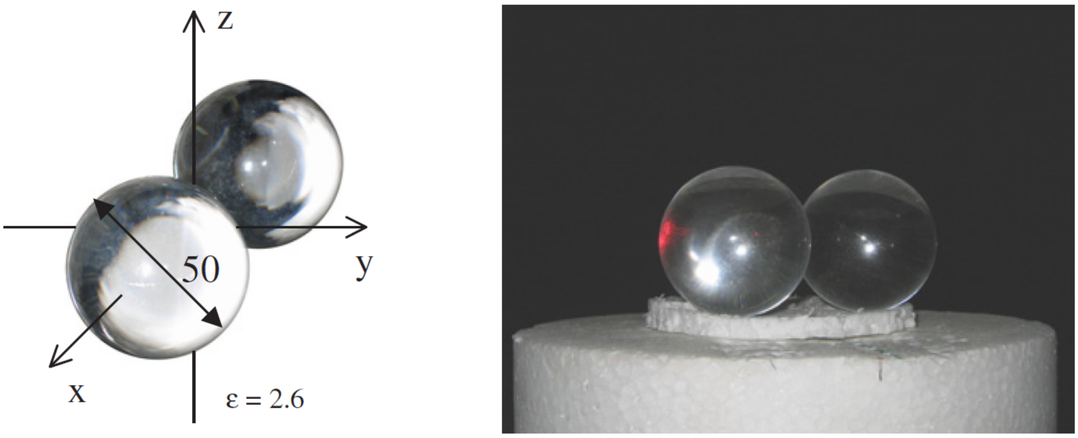
where $\mathcal{E}_{s,r}^{\text{sim}} = \mathcal{E}^{\text{sim}}(s,r,f)$ (resp. $\mathcal{E}_{s,r}^{\text{mes}} = \mathcal{E}^{\text{mes}}(s,r,f)$) denote the simulated (resp. measured) scattered field obtained with the source s and receiver r at the frequency f. The terms N_{src} and N_{rec} are the total number of sources and receivers. The incident field $\mathcal{E}^{\text{inc}}$ is determined as :

$$\mathcal{E}^{\text{inc}} = Ae^{j\phi}e^{j\vec{\text{k}}_s\vec{\text{r}}} \tag{5.14}$$

where A and ϕ are the Amplitude and the Phase of the incident field such that $|\mathcal{E}^{\text{inc}}(0,0,0)| = 1$ and $\arg(\mathcal{E}^{\text{inc}}(0,0,0)) = 0$, as it is precognized in the Fresnel database [80].

5.5.2 TwoSpheres target

This target consists of two dielectric spheres with a diameter of 50 mm and a relative permittivity of 2.6. Their positions are given in Figure (5.17).



(a) (b)

Figure 5.17 – (a) Diagram and (b) picture of the TwoSpheres target on the polystyrene mast in the anechoic chamber.

To start with, we set the working frequency to 4 GHz which corresponds to $\lambda = 0.075$ m. The domain Ω (which for such wavelength is a parallelepiped with a size of $\approx 3.3 \times 3.3 \times 2.7\ \lambda^3$) has been discretized into finite elements (Table 5.14) and then, following the idea of the FETI-DPEM method, has been divided into 25 subdomains (Table 5.15).

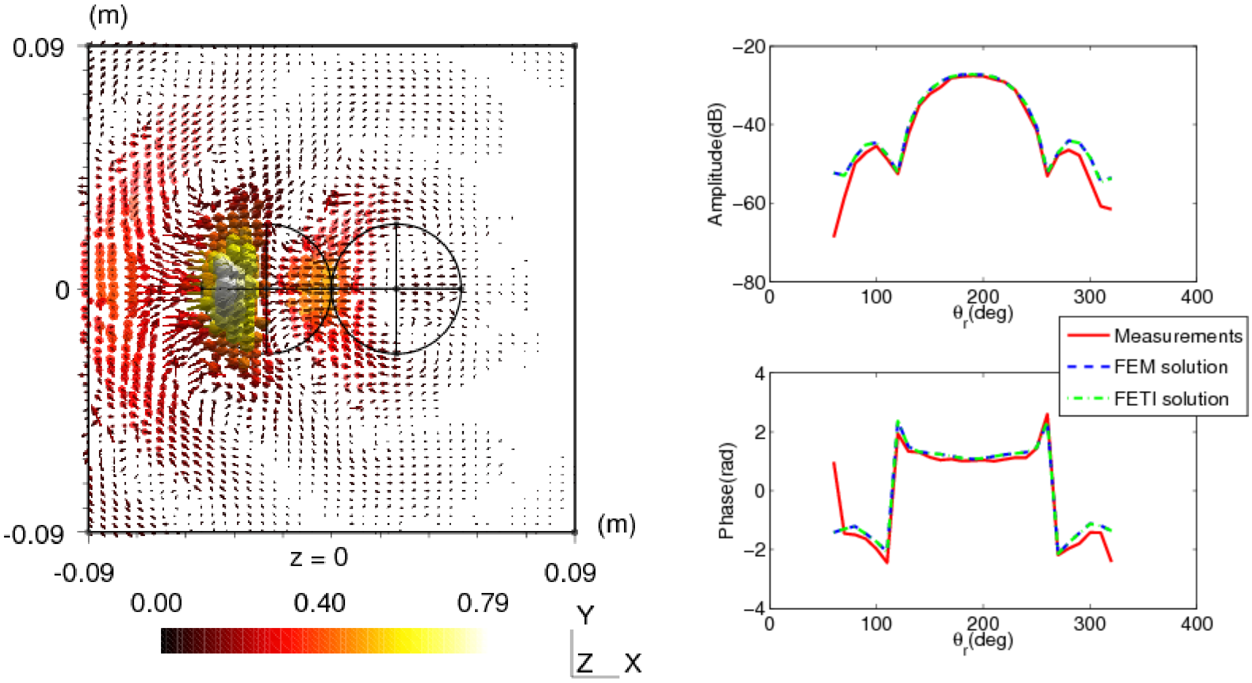
κ	N_p	N_t	N_e
10	26 561	144 981	168 743

TABLE 5.14 – Finite element discretization of the domain with the TwoSpheres target which corresponds to 10 points per wavelength.

N_s	N_{E_c}	N_ξ	N_{λ_r}
25	1 282	3 609	41 350

TABLE 5.15 – Principal decomposition parameters of the domain with the TwoSpheres target which corresponds to a partitioning into 25 subdomains.

Obviously, we calculate this problem for all the source orientations, treating them as a Multiple right-hand side. Nevertheless we are going to present the results obtained with a single target orientation and a single source $\theta_s = 30°$. This problem is calculated with the help of the classical FEM method and the FETI-DPEM2-full method. As previously, we use the normal GMRES parameters (Eq. 5.3) for solving the reduced Interface Problem. To have a good vision of the scattered field inside the domain, we plot its imaginary part obtained with the FETI-DPEM2-full method on the slice $z = 0$ (Figure 5.18a).



(a) (b)

FIGURE 5.18 – (a) $\mathcal{I}m(\mathcal{E}^{\mathrm{sc}})$ created by the source at the angle $\theta_s = 30°$ with the reflection of the TwoSperes target and (b) $\mathcal{E}^{\mathrm{sc}}$ obtained in the $\theta\phi$ polarization with a single source $\theta = 30°$, but all receivers positions at 4GHz. Scattered field magnitudes (up) and phases (down).

Thanks to the database of Institut Fresnel, we can compare the results obtained with the FETI-DPEM2-full method with the measurements. We can see from Figure (5.18b) that the measurements and results obtained with the FETI-DPEM2-full method are in good agreement.

Thanks to the working frequency range ([4-8] GHz, corresponding to wavelengths spanning from 3.75 to 7.5 cm) of the previously introduced microwave device and the Fresnel database, we can verify the results of the proposed method by applying the frequencies $4-8$ GHz with a step of 1 GHz to the TwoSpheres target. As we can see from Table (5.16) the absence of PML-area leads us to the perfect convergence that was discussed previously in Section (5.3.5.2). This error with respect to the measurements is higher than the one obtained when comparing FEM and FETI-DPEM2-full fields.

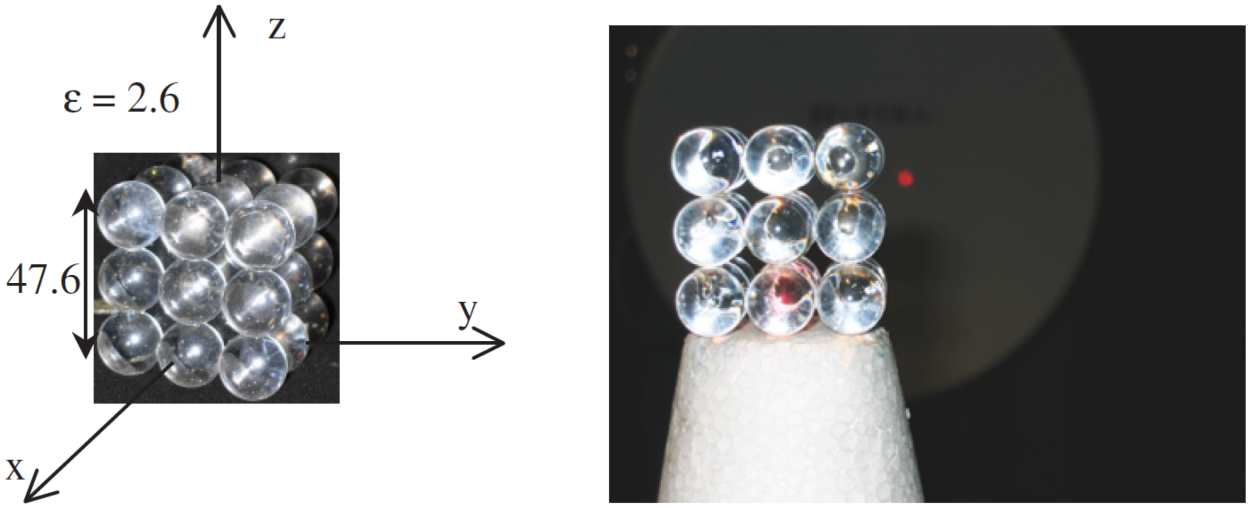
f (GHz)	Size of Ω in λ^3	N_e	N_{λ_r}	L^2-error (FEM)	L^2-error (measure)	N_{iter}
4	$3.3 \times 3.3 \times 2.7$	168 743	41 350	8.6804E-003	0.3964	8
5	$3.7 \times 3.7 \times 2.8$	233 218	53 652	1.1114E-002	0.5477	8
6	$4.0 \times 4.0 \times 3.0$	291 182	59 066	6.5795E-003	0.4805	9
7	$4.3 \times 4.3 \times 3.2$	379 087	69 640	1.7038E-002	0.4326	8
8	$4.7 \times 4.7 \times 3.3$	461 155	83 836	8.1210E-003	0.5295	9

Table 5.16 – Results of simulations with the FETI-DPEM2-full method at various frequencies for the TwoSpheres target.

This is due to the fact that we have not taken into account the measurement error bars during the comparison with our L^2 criterion.

5.5.3 Cube of spheres target

This target consists of an aggregate of dielectric spheres. Each sphere has a diameter of 15.9 mm and a permittivity of 2.6. They were assembled so as to obtain a cube measuring 47.6 mm on each side. The positions of this target is shown in Figure (5.19). Due to the size of the spheres and their arrangement, this target has the finest geometrical details in the Fresnel database.



(a) (b)

Figure 5.19 – (a) Diagram and (b) picture of the CubeSpheres target on the polystyrene mast in the anechoic chamber.

To start with, we set the working frequency to 4 GHz which corresponds to $\lambda = 0.075$ m. The domain Ω (which for such a wavelength is a cube with a size of $\approx 2.5\lambda \times 2.5\lambda \times 2.5\lambda$) has been discretized into finite elements (Table 5.17) and then, following the idea of the FETI-DPEM method, has been divided into 7 subdomains (Table 5.18).

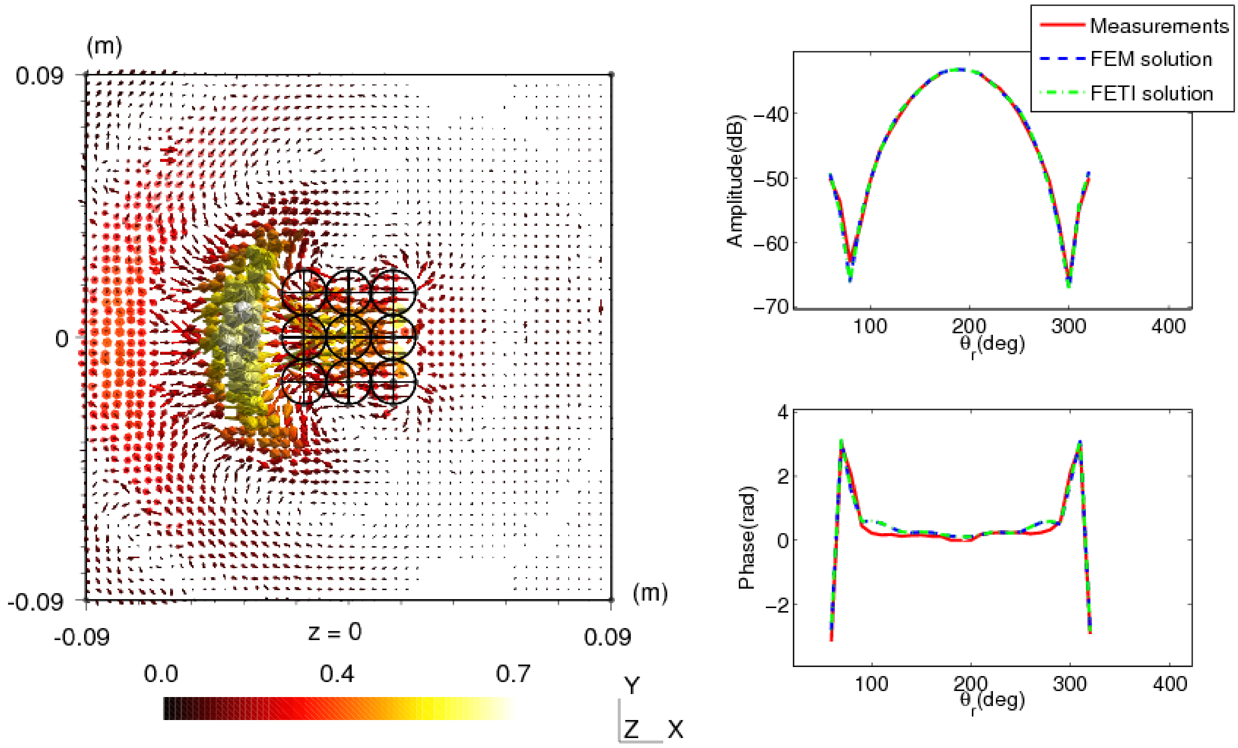
κ	N_p	N_t	N_e
10	60 430	362 855	409 215

TABLE 5.17 – Finite element discretization of the domain with the CubeSpheres target which corresponds to 10 points per wavelength.

N_s	N_{E_c}	N_ξ	N_{λ_r}
7	326	929	45 590

TABLE 5.18 – Principal decomposition parameters of the domain with the CubeSpheres target which corresponds to a partitioning into 7 subdomains.

As previously, we calculate this problem for all the source orientations, treating them as a Multiple right-hand side. For this scatterer we are going to present only the results obtained with a single target orientation and a single source $\theta_s = 30°$, as we have done previously. For the CubeSphere target we present the results obtained with the classical FEM method and the FETI-DPEM2-full. As previously, we use the normal GMRES parameters (Eq. 5.3) for solving the reduced Interface Problem. To have a good vision of the scattered field inside the domain we plot its Imaginary part obtained with the FETI-DPEM2-full method on the slice $z = 0$ (Figure 5.20a).



(a) (b)

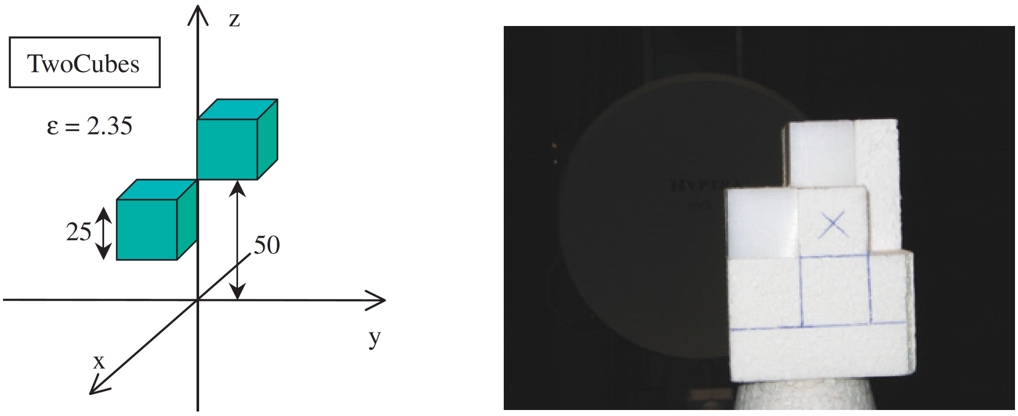
FIGURE 5.20 – (a) $\mathcal{I}m(\mathcal{E}^{\mathrm{sc}})$ created by the source at the angle $\theta_s = 30°$ with the reflection of the CubeSpheres target and (b) $\mathcal{E}^{\mathrm{sc}}$ obtained in the $\theta\phi$ polarization with a single source $\theta = 30°$, but all receivers positions at 4GHz. Scattered field magnitudes (up) and scattered field phases (down).

In Figure (5.20b) we present the scattered far-field obtained with both the FEM and FETI methods at the frequency of 4 GHz compared to the measured field at the angle $\theta_s = 30°$. We can see from this figure that the measurements and results obtained with the proposed method are in a good agreement.

The relative L^2-error (Eq. 5.13) for all the sources is 0.2778. Also we control the results of the method FETI-DPEM2-full with respect to the classical FEM and find a relative L^2-error of 9.1311E-003. The average number of iterations for every source is 8.

5.5.4 TwoCubes object

The last target from the Fresnel database considered in this chapter is called TwoCubes. This object consists of two dielectric cubes measuring 0.025 m on each side with a permittivity of 2.35. The position of this scatterer is shown in Figure (5.21)



(a) (b)

FIGURE 5.21 – (a) Diagram and (b) picture of the TwoCubes target on the polystyrene mast in the anechoic chamber.

As previously, we set the working frequency to 4 GHz and the domain Ω (which for such wavelength is a cube with a size of $\approx 3.1\lambda \times 3.1\lambda \times 3.1\lambda$) has been discretized into finite elements (Table 5.19) and then, following the idea of the FETI-DPEM method, has been divided into 65 subdomains (Table 5.20).

κ	N_p	N_t	N_e
16	91 238	553 678	626 320

TABLE 5.19 – Finite element discretization of the domain with the TwoCubes target which corresponds to 16 points per wavelength.

N_s	N_{E_c}	N_ξ	N_{λ_r}
65	4 732	13 320	152 834

TABLE 5.20 – Principal decomposition parameters of the domain with the TwoCubes target which corresponds to a partitioning into 65 subdomains.

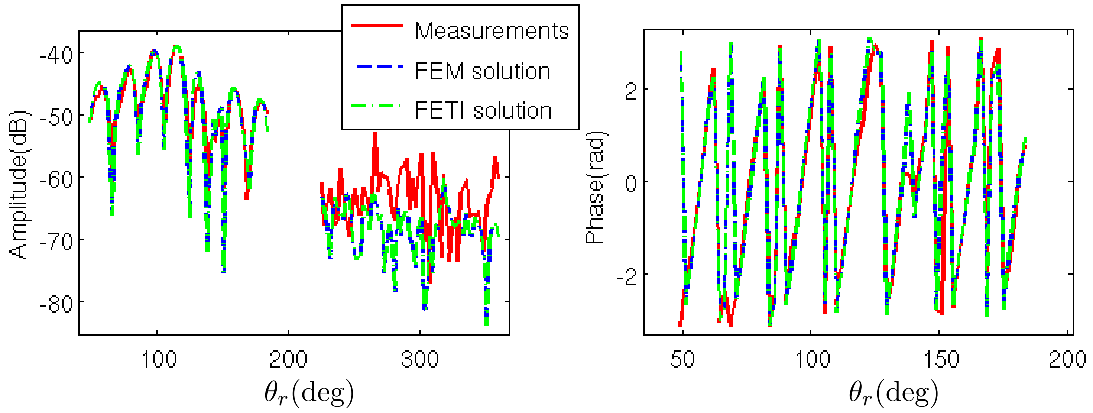


FIGURE 5.22 – $\mathcal{E}^{\mathrm{sc}}$ obtained in the $\theta\phi$ polarization with a single source at the angles $\theta_s = 30°$ at the frequency $f = 4$ GHz. Scattered field magnitudes (left) and scattered field phases (right).

The results obtained with the FEM classical and FETI-DPEM2-full methods are presented in Figure (5.22) with respect to the measurements. We can see that the scattered far-field $\mathcal{E}^{\mathrm{sc}}$ calculated with both the FEM and FETI methods are in good agreement with the measurements performed in the Anechoic chamber.

The relative L^2-error presented in Eq. (5.13) between the measurements and the scattered field obtained with the FETI-DPEM2-full method is equal to 0.3107. The relative L^2-error between the FEM and FETI methods is equal to 3.3138E-004. It took only 9 iterations for the GMRES method to reach this level of error.

5.6 Towards ellipsoid targets

In this section we are going to present some simulated scattered fields generated by ellipsoid targets. This problem becomes challenging in terms of required memory and computations time, as the size of the ellipsoids is much bigger than the wavelength. That is why it is of great interest to test the proposed FETI-DPEM2-full method on such problems.

We use the following source-receivers configuration : there is one source radiating in free space placed in the azimuthal plane at $\theta_s = 0°,\ 45°$ and $90°$ with 2 different polarizations (E_θ and E_ϕ) at the same distance $r = 1.796$ m. The receivers are placed in the azimuthal plane at $\theta_r = 1°$ to $360°$ with a step of 1°. For the given configuration we distinguish 6 different test-cases (Table 5.21)

Test-case	θ_s	Polarisation
TC1	0°	E_θ
TC2	45°	E_θ
TC3	90°	E_θ
TC4	0°	E_ϕ
TC5	45°	E_ϕ
TC6	90°	E_ϕ

Table 5.21 – 6 different test-cases considered for the superellipsoid targets.

5.6.1 One ellipsoid target

This target is an ellipsoid with diameters $D_x,\ D_y$ and D_z of $10,\ 15$ and 10 cm respectively. The relative permittivity ε_r of the ellipsoid is equal to 1.45 which corresponds to foam [127]. The domain Ω is a parallelepiped whose size depends on the wavelength. For the given problem we set the working frequency to $f = 4$ GHz which corresponds to $\lambda = 0.075$ m. The domain Ω (which for such wavelength is a parallelepiped with a size of $0.212 \times 0.262 \times 0.212$ m^3 $(3\lambda \times 3.5\lambda \times 3\lambda)$) has been discretized into finite elements (Table 5.22) and then, following the idea of the FETI-DPEM method, has been divided into 15 subdomains (Table 5.23).

κ	N_p	N_t	N_e
15	74 168	451 920	511 834

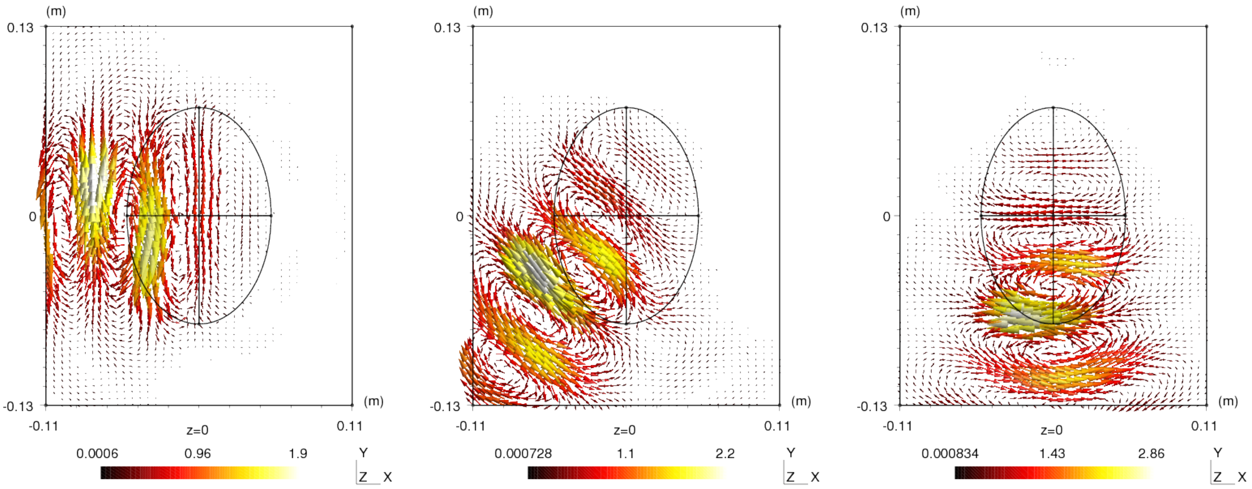
Table 5.22 – Finite element discretization of the problem presented in Section (5.6.1) at the frequency of 4 GHz which corresponds to 15 points per wavelength.

N_s	N_{E_c}	N_ξ	N_{λ_r}
15	1 015	2 868	69 726

Table 5.23 – Principal decomposition parameters of the domain with the Ellipsoid at 4 GHz, which corresponds to a partitioning into 15 subdomains.

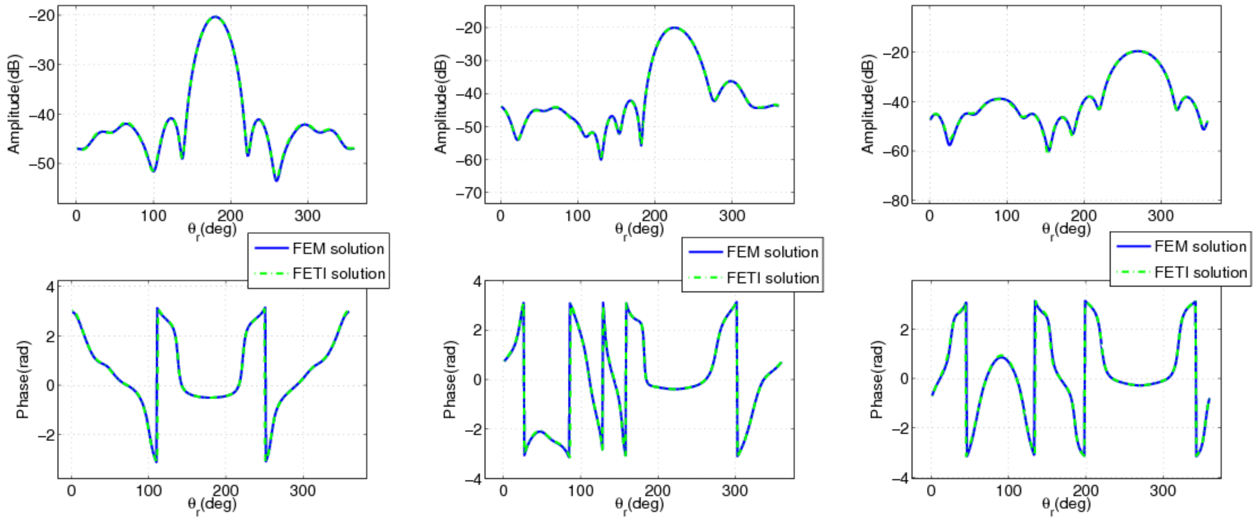
We then calculate this problem for all the source orientations (TC1 .. TC6) with the classical FEM method and the method FETI-DPEM2-full using the domain decomposition presented above.

To have a good vision of the scattered field inside the domain, we plot the Imaginary part of $\mathcal{E}^{\mathrm{sc}}$ obtained with the FETI-DPEM2-full method on the slice $z = 0$ (Figure 5.23) for 3 positions of the source, but only for the polarization E_θ (TC1, TC2 and TC3). Looking at the Amplitude and the Phase of the scattered field presented in Figure (5.24) we can conclude that the results of the FEM method and the method FETI-DPEM2-full are in a good agreement. From a physical point of view, we have good confidence that these results are correct. Indeed, the obtained solution is symmetric for the angles $\theta_s = 0°$ and $90°$.



(a) Test-case 1 (b) Test-case 2 (c) Test-case 3

FIGURE 5.23 – $\mathcal{I}m(\mathcal{E}^{\mathrm{sc}})$ created by the source at the angles (a) $\theta_s = 0°$, (b) $\theta_s = 45°$ and (c) $\theta_s = 90°$ at the frequency $f = 4$ GHz in presence of the Ellipsoid target. The computations are performed with the FETI-DPEM2-full method.

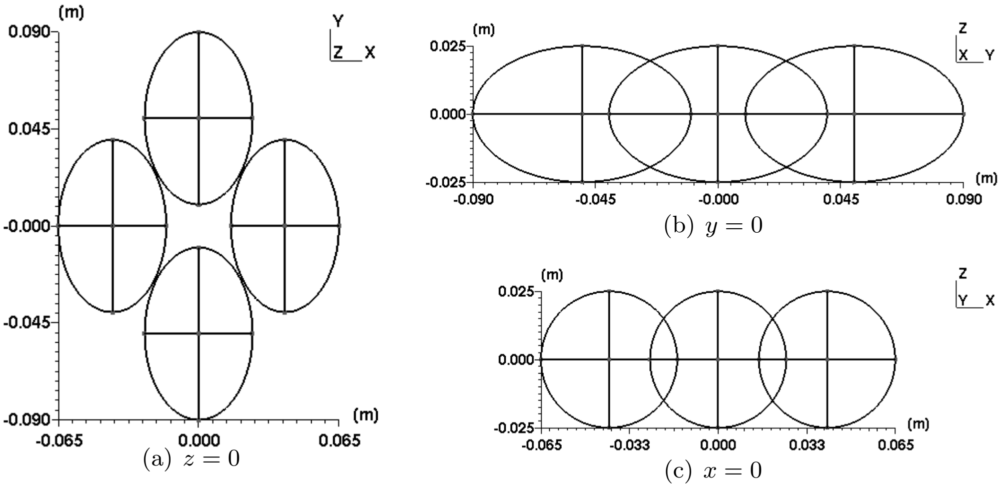


(a) Test-case 1 (b) Test-case 2 (c) Test-case 3

FIGURE 5.24 – $\mathcal{E}^{\mathrm{sc}}$ obtained in the $\theta\phi$ polarization with a single source at the angles (a) $\theta_s = 0°$, (b) $\theta_s = 45°$ and (c) $\theta_s = 90°$ at the frequency $f = 4$ GHz. scattered field magnitudes (up) and scattered field phases (down).

5.6.2 Four ellipsoids target

This target consists of four dielectric ellipsoids with diameters D_x, D_y and D_z of 5, 8 and 5 cm respectively. Their positions are given in Figure (5.25). The relative permittivity ε_r of each ellipsoid is equal to 1.45 which corresponds to foam [127]. For this problem, we use the same source-receiver configuration as previously detailed. The domain Ω is a parallelepiped whose size depends on the wavelength.



(a) $z = 0$

(b) $y = 0$

(c) $x = 0$

FIGURE 5.25 – Schematic map of the 4 ellipsoids target inside the domain Ω.

As compared to the previous target considered in Section (5.6.1) the 4 ellipsoids take much more place and, evidently more memory required for calculations. In this case solving such problems becomes difficult even with the FETI methods. So far we did not talk about the *optimal* domain decomposition in terms of the required memory or computational time. In this subsection we are interested to find it.

Let us set the operating frequency equal to 8 GHz which corresponds to $\lambda = 0.0375$ m. In this case the size of the domain Ω is $0.1862 \times 0.236 \times 0.1062$ m^3 ($4.97\lambda \times 6.29\lambda \times 2.83\lambda$). The resulting finite-element discretization can be fully described by Table (5.24)

κ	N_p	N_t	N_e
15	219 428	1 360 897	1 548 019

TABLE 5.24 – Finite element discretization of the problem presented in Section (5.6.2) at the frequency of 8 GHz which corresponds to 15 points per wavelength.

In this simulation, we decompose the computational domain into $N_s = 20...250$ subdomains in order to find an optimal decomposition. The results of the different domain decompositions in terms of computational time and memory requirement are presented in Figure (5.26).

It can be seen from this figure that the computational time is not very sensitive to the number of subdomains around the optimal one, whereas the memory required for the computations increases very rapidly after the optimal point. Also, we can conclude that the optimal number of subdomains in terms of memory is not the same as for the time of computations. In this case we need to find a compromise. In our opinion the general rule to determine the optimal number of subdomains in terms of memory is to divide the total number of FEM unknowns by a number between 10 000 and 15 000 which corresponds to the approximately number of unknowns in each subdomain, but in terms of computational time we should use a coefficient between 6 500 and 8 000 which agrees with [97].

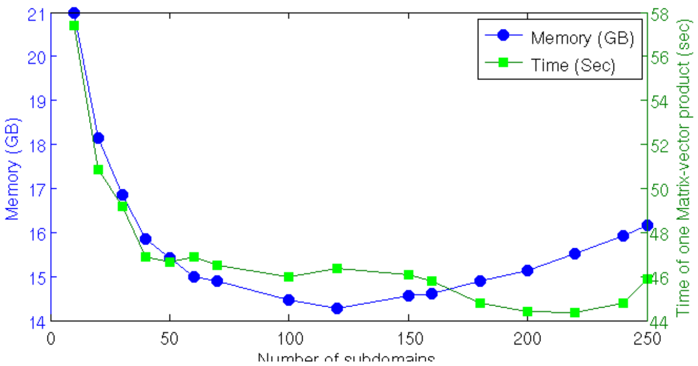
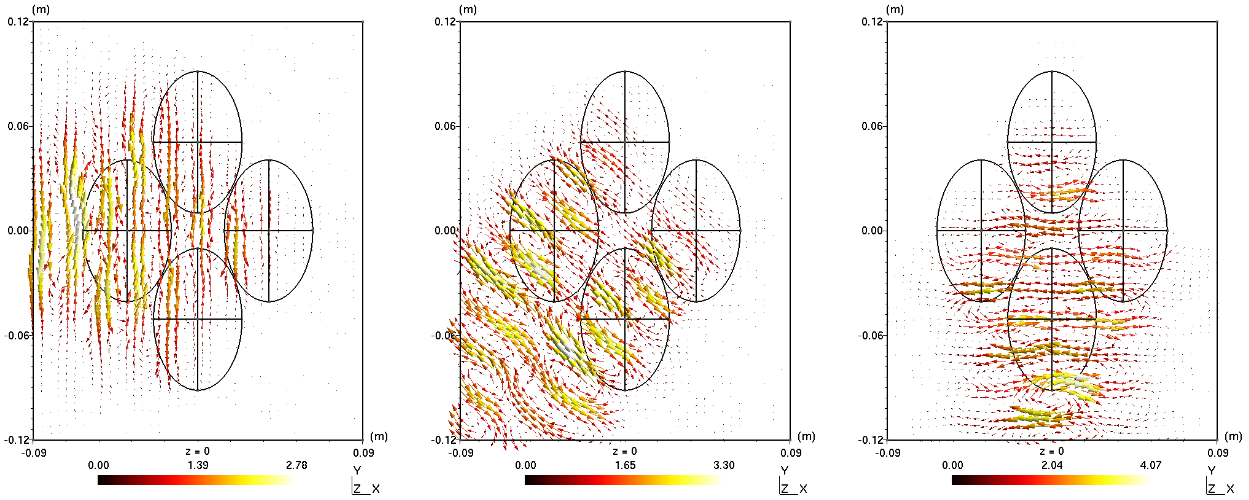


FIGURE 5.26 – Memory required to calculate the problem of 4 ellipsoids and the time of one matrix-matrix multiplication when different partitioning are applied.

The investigation domain Ω is divided then into 170 subdomains which corresponds to a compromise in terms of memory and computational time. As previously, in order to have a good vision of the scattered field inside the domain, we plot its Imaginary part obtained with the FETI-DPEM2-full method on the slice $z = 0$ (Figure 5.27).



(a) Test case 1 (b) Test case 2 (c) Test case 3

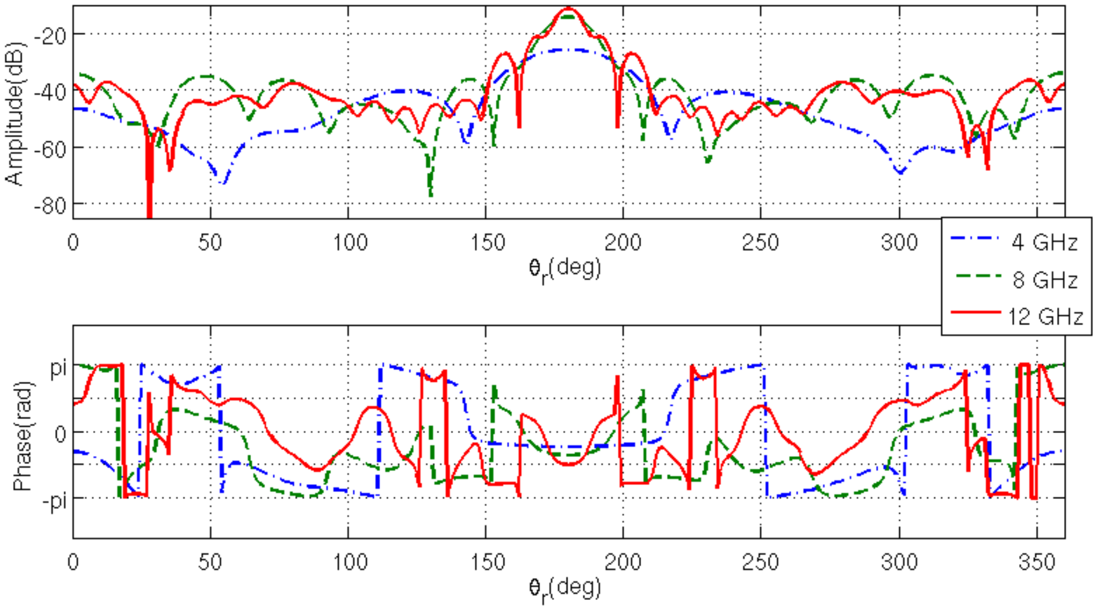
FIGURE 5.27 – $\mathcal{I}m(\mathcal{E}^{\text{sc}})$ created by the source at the angles (a) $\theta_s = 0°$, (b) $\theta_s = 45°$ and (c) $\theta_s = 90°$ at the frequency $f = 8$ GHz in the presence of the 4-Ellipsoids target.

By varying then the working frequencies between 4 and 12 GHz (Figure 5.28), we can conclude that the classical FEM method is effective until 6 GHz for this problem statement on our computer configuration, whereas with the FETI-DPEM2-full method we have been able to calculate all the tasks in this range of frequency (Table 5.25). All computations have been performed on an Intel(R) Xeon(R) CPU X5570 @ 2.93GHz, with 48 GB of RAM, with no parallel programming specificities.

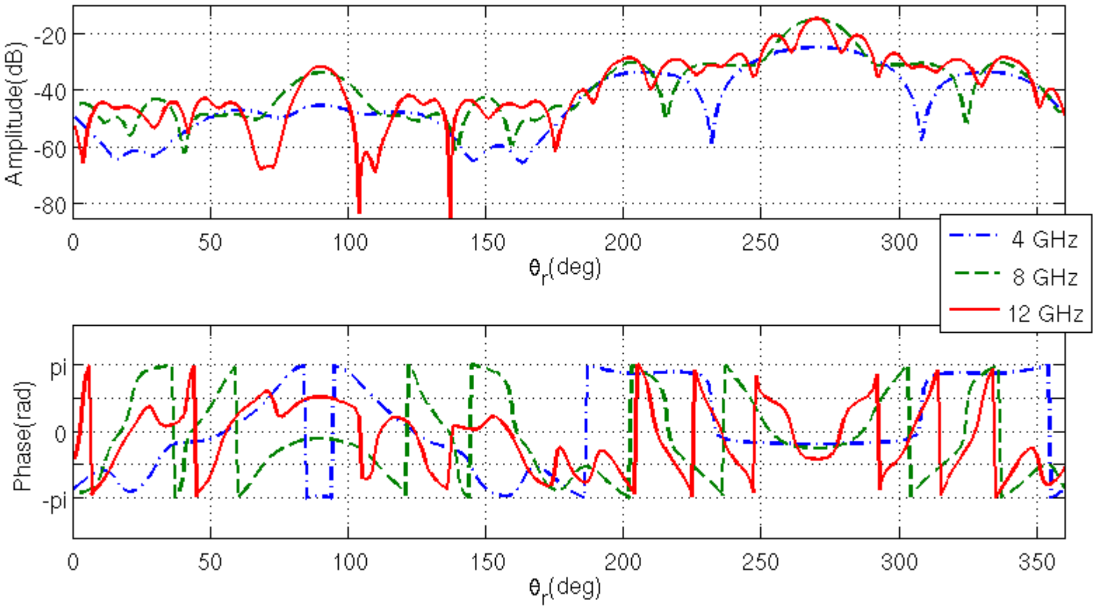
We suppose that these results are correct from a physical point of view, because all of them are symmetric for the angles $\theta_s = 0°$ and 90° (Figure 5.28).

f (GHz)	Size of Ω in λ^3	N_e	N_{λ_r}	Optimal N_s	L^2-error (FEM)	Time of one Fx (Eq. 5.5)	N_{iter}
4	$3.23 \times 3.89 \times 2.17$	475 017	114 636	50	8.8774E-003	13.42	9
5	$3.67 \times 4.50 \times 2.33$	593 046	146 474	62	1.5437E-002	16.72	9
6	$4.12 \times 5.12 \times 2.5$	926 761	231 134	97	7.6544E-003	22.43	8
7	$4.53 \times 5.69 \times 2.66$	1 205 349	313 262	126	–	31.11	8
8	$4.97 \times 6.29 \times 2.83$	1 548 019	398 918	162	–	45.88	9
9	$5.40 \times 6.90 \times 3.00$	1 951 464	519 414	205	–	59.02	10
10	$5.83 \times 7.47 \times 3.17$	2 365 405	637 188	248	–	73.81	10
11	$6.27 \times 8.07 \times 3.34$	2 789 473	754 008	293	–	85.07	10
12	$6.70 \times 8.72 \times 3.50$	3 010 219	805 682	316	–	99.79	9

Table 5.25 – Results of simulations with the FETI-DPEM2-full method at various frequencies, for the 4 ellipsoids target.



(a) Test case 1



(b) Test case 3

FIGURE 5.28 – Im($\mathcal{E}^{\mathrm{sc}}$) created by the source at the angles (a) $\theta_s = 0°$, (b) $\theta_s = 90°$ at the frequencies $f = 4$ GHz, 8 GHz and 12 GHz in the presence of the 4-Ellipsoids target. Scattered field magnitudes (up) and scattered field phases (down).

5.7 Conclusion

In this chapter, we have presented the implementation of the FETI-DPEM2-full method for generic 3D electromagnetic scattering problems. Indeed, a new set of Lagrange multipliers have been added in order to handle differently the continuity conditions at the corner edges between the various subdomains. Moreover, we have been able to partition and to handle internal interfaces which are not necessarily straight lines without major drawback in terms of computational accuracy.

Due to the increasing size of the underlying linear systems, classical iterative methods have been investigated to solve the interface problem. The numerical results presented here have shown that the convergence speed of the iterative method is seriously affected by the presence of the PML and the transparency condition between the internal interfaces. A more accurate approximation of the Dirichlet-to-Neumann operator has thus been investigated. We have been able to numerically show that the use of EMDA strongly enhances the convergence process when PML materials are used in order to mimic open space configuration. Moreover, this approach presents no restriction with respect to heterogeneous and/or anisotropic scatterers. Although, it is concluded in Section (5.3.5.2) that EMDA is outperformed by the Robin-type boundary condition with the classical coefficient $\alpha^i = jk_0$ [123] when we deal with the isotropic media.

So far, we have only investigated a sequential implementation of the FETI-DPEM2-full method for addressing three-dimensional large-scale direct electromagnetic problems. In the future, it would be thus of interest to investigate its parallel implementation.

Chapitre 6

3D Quantitative Inverse Problems

Contents

6.1 Introduction

As we have discussed in the general introduction, quantitative inverse scattering algorithms attempt to estimate from scattering experiments the physical parameters and features (position, form, size and complex permittivity) of a target. We can point out two main challenges which are associated with quantitative microwave imagining. These are ill-posedness of the electromagnetic inverse scattering problem and its non-linearity. The former jeopardizes the robustness of the reconstruction algorithms and the quality of the results and the latter results in a high computational cost because it requires an iterative optimisation into which the forward solver plays a key role.

Various works take profit of the finite element method in order to solve inverse problems in different scientific domains, such as optical imaging [128–130], electroencephalography imaging [131], electrocardiographic imaging [132], elasticity imaging [133–135], electrical impedance tomography [129, 136, 137], electrical capacitance imaging [138], eddy-current imaging [139] and, of coarse, in microwave imaging [140–147]. Very few take profit from the FETI-DPEM methods to handle large-scale problems, apart eventually [148] where a two-dimensional configuration is investigated.

In this work, we focus on the three-dimensional case, because the computational burden here is more important than in 2D and leads us to the necessity of well-defined numerical strategies developments. As we have set our sights on quantitative imagining algorithms, which attempt to recover the value of the complex permittivity in every point of a given investigation domain, without any use of *a-priori* information, the number of the reconstruction unknowns is increasing drastically. We thus must tune the FETI-DPEM2-full method in order to make this process more flexible in terms of required memory.

As inverse scattering problems are ill posed, they are thus very sensitive with respect to the signal-to-noise ratio [126]. It is thus even more of interest to provide robust inversion algorithms. In particular, we can point out a Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton optimization algorithm with line search [77, 126, 149], which distinguishes itself through an approximation of the Hessian matrix in the Newton correction step with a matrix that does not involve the explicit computation of second order derivatives. Taking advantage of the Lagrangian formalism and the definition of an ad-hoc adjoint field, we can calculate the gradients with respect to the permittivity. Thus, we obtain an iterative algorithm which involves two full forward scattering problems at each iteration of the inversion procedure. Again, we will search for effective numerical strategies while implementing the FETI-DPEM2 method in the inversion scheme.

Various solution methods have been proposed and will be detailed afterwards. They are validated against experimental data acquired with real world targets and well-known from the international inverse problem community as they serve as database for testing their algorithms [80, 81].

6.2 Quantitative inverse scattering problem

Consider an isotropic inhomogeneous dielectric object embedded in an isotropic, possibly inhomogeneous background with permittivity $\varepsilon_b(\vec{r})$ (Figure 6.1). The embedding medium is assumed to be infinite and non-magnetic, as well as the object under consideration. Thus, its interaction with electromagnetic fields is entirely determined by its complex permittivity ε, which is function of the 3D position vector $\vec{r}$.

$$\varepsilon(\vec{r}) = \varepsilon_0 \varepsilon_r(\vec{r}) \tag{6.1}$$

where ε_0 is the permittivity of vacuum and $\varepsilon_r(\vec{r})$ represents the relative permittivity in the object.

In this setting, the goal of the electromagnetic inverse scattering problem is to reconstruct the complex relative permittivity function $\varepsilon_r(\vec{r})$ from multi-view scattering data within a bounded domain $\mathcal{D}$ which contains the unknown object. The form and the shape of the investigation domain $\mathcal{D}$ can vary from simplest (cube, sphere, ...) to more complex ones (union of few objects). This knowledge can be obtained from the resolution of the qualitative inverse problem for example [150]. The principle of collecting multi-view scattering data was explained previously (in Section 5.5). [1]

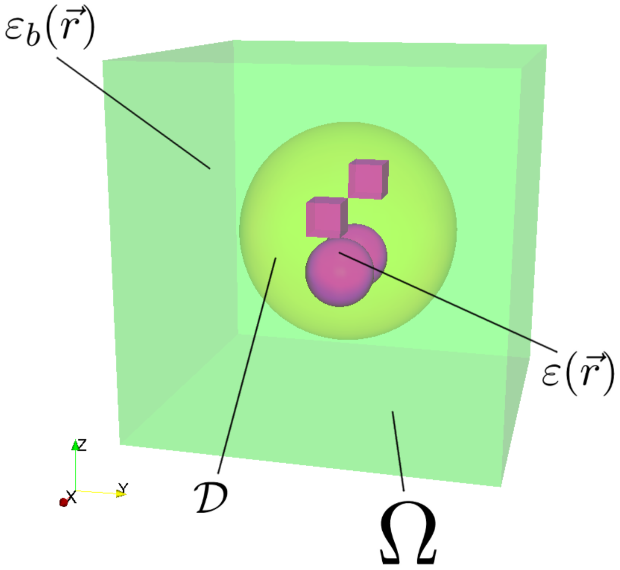


FIGURE 6.1 – Geometry of the problem. A three-dimensional set of objects inside a investigation test domain $\mathcal{D}$ and a computational domain Ω.

From the knowledge of the measured scattered field $\mathcal{E}^{\text{mes}}_{s,r} = \vec{\mathcal{E}}^{\text{mes}}(s,r)$ for $s = 1, ..., N_{src}$ and $r = 1, ..., N_{rec}$, the inverse scattering problem is formulated as an optimization problem (Eq. 6.2), where the least squares data fit cost functional $\mathcal{J}(\varepsilon_r)$ is evaluated for the given measured field $\mathcal{E}^{\text{mes}}_{s,r}$ and the calculated scattered *far-field*[2] $\mathcal{E}^{\text{far}}_{s,r}(\varepsilon_r) = \vec{\mathcal{E}}^{\text{far}}(s, r, \varepsilon_r)$ for an updated permittivity function ε_r.

$$\mathcal{J}(\varepsilon_r) = ||\mathcal{E}^{\text{mes}}_{s,r} - \mathcal{E}^{\text{far}}_{s,r}(\varepsilon_r)||^2_W = \frac{1}{2} \sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} w_{s,r} |\mathcal{E}^{\text{mes}}_{s,r} - \mathcal{E}^{\text{far}}_{s,r}(\varepsilon_r)|^2 \tag{6.2}$$

---

[1] The multi-view scattering data are collected as follows : the investigation test domain $\mathcal{D}$ is illuminated with N_{src} number of incident fields $\mathcal{E}^{\text{inc}}_s$ ($s = 1, ..., N_{src}$). For each such illumination s, the scattered field is measured in a number of measurement positions $r = 1, ..., N_{rec}$ which can differ for different s.

[2] We use the term scattered far-field $\mathcal{E}^{\text{far}}_{s,r}(\varepsilon_r)$ as the sources and receivers are located outside of the investigation domain Ω.

where the matrix w corresponds to appropriate weighting coefficients which are linked to the *experimental noise* [151]. Such cost function assumes white Gaussian noise, which is the case in our free configuration [151]. In order to obtain the field outside the investigated domain Ω from the knowledge of the scattered field $\mathcal{E}^{\text{sc}} = \vec{\mathcal{E}}^{\text{sc}}(\vec{r})$, in the positions of antennas, we use the Near-to-Far-Field transformation presented in Section (2.4), which will be denoted as $\mathcal{F}$:

$$\begin{aligned}\mathcal{F}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}) = \mathcal{E}^{\text{far}}(r) - \oiint_{\Sigma} \{ & -jw\mu \left[\hat{n}' \times \nabla \times \mathcal{E}^{\text{sc}}(r')\right] G_0(r, r') \\ & + \left[\hat{n}' \cdot \mathcal{E}^{\text{sc}}(r')\right] \nabla' G_0(r, r') + \left[\hat{n}' \times \mathcal{E}^{\text{sc}}(r')\right] \times \nabla' G_0(r, r') \} \, dS'\end{aligned} \tag{6.3}$$

As we use the radiation boundary condition, we can take $\Sigma = \partial\Omega$, which is the external boundary of the investigation domain Ω. We could also have only used the field back-propagated on a close surface surrounding the domain, i.e. performing a far-to-near-field transformation. Unfortunately, this linear operator is ill-posed and requires additional regularization strategy. Thus, we prefer to compute the field locally, propagate it to far-field and perform the comparison at the receivers locations.

In the equation above, the calculated scattered field $\mathcal{E}^{\text{sc}}$ created by secondary sources $J^{\text{sc}} = \vec{J}^{\text{sc}}(\vec{r})$ and the unknown permittivity $\varepsilon_r(\vec{r})$ are related by Helmholtz equation (6.4) which will be denoted as $\mathcal{H}$:

$$\mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r) = \nabla \times \left(\frac{1}{\mu_r} \nabla \times \mathcal{E}^{\text{sc}} \right) - k_0^2 \varepsilon_r \mathcal{E}^{\text{sc}} \text{ in } \Omega \tag{6.4}$$

In order to take into account the constraints (6.3) and (6.4) we introduce a following Lagrange functional :

$$\begin{aligned}&\mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = \\ &\sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} \left(w_{r,c} \left| \mathcal{E}^{\text{mes}}_{s,r} - \mathcal{E}^{\text{far}}_{s,r}(\varepsilon_r) \right|^2 + \operatorname{Re} \left\langle \mathcal{U}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}}) \right\rangle_{\Gamma^{\text{far}}} + \operatorname{Re} \left\langle \mathcal{P}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r) - J^{\text{sc}} \right\rangle_{\Omega} \right)\end{aligned} \tag{6.5}$$

where the added extra term represents the constraints of Eqs. (6.3) and (6.4). When $\mathcal{U} = \vec{\mathcal{U}}(\vec{r})$ and $\mathcal{P} = \vec{\mathcal{P}}(\vec{r})$ are the correct Lagrange multipliers, then the saddle point of Eq. (6.5) satisfies Eqs. (6.3) and (6.4).

The saddle point of Eq. (6.5) provides as well an efficient way to compute the gradient of the cost functional by introducing the adjoint field. To find this saddle point, we need to consider the following set of Karush-Kuhn-Tucker (KKT) [110] conditions :

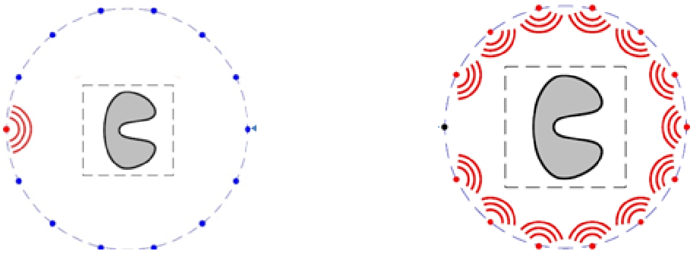
$$\begin{cases} \nabla_{\mathcal{E}^{\text{far}}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\ \nabla_{\mathcal{E}^{\text{sc}}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\ \nabla_{\varepsilon_r} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\ \nabla_{\mathcal{U}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\ \nabla_{\mathcal{P}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \end{cases} \tag{6.6}$$

This supposes that the differentiability of the cost function $\mathcal{L}$ is defined on each direction for each parameter [137, 152]. We will use this fact in order to calculate the derivatives in a given direction (Appendix A).

According to the mathematical determination of Lagrange multipliers, this derivative can be associated with an *Adjoint* total field $\mathcal{P}^{\text{tot}}$ (Appendix A). This field is obtained through the calculation of a scattered $\mathcal{P}^{\text{sc}}$ and incident $\mathcal{P}^{\text{inc}}$ Adjoint fields, as it has been done for the direct problem. Following the reciprocity principle, this adjoint field follows the same type of equations as the direct field except that the receivers act as sources and emit the discrepancy existing between the measured and simulated fields [153], which is expressed as follows for a given source position

$$J_s^{\mathrm{adj}} = \sum_{r=1}^{N_{rec}} 2\, w_{s,r}\, \overline{(\mathcal{E}_{s,r}^{\mathrm{mes}} - \mathcal{E}_{s,r}^{\mathrm{far}})} \tag{6.7}$$

Schematically we can represent these two types of electric fields as shown in Figure (6.2).



(a) Total field (b) Adjoint field

FIGURE 6.2 – Schematic explanation of (a) the total and (b) Adjoint fields in terms of antennas which are illuminating an object.

Finally, it can be shown (Appendix A) that the gradient of the cost functional can be found as follows :

$$\nabla_{\varepsilon_r} \mathcal{J}(\varepsilon_r) = \sum_{s=1}^{N_s} \overline{\mathcal{E}_s^{\mathrm{tot}}} \cdot \mathcal{P}_s^{\mathrm{tot}} \tag{6.8}$$

which is similar to the expressions found in the 2D configuration [99, 145, 146, 153].

Since a solution to the optimization problem has to be sought numerically, a parameter representation of the complex permittivity function $\varepsilon_r(\vec{r})$ is needed to obtain a finite number of optimization variables. Thus, we are going to look for the solution in a space denoted as L_h^2, which is a finite-element subspace of L^2. After finite-element discretization, the number of elements in the domains Ω and $\mathcal{D}$ are equal to N_t and L respectively. The goal is to reconstruct the complex permittivity function $\varepsilon_r(l)$, or, equivalently, the contrast $\chi(l) = \varepsilon_r(l) - \varepsilon_b(l)$ in every element $l = 1, ..., L$ of the investigation domain $\mathcal{D}$.

6.3 Inversion algorithm

The algorithm that we are going to use in this work consists in updating the map of permittivity until Eq. (6.2) is minimized as follows :

1. Let us require that

$$\mathcal{J}(\varepsilon_r^{(i+1)}) < \mathcal{J}(\varepsilon_r^{(i)}) \tag{6.9}$$

   where i corresponds to the *iteration* number. The equation above is equivalent to the demand $\mathcal{J}(\varepsilon_r^{(i)} + \delta_{\varepsilon_r}) < \mathcal{J}(\varepsilon_r^{(i)})$, where δ_{ε_r} is an update direction.
2. Following Taylor formula applied up to second order, we can determine this direction as

$$\delta_{\varepsilon_r} = -2\, \mathbf{H}^{-1}\, \nabla \mathcal{J}(\varepsilon_r) \tag{6.10}$$

   where $\mathbf{H}$ is a Hessian matrix, which is written as follows :

$$\mathbf{H} = \nabla^2 \mathcal{J}(\varepsilon_r) = \begin{bmatrix} \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{(\partial \varepsilon_{r_1})^2} & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_1} \partial \varepsilon_{r_2}} & \cdots & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_1} \partial \varepsilon_{r_L}} \\ \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_2} \partial \varepsilon_{r_1}} & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{(\partial \varepsilon_{r_2})^2} & \cdots & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_2} \partial \varepsilon_{r_L}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_L} \partial \varepsilon_{r_1}} & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{\partial \varepsilon_{r_L} \partial \varepsilon_{r_2}} & \cdots & \dfrac{\partial^2 \mathcal{J}(\varepsilon_r)}{(\partial \varepsilon_{r_L})^2} \end{bmatrix} \tag{6.11}$$

where the derivation takes into account that $\mathcal{J}$ is a real function of complex parameters [154]. In a Newton optimization scheme, the complex permittivity vector ε_r is iteratively updated as

$$\varepsilon_r^{(i+1)} = \varepsilon_r^{(i)} + \delta_{\varepsilon_r}$$

In order to accelerate the convergence rate, it is better to use the correction δ_{ε_r} as a search direction along which next iterate is located as

$$\varepsilon_r^{(i+1)} = \varepsilon_r^{(i)} + \alpha \delta_{\varepsilon_r} \tag{6.12}$$

where the positive parameter α is defined with an approximate line search. It is worth mentioning that the search direction step defined in Eq. (6.12) fully satisfies to the demand (Eq. 6.9). The choice of this parameter is such that $\mathcal{J}(\varepsilon_r^{(i)} + \alpha \delta_{\varepsilon_r})$ is close to a local minimum of the cost function $\mathcal{J}$ along the search direction δ_{ε_r}.

3. There is one more important question left : how can we calculate the term $\mathbf{H}^{-1} = \left[\nabla^2 \mathcal{J}(\varepsilon_r)\right]^{-1}$. In this work we apply the quasi-Newton method which consists in approximating the Hessian matrix $\mathbf{H}$ in the Newton correction step (Eq. 6.10) with a matrix $\mathbf{B} \simeq \mathbf{H}^{-1}$ which does not involve the explicit computation of second order derivatives. At each iteration the approximated Hessian matrix is updated based on the change in the gradient with respect to the previous iteration.

   There are a lot of different approximations for selecting the matrix $\mathbf{B}$. We have decided to implement an approach based on the classical **BFGS** method given by Broyden, Fletcher, Goldfarb and Shanno [155, 156].

   The idea of the method consists in using an approximation of the inverse Hessian matrix $\mathbf{H}$ to steer its search through variable space [156]. But as compared to the classical **BFGS** method (where the dense $L \times L$ matrix is stored), the $\mathbf{L-BFGS}$ method maintains a history of the past m steps that represent the approximation implicitly, which is much more efficient in terms of memory requirement. This memory consumption aspect is compulsory as we are dealing with 3D large-scale problems. In this thesis, we are going to use a package already developed for this algorithm by the Argonne National Laboratory and Northwestern University [157,158].

Taking into account the mathematical formulation of the inverse problem described above, we now present the iterative algorithm in a structured way :

```
Set $i = 0$
Select an initial-guess $\varepsilon_r^{(0)}$
Determine $\mathcal{J}^{(0)} = \mathcal{J}\left(\varepsilon_r^{(0)}\right)$
Calculate $\nabla\mathcal{J}^{(0)} = \nabla\mathcal{J}\left(\varepsilon_r^{(0)}\right)$
While ($\varepsilon_r$ is not convenient & $\exists$ new $\alpha\,\delta_{\varepsilon_r}$ & $i <$ MAXiter)
    - Define the length $\alpha$ and the direction $\delta_{\varepsilon_r}$ of the step with the help of the
      L − BFGS method.
    - Update the relative permittivity as $\varepsilon_r^{(i)} = \varepsilon_r^{(i-1)} + \alpha\,\delta_{\varepsilon_r}$
    - Calculate the direct problem in order to find $\mathcal{E}^{\mathrm{sc}}\left(\varepsilon_r^{(i)}\right)$, and then $\mathcal{E}^{\mathrm{tot}}\left(\varepsilon_r^{(i)}\right)$
    - Calculate the adjoint problem in order to find $\mathcal{P}^{\mathrm{sc}}\left(\varepsilon_r^{(i)}\right)$, and then $\mathcal{P}^{\mathrm{tot}}\left(\varepsilon_r^{(i)}\right)$
    - Calculate the gradient of the cost functional via Eq. (6.8)
      Set $i = i + 1$
End do
```

This iterative algorithm stops if

- the maximum number of iteration is reached
- we can not find a new step α
- there is no evolution in the relative permittivity, or in the gradient. Mathematically, we can write :

$$\frac{|\varepsilon_r^i - \varepsilon_r^{i-1}|}{|\varepsilon_r^{i-1}|} < \xi \qquad \text{or} \qquad \frac{|\nabla\mathcal{J}^i - \nabla\mathcal{J}^{i-1}|}{|\nabla\mathcal{J}^{i-1}|} < \xi \tag{6.13}$$

- the obtained relative permittivity ε_r^i is convenient, i.e.

$$\mathcal{J}(\varepsilon_r^i) < \xi \tag{6.14}$$

The term ξ corresponds to the stopping criterion of the inverse iterative process. In our work, it is set to be equal to ξ = 5.E-2 for all the following calculations.

6.4 Efficient implementation of the FETI method

Note that the considered above algorithm requires two solutions of the Helmholtz equation at every iteration step of the optimization process. These are for the direct and for the adjoint problems. High-frequency electromagnetic scattering problems call for fine meshes, and therefore lead to large-scale systems of equations. For such problems, solving Helmholtz equations with a direct method entails memory and CPU requirements that rapidly overwhelm even the largest resources that are currently available. Thus, in order to make this process more flexible in terms of memory requirement, we have implemented the FETI-DPEM2-full method previously proposed and discussed.

6.4.1 Verification criteria

Let us now discuss which are the principal criteria that we will use in order to verify the well behaviour and efficiency of the inversion algorithm.

1. Convergence of the cost function.
   This criterion shows the correct behaviour of the inversion algorithm, i.e. the choice of the permittivity step $\alpha\delta_{\varepsilon_r}$ (Eq. 6.12), the correct calculation of the gradient of the cost functional, the correct approximations of the matrix $\mathbf{B} \approx \mathbf{H}^{-1}$.

2. Comparison with the exact solution.
   We consider the first criterion as a necessary condition, but unfortunately, this condition is not sufficient. As practice shows, the inversion algorithm can often converge to a wrong solution, i.e. it exists a risk of getting trapped in local minima in addition to the risk of ending up in the wrong global minimum, as the inverse problem stays always underdetermined. So it would be useful, for example, to calculate the relative error between the numerical solution and the exact one. Such error might not be sufficient, because of the noise, or just few pixels which reach a totally wrong permittivity and deteriorate the solution. As it requires further 3D interpolation routines, due to lack of time, this criterion has not unfortunately been computed and only visual comparison are provided between reconstructed and actual maps.
3. Comparison with the FEM-based solution.
   In Sections (4) and (5) devoted to the efficient implementation of the FETI-DPEM2-full method for 2D and 3D Direct electromagnetic calculations, we showed that the limit of the proposed method is a classical FEM solution. Thus, we introduce the third criterion used in this work, that is a comparison with the FEM-based inversion algorithm.
   The principal idea of this criterion is to compare the results obtained with the FETI-based inversion algorithm with the one which is based on the FEM classical method which has been developed in our laboratory and approbate on different 3D electromagnetic inverse scattering problems [81,159]. We would like to point out the main criteria of comparison between FEM- and FETI-based algorithms, that we use in our work.
   — The L^2-error between the maps of the relative permittivity obtained with the two algorithms. The solutions are taken from the ***last*** inversion iteration of each method.
   — The number of iterations of the two algorithms.
   — The L^2-error between the FETI- and FEM-solutions at every iteration step of the inversion algorithm.

6.4.2 Permanent and non permanent information

In order to implement in an efficient way the FETI-DPEM2-full method for the inverse problem, we need to point out the main specificities of the inversion algorithm considered in Section (6.3).

To start with, the inversion algorithm that we are going to use in this work is fully iterative. It means that if we want to implement the FETI method efficiently, we have to conditionally divide it into 2 parts. The first one is going to contain all the information that does not change during all the calculations. The second one, in its turn, changes at every iteration step of the inversion algorithm.

For the part that does not change, we can list :

a) All the geometrical features (number of unknowns, FETI geometrical matrices, interface matrices $M^{i\leftrightarrow j}$, lists of edges and elements, etc)
b) All the subdomains $i = 1, ..., N_{per}$ which do not contain the investigation test domain $\mathcal{D}$. In the future, we will denote such domains as *permanent* ones. For example, the subdomains Ω_1, Ω_2 and Ω_3 in Figure (6.3) are *permanent*.
c) The *permanent* part of the non-geometrical matrix $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$. By *permanent* we call a matrix only obtained from the contribution of all the *permanent* subdomains.

For the second part that changes at every iteration step, we inventoriate :

a) All the R.H.S. for both the direct and adjoint problems.
b) All the subdomains which contain the investigation test domain $\mathcal{D}$. In the future, we will denote such domains as $non-permanent$ ones. For example, the subdomains Ω_4 and Ω_5 in Figure (6.3) are non-*permanent*.

c) The $non-permanent$ part of the $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ matrix.

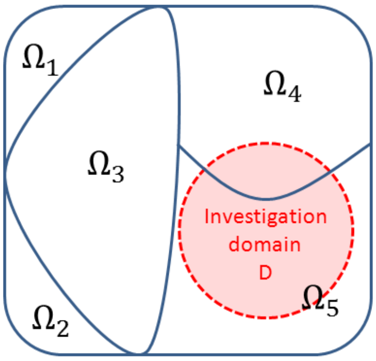
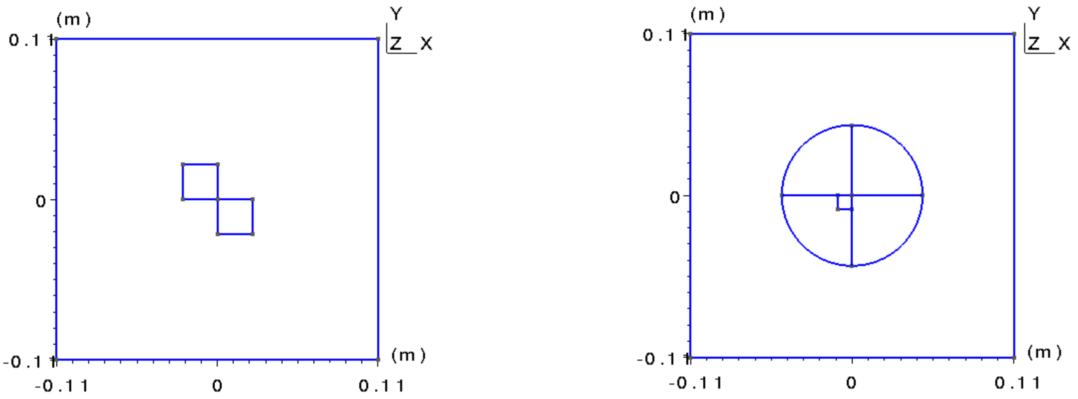


FIGURE 6.3 – Schematic example of the domain partitioning which are contained or not in an investigation test domain $\mathcal{D}$ showing the difference between *permanent* and $non-permanent$ subdomains.

The goal now is to analyse the efficiency of the *permanent/non-permanent* approach from a practical point of view. To start with, we chose the **TwoCubes** object from the Fresnel database. The choice of this target is dictated by its simplicity and size, and by the fact that we have already studied this object in Section (5.5). All the tests in this section will be performed on synthetic noiseless data.

The TwoCubes object consists of two cubes with permittivity $\varepsilon_r = 2.35$ and a side of 0.025 m. For the given problem, we choose the operating frequency equal to 4 GHz that corresponds to $\lambda = 0.075$ m. Obviously, the geometry for the direct problem (Figure 6.4a) is going to differ from the one for the inverse problem (Figure 6.4b). As previously, the domain Ω is a cube with a side of 0.23 m (or, in background wavelengths $\approx 3.1\lambda$). Then we introduce a sphere with a radius of 0.05 m centred in Ω as an investigation test domain $\mathcal{D}$. This test domain does not provide any precision on the target. We than artificially create a little cube with a side of 0.01 m inside the investigation domain (Figure 6.4b), where we set the initial permittivity equal to 1.1, with the aim of avoiding initial zero-solution. Schematically the investigated domain for the inverse problem is presented in Figure (6.4b).



(a) Direct problem

(b) Inverse problem

FIGURE 6.4 – Schematic map of the TwoCubes target scene.

The domain Ω has been discretized into finite elements (Table 6.1) and then, following the main idea of the FETI method, has been divided into 19 subdomains (Table 6.2), that corresponds to the optimal number of partitions in terms of computational time.

κ	N_p	N_t	N_e	L
10	27 820	154 485	189 285	15 197

TABLE 6.1 – Finite element discretization of the domain Ω for the inverse problem at the frequency of 4 GHz which corresponds to 10 points per wavelength.

N_s	N_{per}	N_{λ_r}	$N_{\mathbf{E}_c}$	$N_{\boldsymbol{\lambda}_c}$
19	10	39 388	1 055	2 966

TABLE 6.2 – The principal decomposition parameters of the domain Ω for the inverse problem at the frequency of 4 GHz which corresponds to the partitioning into 19 subdomains, where 10 of them are *permanent*.

As previously detailed in Sections (5.2), we treat the computational mesh obtained after finite-element discretization in terms of N_p points, N_t tetrahedras and N_e edges which corresponds to κ (Eq. 4.1) points per wavelength. We then give as well the size L of the investigation test domain $\mathcal{D}$. Similarly, after having done the domain decomposition of Ω, the given partitioning corresponds to the number of Lagrange multipliers λ_r associated to the Interface-edges (N_{λ_r}), number of corner edges ($N_{\mathbf{E}_c}$) and number of global Lagrange multipliers $\boldsymbol{\lambda}_c$ associated to the corner edges ($N_{\boldsymbol{\lambda}_c}$).

For the mesh obtained after finite-element discretization of the TwoCubes target, we have 19 subdomains, where 10 out of them are *permanent* (Table 6.2). Since all the subdomains are approximately of the same size, the time spent on the LU-decomposition for each of them is also almost the same. For the given problem, the average time of one local LU-decomposition is $t = 0.42$ sec. It means that on each iteration step of the inversion algorithm we will be able to save $N_{per} \cdot t$ sec of time. For the given problem, it is about 4.2 sec. The gain for the $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ matrix is much bigger. At the beginning of the algorithm, we create the basis of this matrix in the *csr* format. Then we allocate a memory for storing a full matrix ($N^2_{\mathbf{E}_c}$ elements of complex type) where we are going to recalculate the contribution of only $non-permanent$ subdomains. For the given problem, we save around 25 seconds per iteration step. Finally, we would like to note that the gain in terms of time is going to increase in a non-linear way when the size of the global problem is increasing. It is thus of interest to separate the various domains with respect to this $permanent/non-permanent$ feature.

6.4.3 Inner and outer iterative loops

Let us now schematically present the FETI-based inversion algorithm developed in this work (Figure 6.5). This algorithm starts with the initialisation of the inversion solution which is denoted here as "Initial guess". As an example, we can refer to the previous test case (Figure 6.4b), where inside the investigated test domain $\mathcal{D}$ of spherical form, a little cube was chosen with a relative permittivity, different from the background one. We also can notice an approach which is called *frequency-hopping* [160], where the initial solution is taken as a result of a similar inverse problem calculated at a close frequency.

The "Initial guess" component of the algorithm is followed by "FETI $permanent/non{-}permanent$" steps. These two parts of the algorithm were described previously in Section (6.4.2). We only would like to pay attention to the fact that the iterative process of the inversion algorithm comes back to the $non-permanent$ part, while the "FETI *permanent*" part is calculated only once.

The step which is called "FETI iterations" is a bit more complex. The 3D electromagnetic inverse scattering problems that we consider here are highly *underdetermined.* This term refers to the fact that the number of unknowns (in our case, pixels in the investigation test domain $\mathcal{D}$) is much smaller than the number of provided data or even of the available degrees of freedom associated to the dataset [161, 162]. Thus, in order to obtain a convergence of the algorithm to a correct solution, we have to provide a large number of input measurements. In our algorithm, the number of right-hand

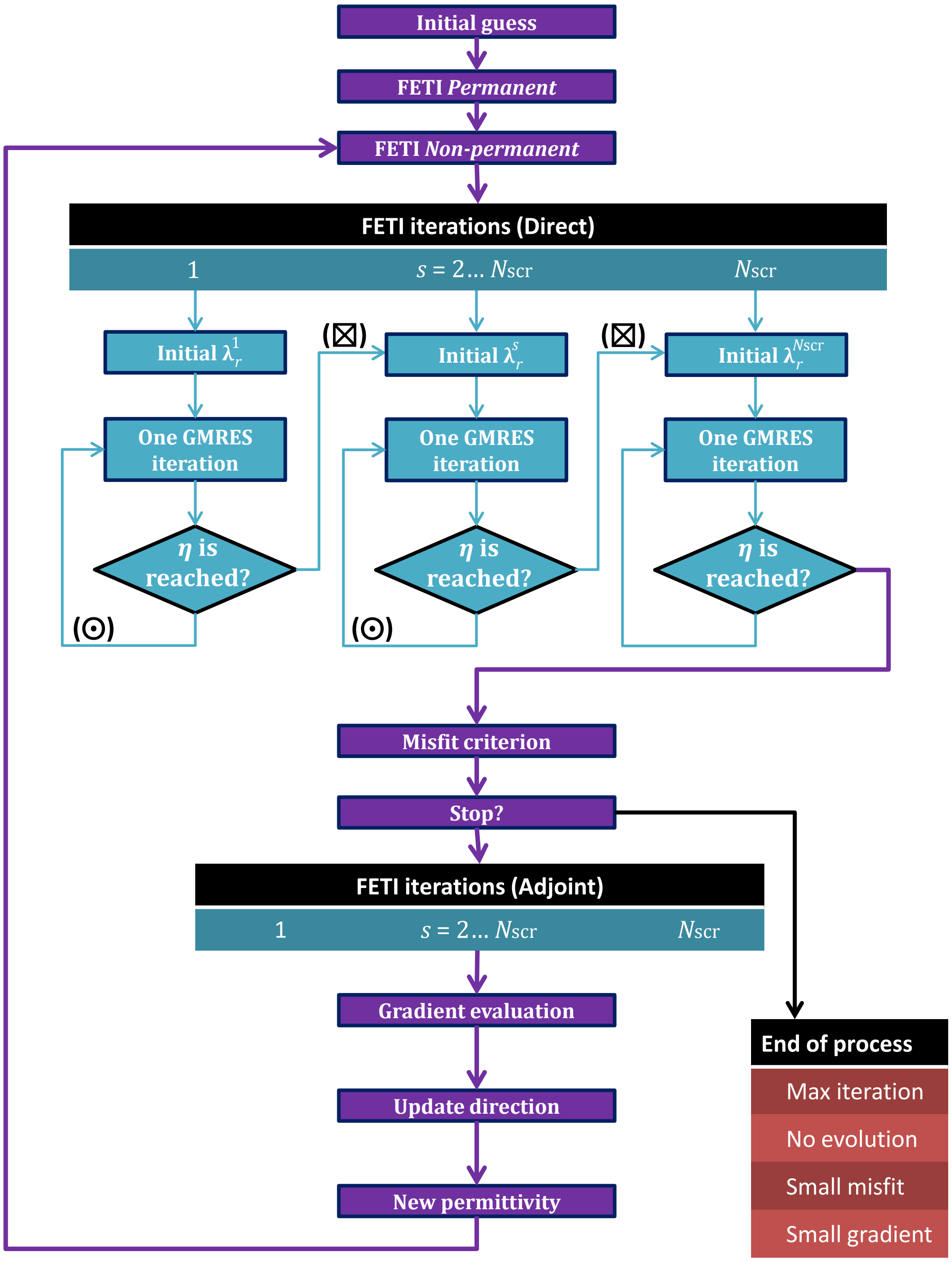


Figure 6.5 – Schematic representation of the FETI-based inversion algorithm used in this work.

sides is equal to $min(N_{src}, N_{rec})$ [3] for both the direct and adjoint problems. It means that at each inverse-iteration step we need to solve a system of linear equations with $2\,min(N_{src}, N_{rec})$ RHS. This can lead to a large number of forward problems in practical situations.

---

[3] Let us assume that $N_{src} < N_{rec}$. In this case the number of RHS for the direct problem is equal to N_{src}. Then, following Eq. (6.8), we can note that for calculating the gradient of the cost functional we only need to know the adjoint field in *source*-positions $s = 1, ..., N_{src}$. That is why the number of RHS for the adjoint problem is also equal to N_{src}. If the assumption above does not take place, i.e. in the case if $N_{src} > N_{rec}$, we can use the principle of reciprocity (exchange sources and receivers) and we will obtain the same number of RHS.

For the classical FEM method, the initialisation procedure (LU-decomposition) takes most of the time. Once it is done, obtaining ($LUx = f$) is performed in a very fast way. For the FETI method, the time management is totally different. As it was previously noted in Section (5.4.3), the convergence process of the FETI-DPEM2-full method takes as much time as the initialisation procedure (LU-decomposition and construction/decomposition of the $\hat{\mathrm{F}}_{\mathbf{E}_c\mathbf{E}_c}$ matrix). As an iterative-type method, the FETI-DPEM suffers a lot from the big number of rhs. Thus, in order to accelerate the proposed method, we will consider few ideas for selecting the solution initialization and the principal FETI parameters.

In order to apply it efficiently, we will discuss :

- The stopping criterion of the FETI solution, which is denoted as η (Section 6.4.3.1).
  Recently, we have said that the FETI method converges towards the solution of the FEM method. But to obtain such result, we need to iterate the Interface problem for a very long time. Due to the lack of time, we obviously need to use a GMRES stopping criterion η of an order which is much less than the computer error. Taking this into account, a logical question would be : what kind of information will we lose by computing the inverse problem with an error of such order and will we obtain a correct solution at all ? In other words, we will try to answer the following question : *how long do we need to repeat the step* ($\odot$) *in the scheme (6.5) in order to obtain an appropriate physical solution ?*
- The initialisation of the GMRES solution (Section 6.4.3.2).
  In order to accelerate the inversion algorithm, we can use as well the initialisation of the GMRES solution. Previously in this section, we discussed the initialisation of the inversion solution which is done only **once**. Unlikely the inversion algorithm, the method FETI can be pre-initialised at **every inversion iteration**. Thus, in this framework, we will answer to the following question : *What kind of initial GMRES solution* λ_r^s *do we need to take for both the direct and adjoint problems, i.e. what information do we have to transmit through the stage* ($\boxtimes$) *in the scheme (6.5), in order to accelerate the inversion algorithm ?*

The study of these two main aspects raised above will allow us to combine them in order to obtain the best mixture of FETI-parameters (Section 6.4.3.3).

6.4.3.1 GMRES stopping criterion

Like the quasi-Newton algorithm used in this work, the FETI-DPEM2-full method is also an iterative one. It means that we have to properly choose the stopping criterion η when solving the Interface Problem (GMRES, BiCGStab, e.t.c.). In order to obtain with the FETI-based algorithm the same[4] solution with respect to the FEM-one, we have to set η to be equal to the computer error. Unfortunately, as practice shows, the multi-sources calculation is going to be tremendously slow in terms of computation time at every iteration step.

This section thus is devoted to the stage ($\odot$) in the inversion algorithm (6.5) and the remaining question here is : can we choose the stopping criterion for the GMRES method as large as possible without loosing the quality of the inversion solution ? And is it possible to perform this choice automatically ? In this Section, we will try to answer to these questions.

To start with, we calculate the problem presented in Section (6.4.2) with different stopping criteria η for the iterative method GMRES. The related results are presented in Figure (6.6).

---

[4] By the *same*, we denote a solution obtained with the FETI-based inversion algorithm whose relative L^2- error is of the computer order with respect to the FEM-based approach

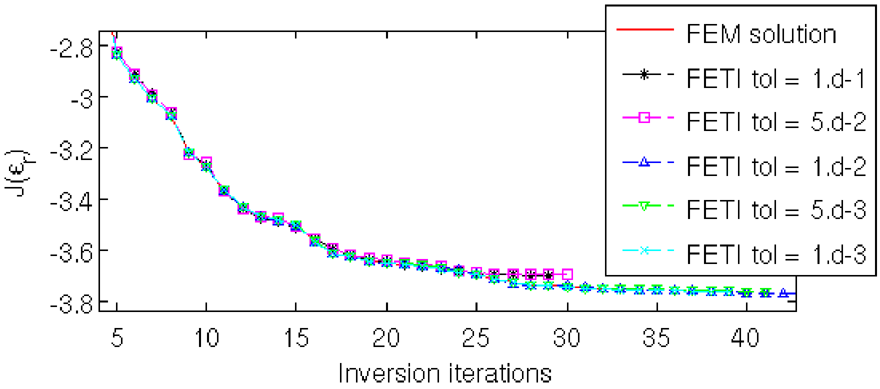
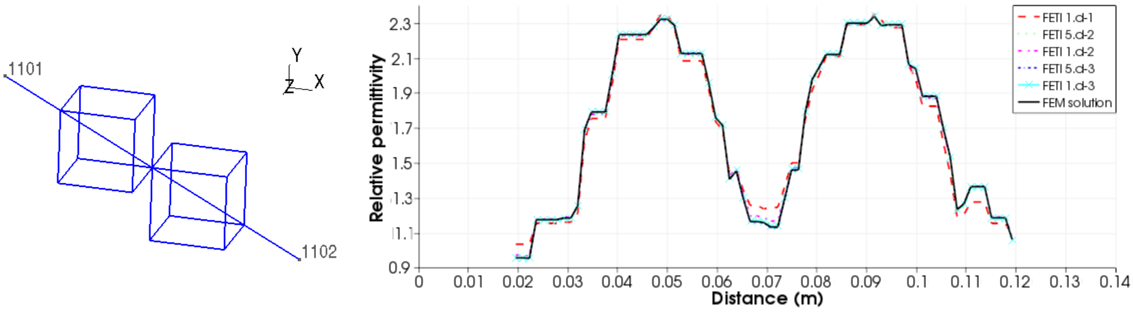


Figure 6.6 – The data fit cost function versus the number of iterations for the reconstruction of the TwoCubes target from noiseless simulated fields at 4 GHz, using the FETI-based inversion algorithm with different stopping criteria η for the GMRES method.

From Figure (6.6) it can be seen that we can conditionally divide these results into two groups. In the first one, we can include the criteria η = 1.E-1 and 5.E-2, while the second one contains all the rest. As we can see, the difference between these two groups is very small. Nevertheless, we can not guarantee that the solutions in these two groups are the same, or even close to each other.



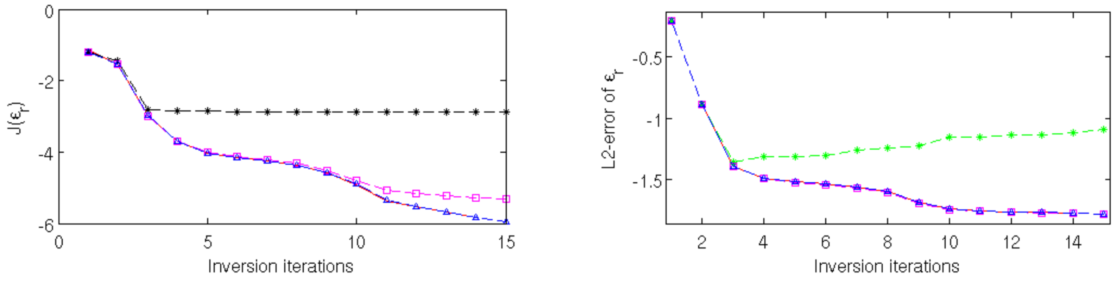
(a) Target with a crossing line (b) $\mathcal{R}e(\varepsilon_r)$

Figure 6.7 – (a) Target under test : TwoCubes object with a line (1101,1102) crossing the diagonals of the two cubes as well as the shared corner and (b) Comparison along the crossing line between the reconstructed real part of the relative permittivity obtained with the FEM- and FETI-based methods with different GMRES stopping criteria η.

That is why in Figure (6.7b) we show the solutions obtained with the different GMRES stopping criteria η in order to compare the results obtained either with the FETI- or the FEM-based methods. A line crossing the diagonals of the two cubes (Figure 6.7a) as well as its shared corner is considered more specifically.

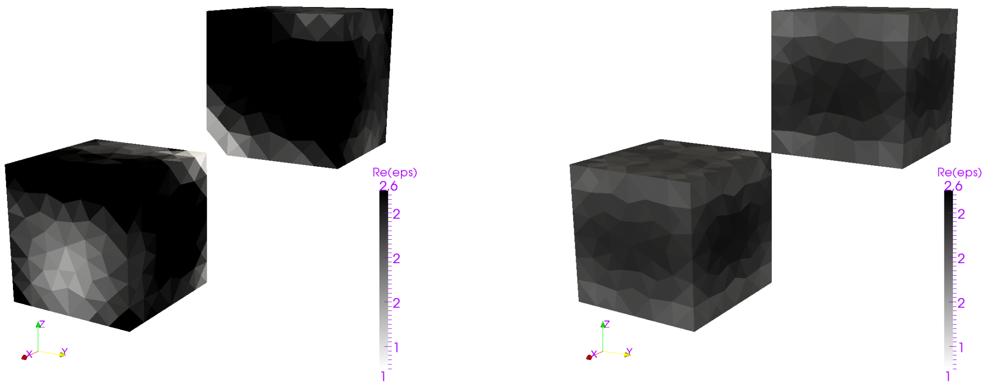
From Figure (6.7b) we can clearly see that, at the end of the inversion algorithm, we obtain a correct solution regardless of the stopping criteria, since the difference between them is not significant. But according to practice, this case is a particular one. To prove it, we will change the initial geometry of the inverse problem. We now conform the mesh of the investigation test domain $\mathcal{D}$ to the one shown in Figure (6.4a). It means that we do take into account a mesh which is conformal to the cubes shape. Then, as previously, we calculate this problem with the FEM- and FETI-based inversion algorithms with different GMRES stopping criteria for the FETI-DPEM2-full method (Figure 6.8).

The maps of the relative permittivity obtained with the FETI-based inversion algorithm corresponding to the GMRES stopping criteria η = 1.E-1 and η = 1.E-2 are presented in Figure (6.9).



(a) Convergence of $\mathcal{J}(\varepsilon_r)$ (b) Relative error of the solution

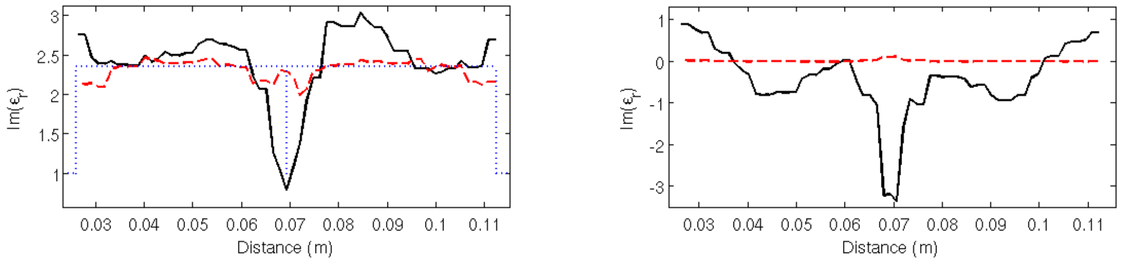
Figure 6.8 – (a) The data fit cost function and (b) the relative L2-error of the permittivity ε_r versus the number of iterations for the reconstruction of the TwoCubes target from noiseless simulated fields at 4 GHz using (- -) the FEM- and FETI-based inversion algorithms with different GMRES criteria : (- $\ast$ -) $\eta = 1.E\text{-}1$, (- □ -) $\eta = 1.E\text{-}2$, (- △ -) $\eta = 1.E\text{-}3$. The domain is described in Figure (6.4a).



(a) η=1.E-1 (b) η=1.E-2

Figure 6.9 – The reconstruction of the TwoCubes target from noiseless simulated fields at 4 GHz with the help of the FETI-based method with GMRES stopping criteria : (a) η=1.E-1 and (b) η=1.E-2

The differences between the results corresponding to $\eta = 1.E\text{-}1$ and $\eta = 1.E\text{-}2$ are presented in Figure (6.10).



(a) $\mathcal{R}e(\varepsilon_r)$ (b) $\mathcal{I}m(\varepsilon_r)$

Figure 6.10 – Comparison between the results of the FETI-based algorithm obtained with different GMRES stopping criteria, such as (—) $\eta = 1.E\text{-}1$ and (- -) $\eta = 1.E\text{-}2$ in terms of the (a) real and (b) imaginary parts of the relative permittivity and the actual one (...) along a line crossing the diagonals of the two cubes as shown in Figure (6.7a).

The results observed from Figures (6.8),(6.9) and (6.10) allow us to conclude that the value of the stopping criterion $\eta = 1.E\text{-}2$ is sufficient to obtain a correct, adequate solution of the inverse problem with respect to a FEM one.

So the question now is : how can we change this parameter dynamically ? We will try to answer this from a practical point of view. First of all, we suppose that this parameter should depend on the cost function $\mathcal{J}(\varepsilon_r)$. In most of the test cases performed in this work, the first order of the stopping criterion η is enough to obtain a correct convergence until the level of 1.E-2 – 1.E-3 in terms of cost function order. That is why in the following reconstructions we are going to use the following strategy for the choice of the GMRES stopping criterion η (Table 6.3).

	$\mathcal{J}(\varepsilon_r)$	
	$>$ 1.E-2	$<$ 1.E-2
η	$\mathcal{J}(\varepsilon_r)$	1.E-2

Table 6.3 – Selection scheme for the GMRES stopping criterion η.

In practice, once we make the stopping criterion η to be equal to 1.E-2, we can use it until the end of the calculations.

6.4.3.2 GMRES initialisation

We want now to accelerate the FETI-resolution process at every iteration step of the inversion algorithm. For this we will use a "smart initialisation" of the GMRES solution by transmitting different types of information through the stage (☒) of the inversion algorithm (6.5).

To start with, we remind that at every iteration step of the inversion algorithm we solve Eq. (3.116) in order to obtain two types of GMRES solutions : λ_r for the direct field and λ_r for the adjoint field. To proceed further, we introduce the GMRES solution obtained at the j^{th} iteration of the GMRES method for the s^{th} source position and at the i^{th} iteration of the inversion algorithm, which will be denoted as

$$\{\lambda_r^j\}_i^s \tag{6.15}$$

Up to now, $\{\lambda_r^0\}_i^s$ was set to be equal to 0 for both the direct and adjoint problems, which corresponds to a classical initialisation denoted as "FETI without".

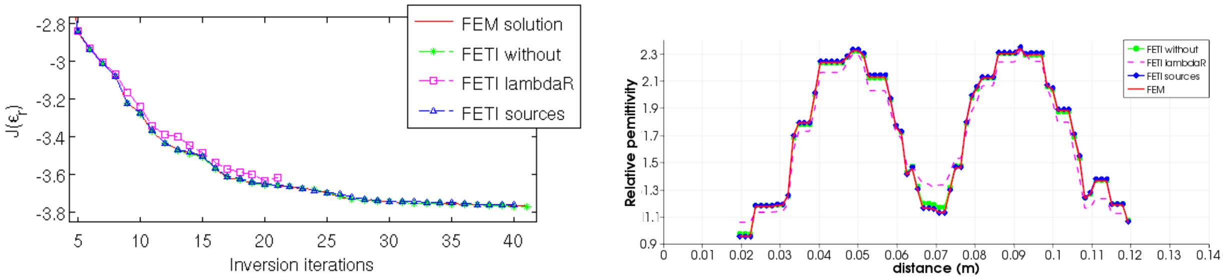
We will now consider two acceleration ideas :

- The first initialisation idea is to take $\{\lambda_r^0\}_i$ as the value $\{\lambda_r^0\}_{i-1}$ obtained at the previous iteration step of the inversion algorithm. This initialisation technique is called "FETI lambdaR".
- The second initialisation idea is based on taking the solution $\{\lambda_r^0\}^{s-1}$ obtained for the previous source as an initial guess for the present source $\{\lambda_r^0\}^s$. In literature this approach is referred as a *marching-on-in* technique [163]. Applied to the FETI method, this initialisation technique will be denoted as "FETI sources".

In order to verify the results of the FETI-based inversion algorithm obtained with the two initialisation techniques described above, we are going to verify these solutions with the classical FEM-based method and with the FETI-based method with classical initialisation.

We will make some numerical experiments based on the reconstruction of the TwoCubes target from noiseless simulated fields at 4 GHz with the GMRES stopping criterion $\eta = 1.E\text{-}2$. The results of the different initialisation techniques are presented in Figure (6.11)

Some conclusions can be drawn from Figure (6.11). First of all, the FETI-based inversion algorithm often stops earlier then the FEM-one. It can be caused by the fact that with a low order GMRES stopping criterion, we are not able to provide high-order information on the solution and do not

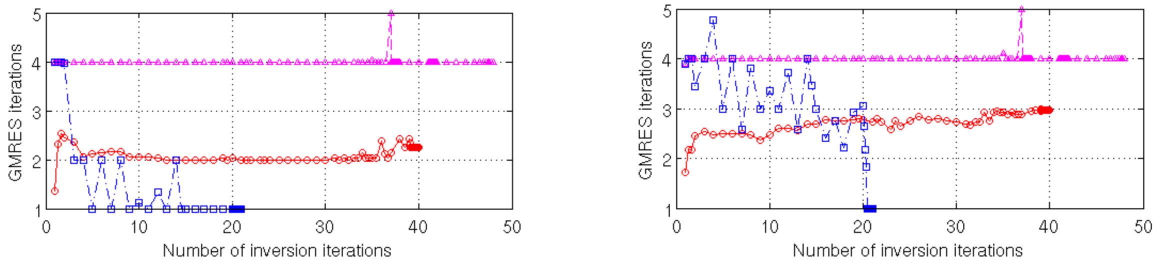


(a) Convergence of the cost function (b) Cut of $\mathcal{R}e(\varepsilon_r)$

FIGURE 6.11 – (a) Data fit cost function and (b) Comparison along the crossing line presented in Figure (6.7a) between the reconstructed real part of the relative permittivity obtained with the FEM- and FETI-based methods with different GMRES initialisation procedures.

find a new direction for ε_r after some point. In the particular case of Figure (6.11), the FETI-based algorithm with the FETI lambdaR initialisation stopped after 21 iterations, because all the solutions obtained after the 20st iteration were equal to the solutions of the 21st iteration up to the error of the η tolerance, so the GMRES method did not make any extra iterations. In spite of this, the solution obtained with FETI lambdaR initialisation is satisfactory, because it provides a similar information about the target. Nevertheless, the solutions obtained with the other types of initialisation are in perfect agreement with the FEM-based algorithm. It is thus difficult to derive the best initialisation so far if we only compared the reconstructed permittivity maps.

Let us now consider the number of iterations of the FETI-DPEM2-full method on each iteration step of the inversion algorithm. For this, we will distinguish two pairs within each inverse-iteration : the direct problem computations and the adjoint ones. We also would like to note that from this moment we are going to compare all the results in terms of the ***Average number*** of the GMRES iterations with respect to the number of sources, which we call the $\mathcal{A}$ number.



(a) (b)

FIGURE 6.12 – The $\mathcal{A}$ number of iterations of the (a) direct field and (b) adjoint field calculations obtained with the (- · □) FETI lambdaR initialisation, (- ○ -) FETI sources initialisation, (- △ -) FETI without initialisation versus the number of iterations of the inversion algorithm.

These results presented in Figure (6.12) were a bit surprising, but after having done other numerical experiments, we concluded that, for the direct field calculations, the FETI lambdaR initialisation is more useful then the FETI sources one, whereas for the adjoint field calculations it is the contrary.

We suppose that such conclusion is due to the fact that for the direct problems, as the sources are far apart, the incident wave changes from one source to another. On the contrary, the map of ε_r varies less. So the initialisation technique ”FETI lambdaR” works better in this case than the ”FETI sources” initialisation. Such results should be mitigated if more sources were present, as in that case, the ”FETI sources” should also work satisfactorily.

As for the adjoint field, it is created according to Eq. (6.7), where the emission corresponds to the discrepancy between the measured and simulated field at the receiver locations. As these receivers are sending at the same time this discrepancy signal, the adjoint field sources correspond to the superposition of all these fields. Thus, from one iteration to the other one, there is no changes for the adjoint sources locations and their amplitude, due to the averaging performed by the superposition, are quite similar, apart for the first iterations. This seems the reason why the "FETI sources" initialisation works better in this case.

6.4.3.3 The best mixture of the FETI parameters

In this Section we would like to summarize all the useful information discussed previously in Sections (6.4.3.1) and (6.4.3.2) in order to create the most efficient FETI-based inversion algorithm. Thus, we will describe the main parameters that we have selected for the FETI-DPEM2-full method.

1. During all the next computations we will use the strategy of the GMRES stopping criterion choice described in Section (6.4.3.1).
2. For the $\{\lambda_r\}$ choice we will take the results of Section (6.4.3.2) with few modifications. First of all, in order to reduce the number of iterations of the inversion algorithm as much as possible, we will change the initialisation technique for the direct field calculation (from lambdaR to sources) when the lambdaR initialisation provides only zero information (for details see Section 6.4.3.2). Secondly, we should not forget about the case without any initialisation. For example, according to the results in Figure (6.12b) there are some cases when the initialised solution converges with larger number of iterations as compared to the case with no initialisation. Thus if the relative error of the first GMRES iteration is bigger then 1.0 we make the initial solution to be equal to 0 (no specific initialisation strategies) and continue the calculations.
3. For the calculation of the adjoint field, we only apply the "FETI sources" initialisation so far.

Taking this into account, let us present the results obtained with the FETI-based algorithm with the most efficient "mixture" of the FETI parameters. In particular, the reconstruction of the chosen target from the noiseless calculations (Figure 6.14) and the $\mathcal{A}$ number of iterations for the both direct and adjoint field computations (Figure 6.13). They show us the efficiency of the selected FETI-parameters.

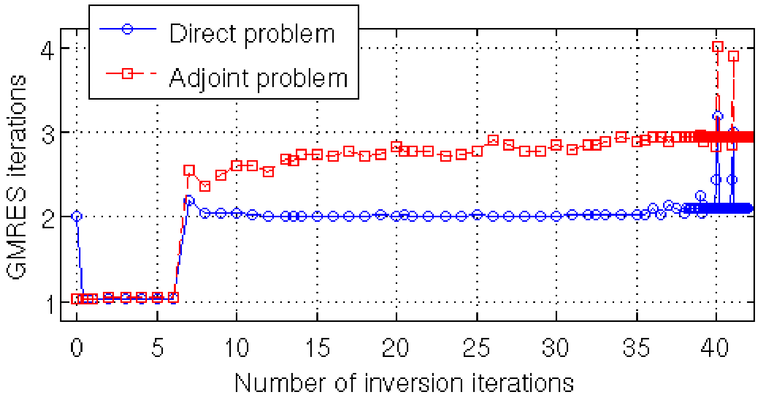
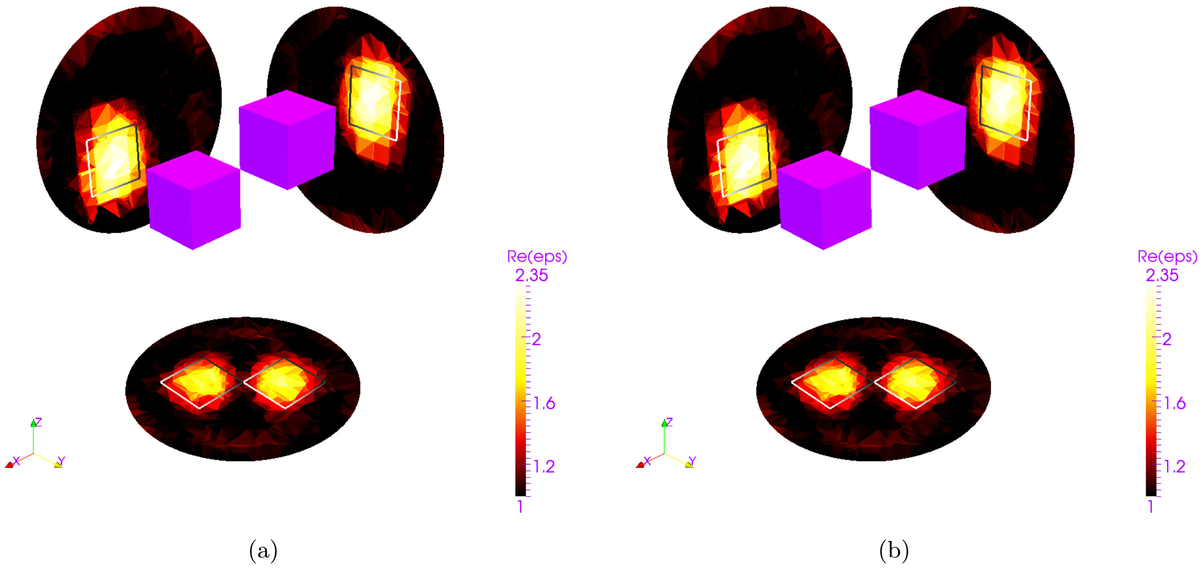


FIGURE 6.13 – $\mathcal{A}$ number of iterations of both the direct and adjoint field computations versus the number of iterations for the FETI-based inversion algorithm with the most efficient "mixture" of the FETI parameters.



(a) (b)

FIGURE 6.14 – The reconstruction of the TwoCubes target from simulated noiseless fields at 4 GHz with the help of the (a) FEM- and (b) FETI-based inversion algorithms.

6.4.4 Memory and time question

There are two important questions that we would like to discuss now. These are related to the technical features of the memory and time requirement. The former is the easiest one, because in order to see the advantage of the FETI method implementation, we will only provide the memory volume required for the FEM- and FETI-based methods. The latter is a bit more complex, because the time required for obtaining a solution with the FETI-based method linearly depends on the number of FETI (GMRES) iterations. From this knowledge, we can nevertheless estimate the time of one iteration of the FETI-based inversion algorithm as Eq. (6.16).

$$\text{Time} = \min(N_{src}, N_{rec})\,(\mathcal{T}_d + \mathcal{T}_a) \tag{6.16}$$

where $\mathcal{T}_d$ and $\mathcal{T}_a$ are the times for solving the Helmholtz equations for the direct and adjoint problems respectively, which highly depend on the number of FETI iterations. Each of these terms was previously considered in Section (5.4.3). In the case of n iterations (for example given in Figure 6.13), for the chosen size of the Krylov subspace m (for the *normal* GMRES parameters $m = 10$), for the partitioning into $N_s = 19$ subdomains and, finally for the times $T_{f\&b}$ and $T^c_{f\&b}$ (see Eq. 5.11) which are equal to 0.3 sec and 0.4 sec respectively, we can estimate the total time as :

$$\text{Time} = 36 \cdot (10\,[2 \cdot 19 \cdot 0.3 + 0.4]) = 4248 \text{ sec}$$

6.5 Inversion of the experimental data

In this section we will discuss the efficiency of the FETI-based inversion algorithm when confronted to measurements, where noise is unfortunately present. For this, we will perform the reconstruction of few targets from the 3D Fresnel database [80, 81] which have been used intensively for testing 3D inversion algorithms. The Fresnel database involves the set of homogeneous objects, i.e. objects where only two permittivity values are present in the permittivity profiles, the background permittivity and the permittivity of the target. Most of these objects have been studied and simulated in Section (5.5) where the proposed FETI-DPEM2-full method was applied for solving 3D Large-scale electromagnetic scattering problems. Thus, in this section we will not detail again the structure of each target. For its definitions the reader is refered to Section (5.5).

To start with, we set the working frequency to 4 GHz to obtain the reconstruction for all the objects. Then, for every target, we are going to use the inversion algorithm, described previously in Section (6.3). For the solution of the direct problem we will either use the classical FEM method or the proposed method FETI-DPEM2-full with the *optimal* set of parameters discussed in Section (6.4.3.3). For every target, these two resulting inversion schemes are employed in independent reconstructions and the results are compared.

For the two inversion algorithms (FEM- and FETI-based) the investigation test-domain $\mathcal{D}$ is a sphere with a diameter of 0.16 m ($\approx 2.1\lambda$). It is contained within a global cubic domain Ω with a side of 0.26 m ($\approx 3,5\lambda$). As previously, we artificially create a little cube with a side of 10 mm inside the investigation domain, where we set the initial permittivity equal to 1.1, with the aim of avoiding initial zero-solution. The geometry of the domain Ω with the investigation domain $\mathcal{D}$ is presented in Figure (6.15). As you can see from this figure, the center of domain/test-sphere is located at (0.,0.,0.025) m. The investigation domain $\mathcal{D}$ is such that it contains all the objects from the Fresnel database.

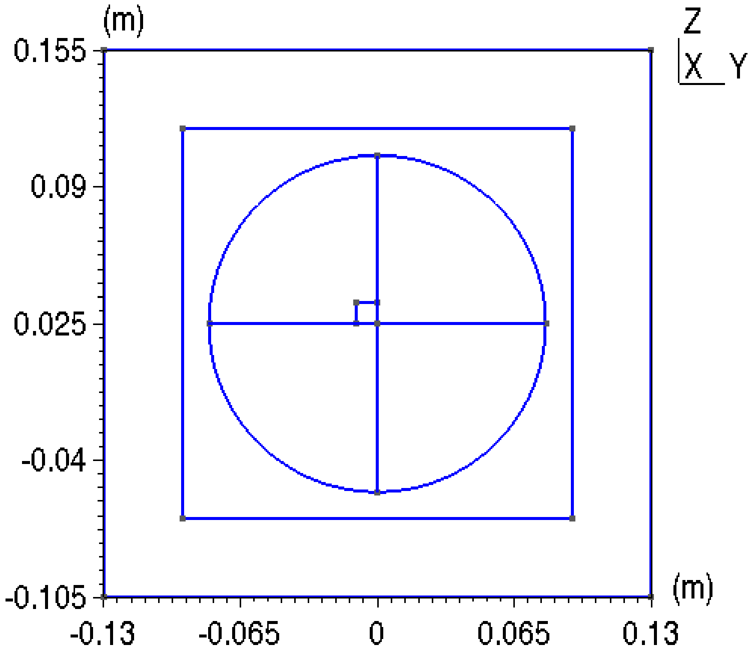


FIGURE 6.15 – Geometry of the spherical investigation domain $\mathcal{D}$ with an artificial cube, within the computational domain Ω.

The principal domain Ω has been discretized into finite elements (Table 6.4) and then, following the main idea of the FETI method, has been divided into 38 subdomains (Table 6.5), that corresponds to the optimal number of partitions in terms of computation time. In this case, the number of *permanent* subdomains is 12 out of 38.

κ	N_p	N_t	N_e	L
10	51 050	302 244	362 280	77 539

TABLE 6.4 – Finite element discretization of the domain Ω for the Fresnel database reconstruction at the frequency of 4 GHz which corresponds to 10 points per wavelength.

N_s	N_{per}	N_{λ_r}	$N_{\mathcal{E}_c}$	N_ξ
38	12	89 308	2 818	7 963

TABLE 6.5 – The principal decomposition parameters of the domain Ω which corresponds to the partitioning into 38 subdomains, where 12 of them are *permanent*.

As the chosen mesh is common for the reconstruction of all the targets from the Fresnel database, we would like to mention here the memory requirement. For the classical FEM method, we need around 15.3 GB of operative memory in order to reconstruct the objects with the mesh described above. It takes only 5.7 GB for the FETI-DPEM2-full method. As the time requirement highly depends on the number of FETI (GMRES) iterations, we will discuss it for each object under reconstruction.

6.5.1 TwoCubes target

Let us start with the object that we used for applying and testing the FETI-DPEM2-full method in Section (5.5.4). This object consists of two cubes with permittivity $\varepsilon_r = 2.35$ and with a side length of 0.025 m (Figure 6.16).

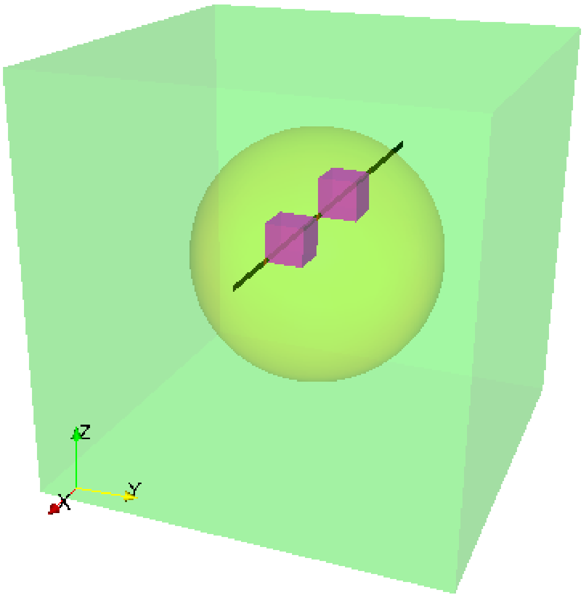
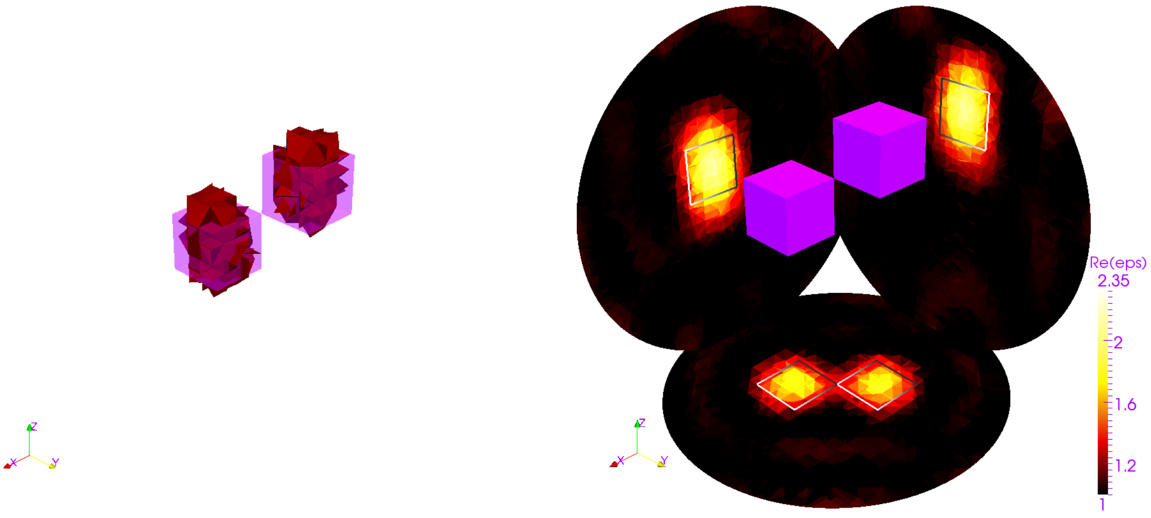


FIGURE 6.16 – Geometry of the TwoCubes target compared to the investigation domain $\mathcal{D}$ which is a sphere. A crossing black line is also presented.

The initial reconstruction of this target at 4 GHz with the FETI-based inversion algorithm yields the result of Figure (6.17). The cubes are located at the correct position and have more or less the correct size. As we can see, the permittivity values are a bit too low ($min = 0.58, max = 2.12$). However, we can conclude that for the given target the optimal set of FETI-parameters discussed in Section (6.4.3.3) is sufficient to obtain a physically correct solution.



(a) (b)

FIGURE 6.17 – Reconstruction of the TwoCubes target from measured fields at 4 GHz obtained with the FETI-based inversion algorithm. (a) 3D view of the reconstructed iso-surface corresponding to values $\varepsilon_r = 1.5 \ldots 2.35$ and (b) Vertical and horizontal cuts at the cube centres of the real part of the permittivity maps. The semi-transparent object inside represents the actual boundaries of the target.

In order to comfort the obtained results, we then reconstruct the same target with the FEM-based inversion algorithm. As it can be seen from Figure (6.18), both the FEM- and FETI-based algorithms stopped after 25 iterations with the data fit reduced to 1.30d-3 and 1.49d-3 respectively.

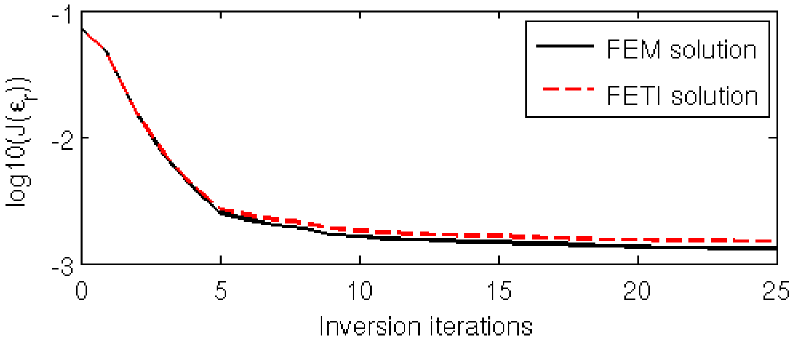
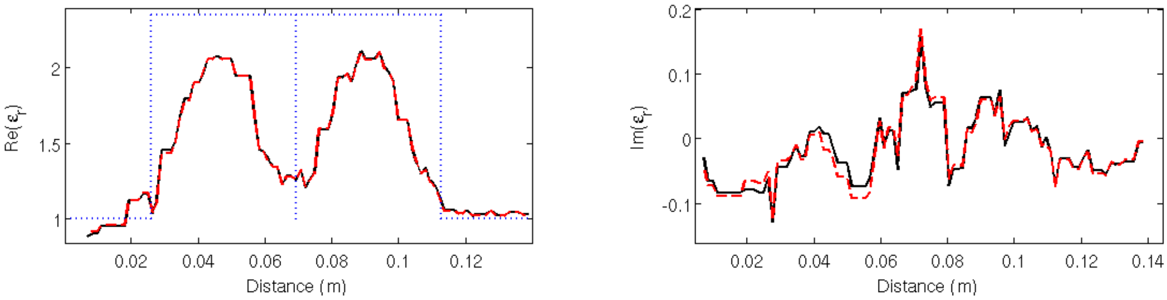


FIGURE 6.18 – Evolution of the least squares data fit cost function $\mathcal{J}(\varepsilon_r)$ over iterations for the reconstruction of the TwoCubes target from the measured fields at 4GHz using the FEM- and FETI-based inversion algorithms.

As we can see, the behaviour of the cost functions is similar for the two algorithms, but in order to be sure that we obtain similar solutions as well, we need to provide another type of comparison, as shown in Figure (6.19).

From Figure (6.19) we can conclude that the results obtained with the FEM- and FETI-based algorithms are in good agreement. It means that the proposed FETI-DPEM2-full method was implemented in a correct way. In fact, we are not interested in looking for the imaginary part of the solution ε_r, because for all the objects from the Fresnel database $\mathcal{I}m(\varepsilon_r) = 0$. Nevertheless, our inversion algorithm is based on the independent search of both the real and imaginary parts. That is why in the framework of the TwoCubes target reconstruction, we would like to show that at the end of the



(a) $\mathcal{R}e(\varepsilon_r)$ (b) $\mathcal{I}m(\varepsilon_r)$

FIGURE 6.19 – Comparison between the (—) FEM- and (- -) FETI-based solutions in terms of the (a) real and (b) imaginary parts of the relative permittivity and the actual one (. . .) along a line crossing the diagonals of the two cubes as shown in Figure (6.16).

inversion algorithm, the imaginary part of the solution is not equal to 0 due to the noise existence and to the fact that the problem is undetermined (Figure 6.19b). Nevertheless, the reconstructed imaginary parts are reasonably low. As no *a-priory* constraints on the values are given, it is also possible to obtain permittivity values with $\varepsilon_r' < 1$ or $\varepsilon_r'' < 0$.

The technical side of the FETI-based inversion algorithm is represented in Figure (6.20). As we

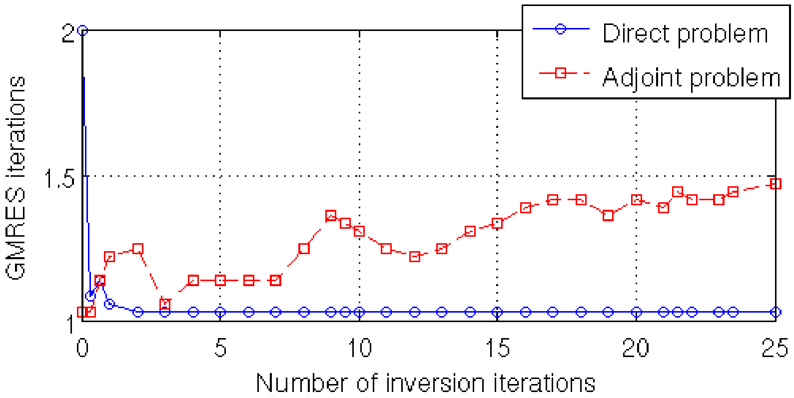


FIGURE 6.20 – Evolution of the average GMRES iterations for one source for the direct and adjoint problems.

can see, the average number of iterations for one source is equal to 1 for the direct problem and varies from 1 to 1.5 for the adjoint problem. It means that the initialisation procedure discussed in Section (6.4.3.2) is good enough for both the direct field and the adjoint field computations. We remind that the minimal number of antennas is 36 and the time of one GMRES iteration is ≈ 11 sec. In order to estimate the computational time, we use Eq. (6.16). For this target, the approximate time of one inverse-iteration is equal to 1030 sec against 450 sec for the FEM solution when no parallelization procedure is followed. This knowledge gives us an idea of the total time of the inversion algorithm.

6.5.2 TwoSpheres target

The TwoSpheres target consists of two dielectric spheres with a diameter of 0.05 m and a permittivity $\varepsilon_r = 2.6$. As compared to the previous object, the TwoSpheres target is approximately twice bigger. Its actual position in the investigation domain $\mathcal{D}$ is presented in Figure (6.21).

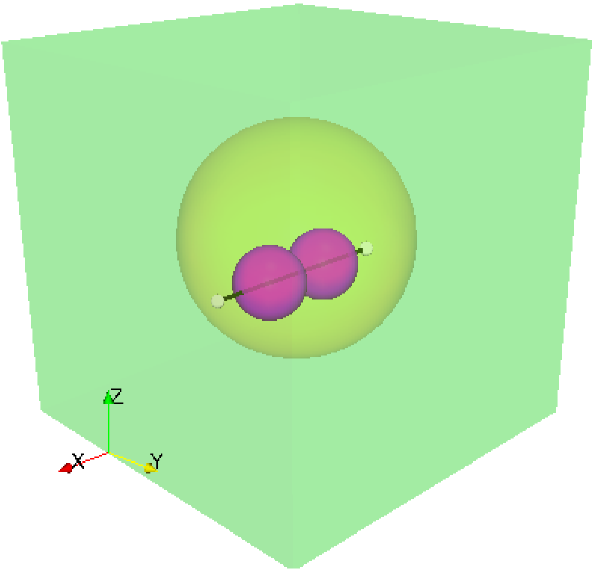
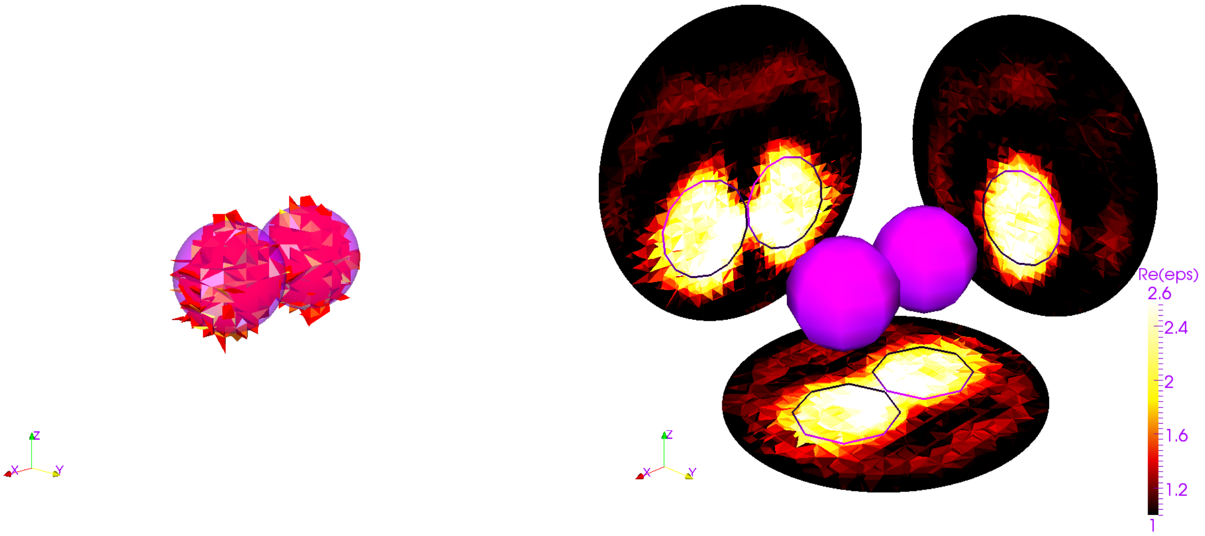


FIGURE 6.21 – Geometry of the TwoSpheres target within the spherical investigation domain $\mathcal{D}$. A crossing black line is also presented.

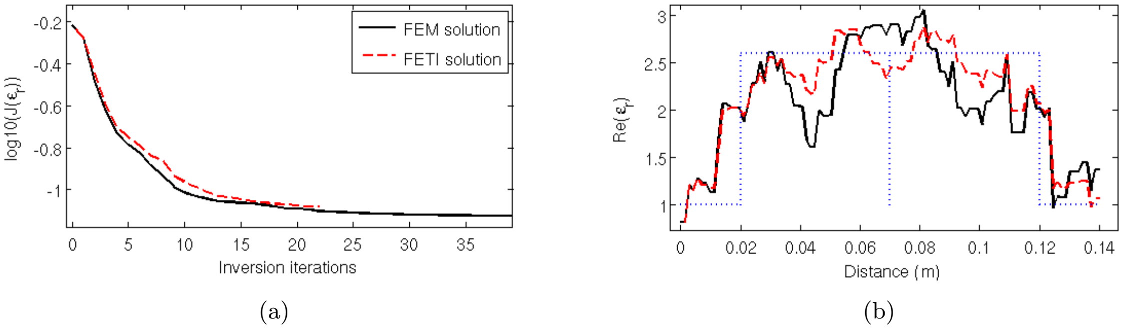
The initial reconstruction of this target at 4 GHz with the FETI-based inversion algorithm yields the result of Figure (6.22). The spheres are located at the correct position and have more or less the correct size. As we can see, the permittivity values are totally different from the previous target ($min = -4.55, max = 3.34$), however their spreading yields the determination of value $\varepsilon_r = 2.6$. That leads us to a conclusion that for the given target the optimal set of FETI-parameters discussed in Section (6.4.3.3) is totally sufficient to obtain a physically correct solution.



(a) (b)

FIGURE 6.22 – Reconstruction of the TwoSpheres target from measured fields at 4 GHz obtained with the FETI-based inversion algorithm. (a) 3D view of the reconstructed iso-surface corresponding to values $\varepsilon_r = 2.4 \ldots 2.6$ and (b) Vertical and horizontal cuts of the real part of the permittivity maps. The semi-transparent object inside represents the actual boundaries of the target.

We then reconstruct the same target with the FEM-based inversion algorithm.



(a) (b)

Figure 6.23 – Comparison between the (—) FEM- and (- -) FETI-based solutions for the reconstruction of the TwoSpheres target from measured fields at 4GHz in terms of (a) the evolution of the least squares data fit cost function $\mathcal{J}(\varepsilon_r)$ over iterations and (b) the real part of the relative permittivity with the actual one (. . .) along a line crossing the diameters of the two spheres as in Figure (6.21).

As we can see from Figure (6.23a) the behaviour of the cost functions is similar for the two algorithms, in spite of the different number of iterations (23 for FETI and 39 for FEM). Thus, similarly to the previous target, we provide as well another type of comparison, as shown in Figure (6.23b). We can clearly see the difference between the two solutions. Nevertheless we are fully satisfied by the obtained results, because this difference is of the same level as the noise. Thus, we suppose that the results obtained with the FEM- and FETI-based algorithms are in a good agreement. The imaginary part of the solution is different from zero, but its value is not significant.

The technical side of the FETI-based inversion algorithm is represented by Figure (6.24). As we can see, the average number of iterations for one source is varying from 1 to 3 for the direct problem and from 1 to 6 for the adjoint problem. It is twice bigger with respect to the previous target. This can be explained by the size of the TwoSpheres object, which is almost twice bigger than the object TwoCubes. For this target the approximate time of one inverse-iteration is equal to 2100 sec against 450 sec for the FEM solution. This knowledge gives us an idea of the total time of the inversion algorithm.

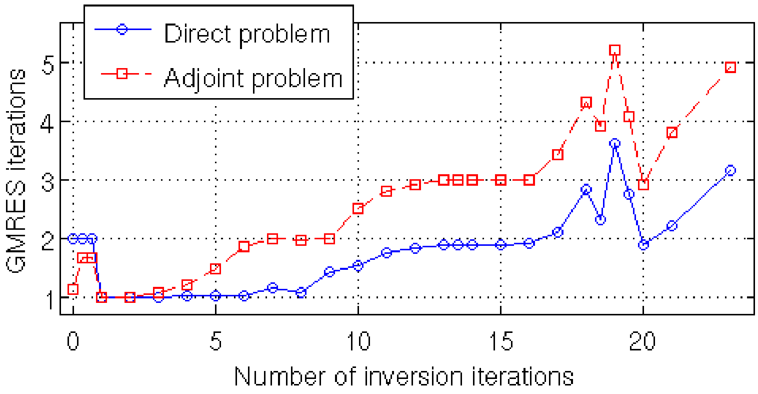


Figure 6.24 – The evolution of the average GMRES iterations for one source for the direct and adjoint problems.

6.5.3 CubeSpheres target

The CubeSpheres target consists of an aggregate of dielectric spheres. Each of them has a diameter of 0.0159 m and a permittivity of 2.6. The position of the target in the investigation domain $\mathcal{D}$ is presented in Figure (6.25).

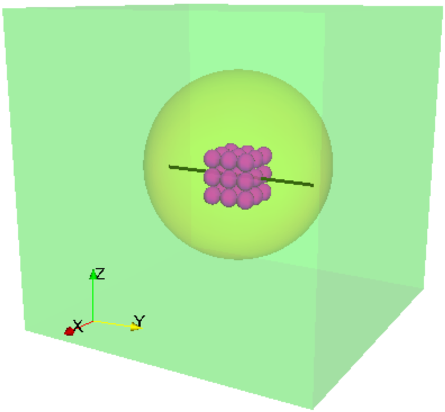
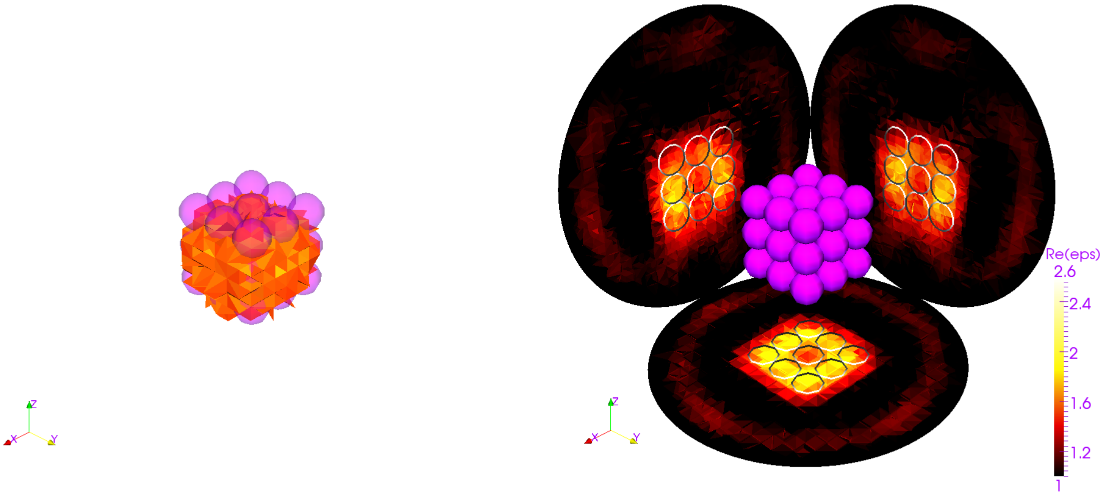


FIGURE 6.25 – Geometry of the CubeSpheres target within the spherical investigation domain $\mathcal{D}$. with the coupling black line.

The initial reconstruction of this target from measured fields at 4 GHz with the FETI-based inversion algorithm yields the result of Figure (6.26). The spheres are located at the correct position, nevertheless we are faced with a problem of size and form determination of each individual sphere. Such results can be explained as follows : the diameter of a sphere in the wavelength background is $0.0159\ m \approx 0.2\lambda$, which is much less then the wavelength. As we know, in the classical tomography, the expected spatial resolution is $\sim \lambda/2$ [59]. With non-linear scattering algorithms, the expected spatial resolution might be lower. Nevertheless, without adding *a-priori* information, it is doubtless

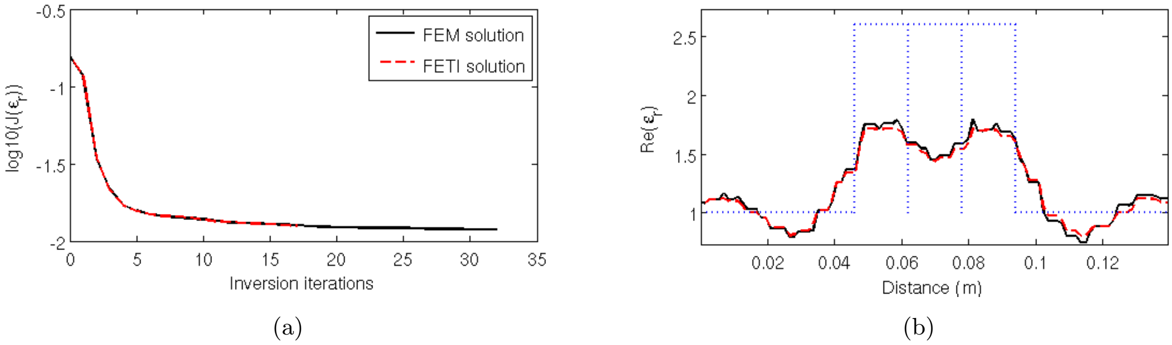


(a) (b)

FIGURE 6.26 – Reconstruction of the CubeSpheres target from measured fields at 4 GHz obtained with the FETI-based inversion algorithm. (a) 3D view of the reconstructed iso-surface corresponding to values $\varepsilon_r = 1.7 \dots 1.94$ and (b) Vertical and horizontal cuts of the real part of the permittivity maps. The semi-transparent object inside represents the actual boundaries of the target.

to obtain better results. That is why the obtained results are not surprising for us. As we can see, the permittivity values are low ($min = 0.57, max = 1.94$). Indeed, as the reconstructed target is bigger, its estimated permittivity is lower in order to provide the same scattering power.

The logical question now is : could the FEM-based method manage this problem better ? In order to answer this question, we reconstruct the same target with the FEM-based inversion algorithm.



(a) (b)

FIGURE 6.27 – Comparison between the (—) FEM- and (- -) FETI-based solutions for the reconstruction of the CubeSpheres target in terms of (a) the evolution of the least squares data fit cost function $\mathcal{J}(\varepsilon_r)$ over iterations and (b) the real part of the relative permittivity with the actual one (...) along a line crossing the diameters of the spheres as in Figure (6.25).

As we can see from Figure (6.27a) the behaviour of the cost functions is similar for the two algorithms, in spite of the different number of iterations (17 for FETI and 33 for FEM). Thus, similarly to the previous target, we provide as well another type of comparison, as shown in Figure (6.27b). In contradiction to the previous target (Section 6.5.2) where the difference between the number of iterations was almost the same, for the given object we can hardly see the difference between the FEM- and FETI- solutions. We suppose that the results obtained with the FEM- and FETI-based algorithms are in a good agreement. The imaginary part of the solution is different from zero, but its value is not significant.

The technical side of the FETI-based inversion algorithm is represented by Figure (6.28). As we can see, the average number of iterations for one source is varying from 1 to 2 for the direct problem and from 1 to 5 for the adjoint problem. For this target, the approximate time of one inverse-iteration is equal to 2650 sec against 450 sec for the FEM solution. This knowledge gives us an idea of the total time of the inversion algorithm. It is worth mentioning here that the FEM-based algorithm requires more iterations (twice in fact) than the FETI-based inversion approach and that they stop for different reasons.

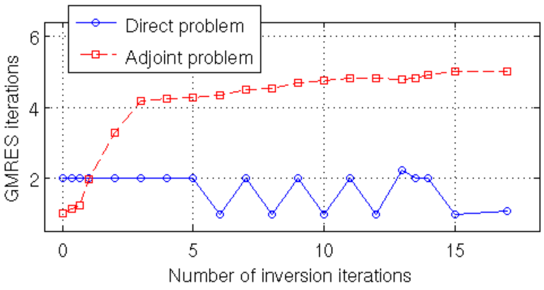


FIGURE 6.28 – Evolution of the average GMRES iterations for one source for the direct and adjoint problems.

6.5.4 Myster target

The Myster target consists of 12 dielectric spheres with a diameter of 0.0238 m and a permittivity $\varepsilon_r = 2.6$ which were glued together to compose a regular geometrical figure. Its actual position in the investigation domain $\mathcal{D}$ is presented in Figure (6.29).

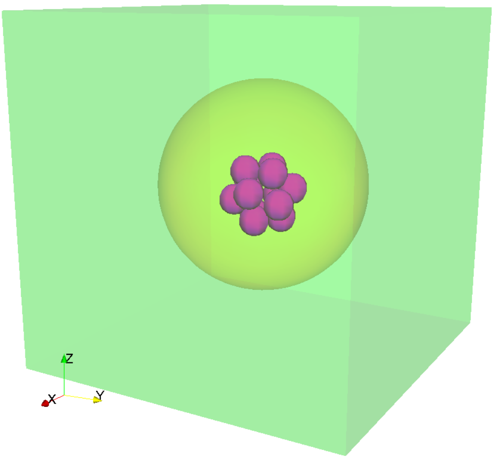
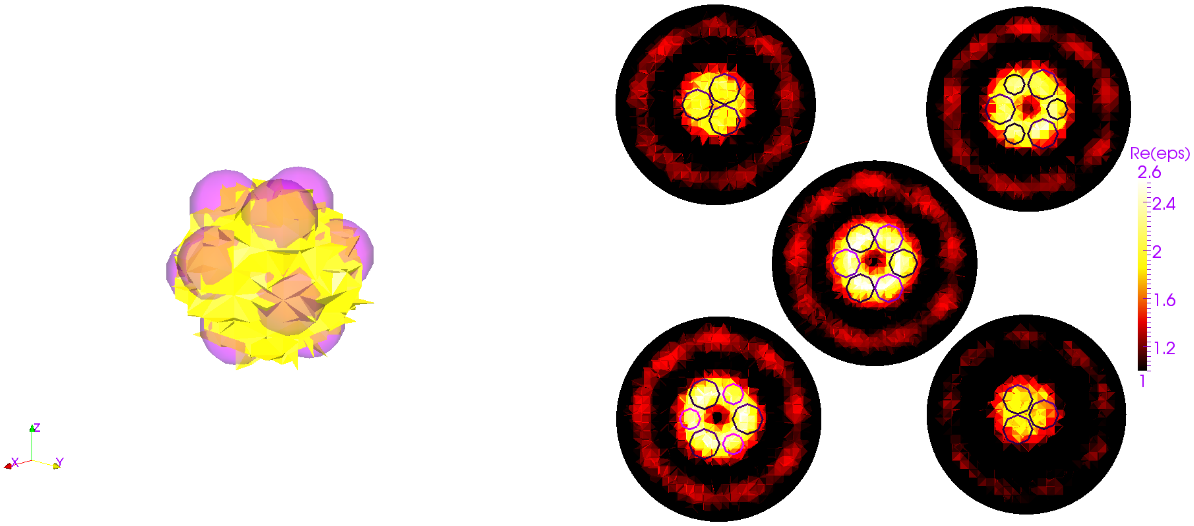


FIGURE 6.29 – Geometry of the Myster target compared to the investigation domain $\mathcal{D}$.

This target was not unveiled to the inverse problem community before the publication of the special section in Inverse Problems. This target was indeed, at that time, an effective mysterious target for blind testing the various algorithms.

The initial reconstruction of this target at 4 GHz with the FETI-based inversion algorithm yields the result of Figure (6.30). The spheres are located at the correct position, nevertheless we are faced



(a) (b)

FIGURE 6.30 – Reconstruction of the Myster target from measurements at 4 GHz obtained with the FETI-based inversion algorithm. (a) 3D view of the reconstructed iso-surface corresponding to values $\varepsilon_r = 1.72 \ldots 2.50$ and (b) Real part of the relative permittivity depicted in five horizontal slices (top : $z = 0,\ 13.74$, down : $z = 22.23,\ 35.97$, center : $z = 17.985$ mm). The semi-transparent object inside represents the actual boundaries of the target.

with the same type of problems as for the CubeSpheres target. As we can see, the permittivity values are also a little bit low ($min = 0.34, max = 2.50$).

In order to prove the obtained results from the numerical point of view, we then reconstruct the same target with the FEM-based inversion algorithm. As it can be seen from Figure (6.31), the data fit is reduced to 2.17d-2 after 63 FEM-iterations and 2.50d-2 after 36 FETI-iterations. The behaviour of the cost functions is similar for the two algorithms, even if the error remains higher with the FETI solution. The Myster target is geometrically more complicated than all the previously considered targets. Thus, we need to provide another type of comparison, shown in Figure (6.32) in order to be sure that we obtain similar solutions as well.

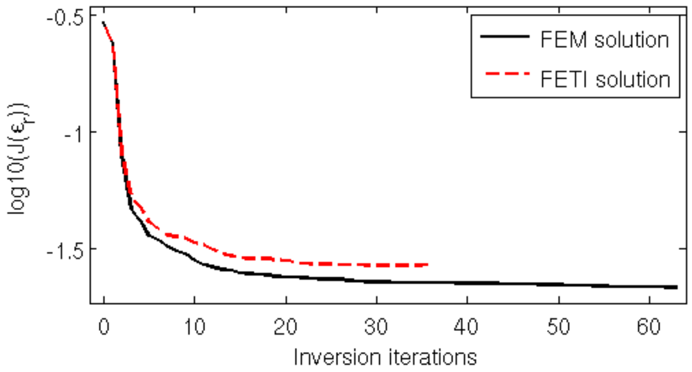


FIGURE 6.31 – Evolution of the least squares data fit cost function $\mathcal{J}(\varepsilon_r)$ over iterations for the reconstruction of the Myster target from the measured fields at 4GHz using FEM- and FETI-based inversion algorithms.

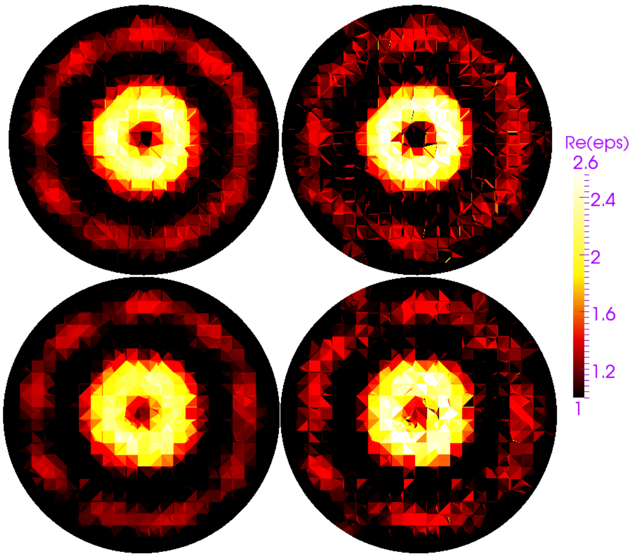


FIGURE 6.32 – Reconstruction of the Myster target from measurements at 4 GHz obtained with the (left) FETI- and (right) FEM-based inversion algorithm. The real part of the relative permittivity is depicted in horizontal slices (top : $z = 0.01374$ and down : $z = 0.02223$).

The results obtained with the FEM method appear sharper, but also we can note that there is a brighter unwanted ring around the target. To avoid it, we have to stop the inversion algorithm at an earlier iteration. In this case, if we compare the results obtained with the FEM- and FETI-based algorithms after the *same* number of iterations, we will hardly note the difference both numerically and visually. Indeed, at the end of the inversion process, when the cost functional is sufficiently reduced, it reaches the level of the signal to noise of the experiment. Thus, the last iterations only provide oscillations in the reconstructions which are linked to noise. One way of regularising is thus simply to stop the inversion iteration rapidly.

The technical side of the FETI-based inversion algorithm is represented by Figure (6.33). The

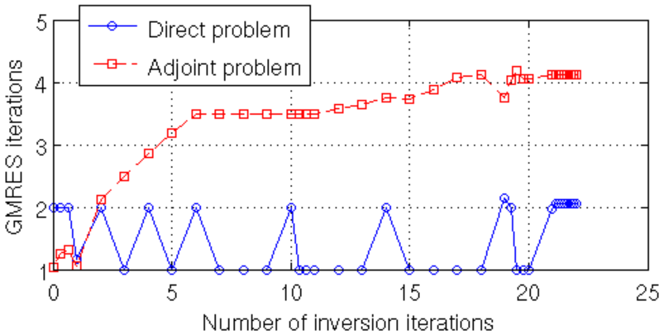


FIGURE 6.33 – Evolution of the average GMRES iterations for one source for the direct and adjoint problems.

average number of iterations for one source is varying from 1 to 2 for the direct problem and varying 1 to 4 for the adjoint problem. The approximate time of one inverse-iteration is equal to 1980 sec against 450 sec for the FEM solution. This knowledge gives us an idea of the total time of the inversion algorithm.

6.6 Conclusion

In this chapter we were dealing with the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton optimization algorithm with line search, which distinguishes itself through an approximation of the Hessian matrix in the Newton correction step with a matrix that does not involve the explicit computation of second order derivatives. Taking advantage of the Lagrangian formalism and the definition of an ad-hoc adjoint field, we calculated the gradient with respect to the permittivity. In order to make this process less depending on memory, we applied the previously discussed FETI-DPEM2-full method. Then we considered different computational aspects of the proposed method, such as the pre-initialisation and the influence of the stopping criterion. Taking this into account, we tested the FETI method on various objects from the Fresnel database. All of these tests have shown the efficiency and the gain of the FETI method.

Chapitre 7

Conclusion and perspectives

In this doctoral research, algorithms for solving two- and three-dimensional electromagnetic problems have been studied. The emphasis was on the Domain Decomposition idea which has been recognized as one of the most efficient parallel techniques in the domain of computational electromagnetics. In particular, the FETI-DPEM2 method has shown its efficiency in various implementations [43, 48, 97]. Starting with the classical formulation of the FETI-DPEM2, we have extended this method by imposing everywhere, even on the corner degrees of freedom (DOFs), more flexible Robin-type boundary conditions. The resulting interface problem was completely rewritten which lead us to a new global coarse preconditioner.

We have then presented the implementation of this modified method, denoted as FETI-DPEM2-full [55], for generic 2D and 3D electromagnetic problems. The scattered field formulation was more specifically considered. Note that we pursued different goals in these two cases. In 2D, we aimed at showing that the new treatment of the corner DOFs enables to provide more accurate results with respect to the classical FETI-DPEM2 method. We also have been able to show that essentially, by playing with the mathematical parameter α^i, arising from the Robin-type boundary condition, we would not degrade the final solution of the method. For this, we compared the results of the FETI methods with the ones obtained with the FEM method. Moreover, the results obtained in the 2D case have shown that there is no restriction in taking into account scatterers made of heterogeneous or anisotropic materials, and we are not limited in taking a strict partitioning. Indeed, we have been able to partition and handle internal interfaces which are not necessarily straight lines without loss of the solution quality.

Note that the electromagnetic problems are strongly dimension dependent and, if in 2D, the use of direct solvers for the resolution of the linear system arising from the finite-element discretization was obvious, the one related to the 3D configurations with a direct solver was much more tricky, time and memory consuming. Indeed, according to the dense structure of the FETI-DPEM related interface problem, its resolution with a direct method could be over-performed by just solving a regular FEM system of linear equations. We thus had to implement an iterative technique which, as we have shown, was not so obvious for problems containing anisotropic materials. In order to overcome the convergence related problems, the Evanescent Modes Damping algorithm (EMDA) has been applied. Indeed, this algorithm is recognised as a ”numerically simple” technique which enables to extend the transmission boundary conditions to evanescent modes in the anisotropic area. In order to validate the proposed method from a practical point of view, we then performed a comparison between the results of simulations (FETI-DPEM2-full, FEM-classical) and scattered field measurements from the Fresnel database. Finally, we have shown the effectiveness of the proposed method with respect to FEM by performing calculations on large objects with respect to the wavelength. Indeed, it was concluded that the FETI-DPEM2-full method is able to deal with much bigger domains in terms of wavelengths than the classical FEM method.

Finally, we arrive at the main application of the FETI-DPEM2-full method. It is well known that a fast forward simulator is a key-point in every electromagnetic inverse scattering problems. In this thesis, we focused on the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton optimization algorithm with line search. The finite element method previously applied for solving two direct problems at each iteration of the inverse process has been replaced by the more flexible, in terms of memory, FETI-DPEM2- full method. In order to accelerate the process of a multi-source calculation, we have studied some implementation issues of the proposed method, such as the stopping criterion or different choices of initialisation. It was concluded that, for calculating the direct problem, it is better to take the solution from the previous inversion iteration, while for calculating the adjoint problem we use the marching-on-in technique [163]. We have also proposed in this work to choose the stopping criterion of the FETI method automatically at each iteration of the inversion algorithm.

Despite the advances made in two- and three-dimensional electromagnetic problems, several improvements are still possible and should be integrated. Let us focus at first on the ”direct” part of this thesis. As far as the forward electromagnetic problem is concerned, the following suggestions can be made.

— As it was mentioned before, the domain decomposition technique has been recognized as one of the most important methodologies for constructing efficient *parallel computing* algorithms [97]. Indeed, the proposed FETI-DPEM2-full method based on the Domain Decomposition idea is highly parallelizable and can be implemented for the simulations of large-scale electromagnetic problems using massively parallel systems. Although, in this PhD thesis, we preferred to focus on the methodology and development of a new approach, rather than its numerical realisation. Nevertheless, almost a linear speed-up [97] can be expected from the parallelization of the algorithm. This, without any doubts, can be treated as the first way to improve our computational package.

— Improving the convergence properties of the iterative process constitutes the key in designing effective algorithms, in particular in mid and high frequency. Recently, a lot of effort has been put on creating various techniques based on local transmission conditions in order to improve the convergence : these include the class of FETI-H methods [36, 37, 101, 103],, the optimized Schwarz approach [164], and the evanescent modes damping algorithm [106, 165, 166]. However, the related impedance operators do not accurately approximate the exact Dirichlet-to-Neumann (DtN) operator on all the modes of the solution, which makes the resulting iterative methods suboptimal. It is thus of great interest to apply a new square-root based transmission condition, localized using Padé approximants, which accurately approximates the DtN operator, as it has been done in [45] for the classical Lions-Després domain decomposition method [25, 26, 28].

— Finally, in order to increase the domain of applications, it is necessary to ”clean” the computational code for a resolution of three-dimensional electromagnetic problems and, of course to create a more simple and clear user interface.

Concerning the quasi-Newton approach to microwave imaging, since electromagnetic measured fields essentially have a limited number of degrees of freedom in finite precision [167], the information content of the data vector is always limited and can not be increased by adding more illuminations and measurement positions. On the other hand, it is clear that increasing the resolution leads to a smaller cell size and a bigger number of DOFs. This will unfortunately result in over-fitting issues. In order to overcome this problem, it is of great interest to apply different types of regularization schemes, such as the value picking regularization [168], or the level-set approach [169, 170].

Annexe A

Calculation of derivatives of the Lagrangian

The parameter of interest, here the relative permittivity ε_r, must minimize a properly defined cost functional $\mathcal{J}(\varepsilon_r)$ (Eq. 6.2). This problem then has been reformulated into a Lagrange functional as in Eq. (A.1) in order to take into account the constraints of the Near-to-Far-Field transformation and the Helmholtz equation for the scattered field.

$$\begin{aligned}
&\mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = \\
&\quad \sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} \mathcal{L}_{s,r}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}_{s,r}, \mathcal{P}_{s,r}) = \\
&\qquad \sum_{s=1}^{N_{src}} \sum_{r=1}^{N_{rec}} \left(w_{s,r} |\mathcal{E}^{\text{far}}_{s,r} - \mathcal{E}^{\text{mes}}_{s,r}|^2 + \mathcal{R}e \left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}}) \right\rangle_{\Gamma^{\text{far}}} + \mathcal{R}e \left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r) - J^{\text{sc}} \right\rangle_{\Omega} \right)
\end{aligned} \tag{A.1}$$

The saddle point of Eq. (A.1) provides an efficient way for computing the gradient of the cost functional by introducing an adjoint field. To find this saddle point, we need to consider the following set of Karush-Kuhn-Tucker (KKT) conditions :

$$\begin{cases}
\nabla_{\mathcal{E}^{\text{far}}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\mathcal{E}^{\text{sc}}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\varepsilon_r} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\mathcal{U}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\mathcal{P}} \mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0
\end{cases} \tag{A.2}$$

Assuming that the functional $\mathcal{L}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P})$ is differentiable in the Frechet sense we will calculate the derivative in the following directions :

A.1 Derivation with respect to the far-field

To start with, we will calculate the derivative with respect to the far-field, which is equal to

$$\begin{aligned}
< \nabla_{\mathcal{E}^{\text{far}}} \mathcal{L}_{s,r}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}_{s,r}, \mathcal{P}_{s,r}),\ h >_{\Gamma^{\text{far}}} &= \\
\lim_{t \to 0} \frac{1}{t} \{ \mathcal{L}_{s,r}(\mathcal{E}^{\text{far}} + th) - \mathcal{L}_{s,r}(\mathcal{E}^{\text{far}}) \} &= \\
\mathcal{P}art1 + \mathcal{P}art2 &
\end{aligned} \tag{A.3}$$

where $\mathcal{P}art1$ can be calculated as follows :

$$\begin{aligned}
&\mathcal{P}art1 = \\
&\lim_{t \to 0} \frac{1}{t} \{ w_{s,r} |(\mathcal{E}^{\text{far}}_{s,r} + th) - \mathcal{E}^{\text{mes}}_{s,r}|^2 - w_{s,r} |\mathcal{E}^{\text{far}}_{s,r} - \mathcal{E}^{\text{mes}}_{s,r}|^2 \}
\end{aligned} \tag{A.4}$$

Let us consider the first term here, which is equal to :

$$\begin{aligned} w_{s,r}|(\mathcal{E}^{\text{far}}_{s,r}+th)-\mathcal{E}^{\text{mes}}_{s,r}|^2 =& w_{s,r}|(\mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r})+th|^2 = \\ w_{s,r}|\mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r}|^2 + w_{s,r}|th|^2 &- 2w_{s,r}\mathcal{R}e\left\langle th, \mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r}\right\rangle_{\Gamma^{\text{far}}} \end{aligned} \tag{A.5}$$

Taking this into account, $\mathcal{P}art1$ can be rewritten as follows :

$$\begin{aligned} &\mathcal{P}art1 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{w_{s,r}\left\langle th, th\right\rangle_{\Gamma^{\text{far}}} - 2w_{s,r}\mathcal{R}e\left\langle th, \mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r}\right\rangle_{\Gamma^{\text{far}}}\right\} = \\ &-2w_{s,r}\mathcal{R}e\left\langle h, \mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r}\right\rangle_{\Gamma^{\text{far}}} \end{aligned} \tag{A.6}$$

Similarly, we obtain for $\mathcal{P}art2$:

$$\begin{aligned} &\mathcal{P}art2 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}+th, \mathcal{E}^{\text{sc}})\right\rangle_{\Gamma^{\text{far}}} - \mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}})\right\rangle_{\Gamma^{\text{far}}}\right\} \end{aligned} \tag{A.7}$$

Taking advantage of Eq. (6.3) we can rewrite the equation above as follows :

$$\begin{aligned} &\mathcal{P}art2 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{E}^{\text{far}}_{s,r}+th+\oiint_{\Sigma}(...)-\mathcal{E}^{\text{far}}_{s,r}+\oiint_{\Sigma}(...)\right\rangle_{\Gamma^{\text{far}}}\right\} = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, th\right\rangle_{\Gamma^{\text{far}}}\right\} = \mathcal{R}e\left\langle \mathcal{U}_{s,r}, h\right\rangle_{\Gamma^{\text{far}}} \end{aligned} \tag{A.8}$$

Combining these two parts we obtain :

$$\boxed{\nabla_{\mathcal{E}^{\text{far}}}\mathcal{L}_{s,r} = \mathcal{U}_{s,r} - 2w_{s,r}\left(\mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{mes}}_{s,r}\right)^*, \quad \text{on } \Gamma^{\text{far}}} \tag{A.9}$$

A.2 Derivation with respect to the scattered field

Following the same idea we can calculate the derivative with respect to the scattered field :

$$\begin{aligned} < \nabla_{\mathcal{E}^{\text{sc}}}\mathcal{L}_{s,r}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}_{s,r}, \mathcal{P}_{s,r}),\ h >_{\Omega} =& \\ \lim_{t\to 0}\frac{1}{t}\{\mathcal{L}_{s,r}(\mathcal{E}^{\text{sc}}+th)-\mathcal{L}_{s,r}(\mathcal{E}^{\text{sc}})\} =& \\ \mathcal{P}art1+\mathcal{P}art2& \end{aligned} \tag{A.10}$$

where $\mathcal{P}art1$ can be obtained as follows :

$$\begin{aligned} &\mathcal{P}art1 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}}+th)\right\rangle_{\Omega} - \mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}})\right\rangle_{\Omega}\right\} = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}}+th) - \mathcal{F}(\mathcal{E}^{\text{far}}_{s,r}, \mathcal{E}^{\text{sc}}\right\rangle_{\Omega}\right\} = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{E}^{\text{far}}_{s,r}-\mathcal{E}^{\text{far}}_{s,r}+\oiint_{\Sigma}\{\mathcal{Q}(\mathcal{E}^{\text{sc}}+th)-\mathcal{Q}(\mathcal{E}^{\text{sc}})\}\ dS'\right\rangle_{\Omega}\right\} \end{aligned} \tag{A.11}$$

where $\mathcal{Q}(x)$ is the subintegral value of Eq. (6.3). Taking advantage of the linearity of the operator $\mathcal{Q}(x)$ with respect to x, we then obtain :

$$\begin{aligned} &\mathcal{P}art1 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{Q}(th)\right\rangle_\Omega\right\} = \mathcal{R}e\left\langle \mathcal{U}_{s,r}, \mathcal{Q}(h)\right\rangle_\Omega \end{aligned} \tag{A.12}$$

Note, that the operator $\mathcal{Q}$ is not selfadjoint. Moreover, it is hard to find explicitly its adjoint operator $\mathcal{Q}^*$. Physically speaking, it corresponds to the attempt of finding the scattered field on the boundary of the investigation domain Ω from the knowledge of the scattered far-field on the line Γ^{far}. That is a very difficult problem, as part of the information about the electromagnetic wave vanishes when it propagates to far distance. The second part of this equation can be calculated as :

$$\begin{aligned} &\mathcal{P}art2 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}+th, \varepsilon_r) - J^{\text{sc}}\right\rangle_\Omega - \mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r) - J^{\text{sc}}\right\rangle_\Omega\right\} \end{aligned} \tag{A.13}$$

Using the linearity of the operator $\mathcal{H}$ with respect to the first parameter, we can write :

$$\begin{aligned} &\mathcal{P}art2 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}+th-\mathcal{E}^{\text{sc}}, \varepsilon_r)\right\rangle_\Omega\right\} = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(th, \varepsilon_r)\right\rangle_\Omega\right\} = \mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(h, \varepsilon_r)\right\rangle_\Omega \end{aligned} \tag{A.14}$$

Note that unlikely the operator $\mathcal{Q}$, the operator $\mathcal{H}$ is selfadjoint with respect to the first parameter, at least in lossless media. Thus we can rewrite the equation above as follows :

$$\begin{aligned} &\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(h, \varepsilon_r)\right\rangle_\Omega = \\ &\mathcal{R}e\left\langle \mathcal{H}(\mathcal{P}_{s,r}, \varepsilon_r), h\right\rangle_\Omega \end{aligned} \tag{A.15}$$

Finally, combining Eqs. (A.12) and (A.15) we obtain the second derivative expression :

$$\boxed{\nabla_{\mathcal{E}^{\text{sc}}}\mathcal{L}_{s,r} = \mathcal{H}(\mathcal{P}_{s,r}, \varepsilon_r) + \mathcal{Q}^*(\mathcal{U}_{s,r}), \quad \text{in } \Omega} \tag{A.16}$$

A.3 Derivation with respect to the relative permittivity

Let us now calculate the third derivative, which is going to consist as well on 2 parts :

$$\begin{aligned} < \nabla_{\varepsilon_r}\mathcal{L}_{s,r}(\mathcal{E}^{\text{far}}, \mathcal{E}^{\text{sc}}, \varepsilon_r, \mathcal{U}_{s,r}, \mathcal{P}_{s,r}),\ h >_\Omega &= \\ \lim_{t\to 0}\frac{1}{t}\{\mathcal{L}_{s,r}(\varepsilon_r+th) - \mathcal{L}_{s,r}(\varepsilon_r)\} &= \\ \mathcal{P}art1 + \mathcal{P}art2 & \end{aligned} \tag{A.17}$$

This time we will start with the second part :

$$\begin{aligned} &\mathcal{P}art2 = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r+th)\right\rangle_\Omega\right\} - \lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r)\right\rangle_\Omega\right\} = \\ &\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r+th) - \mathcal{H}(\mathcal{E}^{\text{sc}}, \varepsilon_r)\right\rangle_\Omega\right\} \end{aligned} \tag{A.18}$$

Taking into account the definition of the operator $\mathcal{H}$ (Eq. 6.4) we can rewrite equation above as follows :

$$\begin{aligned}
&\mathcal{P}art2 = \\
&\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, -k_0^2[\varepsilon_r + th]\mathcal{E}^{\mathrm{sc}} + k_0^2[\varepsilon_r]\mathcal{E}^{\mathrm{sc}}\right.\right. \\
&\qquad\left.\left. -k_0^2[\varepsilon_r - \varepsilon_b + th]\mathcal{E}^{\mathrm{inc}} + k_0^2[\varepsilon_r - \varepsilon_b]\mathcal{E}^{\mathrm{inc}}\right\rangle_\Omega\right\} = \\
&\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, -k_0^2\ th\ (\mathcal{E}^{\mathrm{sc}} + \mathcal{E}^{\mathrm{inc}})\right\rangle_\Omega\right\} = \\
&\lim_{t\to 0}\frac{1}{t}\left\{\mathcal{R}e\left\langle \mathcal{P}_{s,r}, -k_0^2\ th\ \mathcal{E}^{\mathrm{tot}}\right\rangle_\Omega\right\} = \\
&\mathcal{R}e\left\langle \mathcal{P}_{s,r}, -k_0^2\ h\ \mathcal{E}^{\mathrm{tot}}\right\rangle_\Omega
\end{aligned} \tag{A.19}$$

For this derivative the first part represents the gradient of the cost function :

$$\mathcal{P}art1 = \left\langle \frac{\partial \mathcal{J}}{\partial \varepsilon_r}, h \right\rangle_\Omega \tag{A.20}$$

Combining Eq. (A.20) with the second part of the derivative (Eq. A.18), we obtain the third derivative :

$$< \nabla_{\varepsilon_r}\mathcal{L}_{s,r},\ h >_\Omega = \left\langle \frac{\partial \mathcal{J}_{s,r}}{\partial \varepsilon_r}, h \right\rangle_\Omega - \mathcal{R}e\left\langle k_0^2\ \mathcal{P}_{s,r}\ \overline{\mathcal{E}_{s,r}^{\mathrm{tot}}}, h \right\rangle_\Omega \tag{A.21}$$

From Eq. (A.21) we can obtain the explicit expression for the gradient of the cost function $\mathcal{J}$ as :

$$\boxed{\nabla_{\varepsilon_r}\mathcal{L}_{s,r} = \frac{\partial \mathcal{J}_{s,r}}{\partial \varepsilon_r} - k_0^2\ \mathcal{P}_{s,r}\ \overline{\mathcal{E}_{s,r}^{\mathrm{tot}}},\quad \text{in } \Omega} \tag{A.22}$$

A.4 Cost function derivation

Let us assume that we are at the saddle point condition. We thus have for the three derivatives above :

$$\begin{cases}
\nabla_{\mathcal{E}^{\mathrm{far}}}\mathcal{L}(\mathcal{E}^{\mathrm{far}}, \mathcal{E}^{\mathrm{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\mathcal{E}^{\mathrm{sc}}}\mathcal{L}(\mathcal{E}^{\mathrm{far}}, \mathcal{E}^{\mathrm{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0 \\
\nabla_{\varepsilon_r}\mathcal{L}(\mathcal{E}^{\mathrm{far}}, \mathcal{E}^{\mathrm{sc}}, \varepsilon_r, \mathcal{U}, \mathcal{P}) = 0
\end{cases} \tag{A.23}$$

Taking into account Eqs. (A.9), (A.16) and (A.22), we can rewrite them as

$$\begin{cases}
\dfrac{\partial \mathcal{J}_{s,r}}{\partial \varepsilon_r} = k_0^2\ \mathcal{P}_{s,r}\ \overline{\mathcal{E}_{s,r}^{\mathrm{tot}}}, \quad \text{in } \Omega \\
\mathcal{H}(\mathcal{P}_{s,r}, \varepsilon_r) = -\mathcal{Q}^*(\mathcal{U}_{s,r}), \quad \text{in } \Omega \\
\mathcal{U}_{s,r} = 2w_{s,r}\left(\mathcal{E}_{s,r}^{\mathrm{far}} - \mathcal{E}_{s,r}^{\mathrm{mes}}\right)^*, \quad \text{on } \Gamma^{\mathrm{far}}
\end{cases} \tag{A.24}$$

Because of the lack of the explicit expression for the adjoint operator $\mathcal{Q}^*$, we calculate at first the Lagrange multiplier $\mathcal{U}_{s,r}$ as it is given above in the last equation of Eq. (A.24). Then we propagate this value, which represents nothing but the discrepancy between the measured and simulated fields, from the far-field region to the investigation domain Ω. This propagation is performed assuming that $\varepsilon_r = 1$ everywhere in Ω. We thus use the classical free space Green function for obtaining $\mathcal{P}_{s,r}^{\mathrm{inc}}$ everywhere inside the investigation domain. In order to compute the total field $\mathcal{P}_{s,r}^{\mathrm{tot}}$, instead of using the second equation of Eq. (A.24) we compute the adjoint scattered field $\mathcal{P}_{s,r}^{\mathrm{sc}}$ at first as we did in Section (2.2.1) taking into account the variations in permittivity as well as the incident adjoint field. The adjoint total field $\mathcal{P}_{s,r}^{\mathrm{tot}}$ is finally obtained by summing $\mathcal{P}_{s,r}^{\mathrm{inc}}$ and $\mathcal{P}_{s,r}^{\mathrm{sc}}$.

Finally, we use the first line in Eq. (A.24) in order to find the gradient of the cost function.

$$\boxed{\frac{\partial \mathcal{J}}{\partial \varepsilon_r} = k_0^2 \, \sum_{s=1}^{N_{src}} \, \mathcal{P}_s^{\mathrm{tot}} \, \overline{\mathcal{E}_s^{\mathrm{tot}}}} \tag{A.25}$$

Bibliographie

[1] James Clerk Maxwell. *A treatise on electricity and magnetism*, volume 1. Clarendon Press, 1881. (Cited on pages 1 and 5.)

[2] C-D Munz, P Omnes, R Schneider, E Sonnendrücker, and U Voss. Divergence correction techniques for Maxwell solvers based on a hyperbolic model. *Journal of Computational Physics*, 161(2) :484–511, 2000. (Cited on page 1.)

[3] Andreas Dedner, Friedemann Kemm, Dietmar Kröner, C-D Munz, Thomas Schnitzer, and Matthias Wesenberg. Hyperbolic divergence cleaning for the mhd equations. *Journal of Computational Physics*, 175(2) :645–673, 2002. (Cited on page 1.)

[4] Kane S Yee et al. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3) :302–307, 1966. (Cited on page 1.)

[5] Allen Taflove and Korada R Umashankar. The finite-difference time-domain method for numerical modeling of electromagnetic wave interactions. *Electromagnetics*, 10(1-2) :105–126, 1990. (Cited on page 1.)

[6] M Thevenot, A Reineix, and B Jecko. A new FDTD surface impedance formalism to study PBG structures. *Microwave and optical technology letters*, 18(3) :203–206, 1998. (Cited on page 1.)

[7] Dennis M Sullivan. *Electromagnetic simulation using the FDTD method.* John Wiley & Sons, 2013. (Cited on page 1.)

[8] Paul Houston, Ilaria Perugia, Anna Schneebeli, and Dominik Schötzau. Interior penalty method for the indefinite time-harmonic Maxwell equations. *Numerische Mathematik*, 100(3) :485–518, 2005. (Cited on page 1.)

[9] Bernardo Cockburn, George E Karniadakis, and Chi-Wang Shu. *The development of discontinuous Galerkin methods.* Springer, 2000. (Cited on page 1.)

[10] David Colton and Rainer Kress. *Integral equation methods in scattering theory*, volume 72. SIAM, 2013. (Cited on page 1.)

[11] Jack Richmond. Scattering by a dielectric cylinder of arbitrary cross section shape. *IEEE Transactions on Antennas and Propagation*, 13(3) :334–341, 1965. (Cited on page 1.)

[12] Roger F Harrington. The method of moments in electromagnetics. *Journal of Electromagnetic Waves and Applications*, 1(3) :181–200, 1987. (Cited on page 1.)

[13] Andrew F Peterson, Scott L Ray, and Raj Mittra. *Computational methods for electromagnetics*, volume 24. IEEE Press, New York, 1998. (Cited on pages 1 and 8.)

[14] Jürgen De Zaeytijd. *On the 3D electromagnetic quantitative inverse scattering problem : algorithms and regularization.* PhD thesis, Ghent University, 2009. (Cited on page 1.)

[15] Prasanta Kumar Banerjee and Roy Butterfield. *Boundary element methods in engineering science*, volume 17. McGraw-Hill London, 1981. (Cited on page 1.)

[16] Jian-Ming Jin, Jiamming Jin, and Jian-Ming Jin. *The finite element method in electromagnetics.* Wiley New York, 2002. (Cited on pages 1, 5, 10 and 39.)

[17] P Collino, Gianluca Delbue, Patrick Joly, and Andrea Piacentini. A new interface condition in the non-overlapping domain decomposition method for the Maxwell equations. *Computer methods in applied mechanics and engineering*, 148(1) :195–207, 1997. (Cited on pages 1 and 2.)

[18] Youcef Saad and Martin H Schultz. Gmres : A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3) :856–869, 1986. (Cited on pages 1, 69, 70, 71 and 75.)

[19] Cornelius Lanczos. Solution of systems of linear equations by minimized iterations. *J. Res. Nat. Bur. Standards*, 49(1) :33–53, 1952. (Cited on page 1.)

[20] Henk A Van der Vorst. Bi-CGSTAB : A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM Journal on scientific and Statistical Computing*, 13(2) :631–644, 1992. (Cited on page 1.)

[21] Oliver G Ernst and Martin J Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In *Numerical Analysis of Multiscale Problems*, pages 325–363. Springer, 2012. (Cited on pages 2, 14 and 17.)

[22] Wolfgang Hackbusch. *Iterative solution of large sparse systems of equations.* Springer, 1994. (Cited on pages 2 and 14.)

[23] Yousef Saad. *Iterative methods for sparse linear systems.* Siam, 2003. (Cited on pages 2, 14, 69, 70, 71 and 75.)

[24] N Marsic and C Geuzaine. Efficient finite element assembly of high order Whitney forms. 2014. (Cited on page 2.)

[25] Pierre-Louis Lions. On the Schwarz alternating method. I. In *First international symposium on domain decomposition methods for partial differential equations*, pages 1–42. Paris, France, 1988. (Cited on pages 2 and 126.)

[26] Pierre-Louis Lions. On the Schwarz alternating method. III : a variant for nonoverlapping subdomains. In *Third international symposium on domain decomposition methods for partial differential equations*, volume 6, pages 202–223. SIAM, Philadelphia, PA, 1990. (Cited on pages 2 and 126.)

[27] Francis Collino, Souad Ghanemi, and Patrick Joly. Domain decomposition method for harmonic wave propagation : a general presentation. *Computer methods in applied mechanics and engineering*, 184(2) :171–211, 2000. (Cited on page 2.)

[28] Bruno Després. *Une méthodes de décomposition de domaine pour les problèmes de propagation d'ondes en régime harmonique. Le théorème de Borg pour l'équation de Hill vectorielle.* PhD thesis, 1991. (Cited on pages 2 and 126.)

[29] Xiao-Chuan Cai and Olof B Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM Journal on Scientific and Statistical Computing*, 13(1) :243–258, 1992. (Cited on page 2.)

[30] Xiao-Chuan Cai and Olof B Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM Journal on Numerical Analysis*, 30(4) :936–952, 1993. (Cited on page 2.)

[31] Olivier Cessenat and Bruno Despres. Une nouvelle formulation variationnelle des équations d'onde en fréquence(application au problème de helmoltz 2 d). *Rapport- C. E. A.*, 1994. (Cited on page 2.)

[32] Jean-David Benamou. A massively parallel algorithm for the optimal control of systems governed by elliptic pde's. In *PPSC*, pages 90–95, 1995. (Cited on page 2.)

[33] Jean-David Benamou. A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *SIAM journal on numerical analysis*, 33(6) :2401–2416, 1996. (Cited on page 2.)

[34] Bruno Després. Domain decomposition method and the Helmholtz problem. *Mathematical and numerical aspects of wave propagation phenomena (Strasbourg, 1991)*, pages 44–52, 1991. (Cited on page 2.)

[35] Bruno Stupfel. A fast-domain decomposition method for the solution of electromagnetic scattering by large objects. *IEEE Transactions on Antennas and Propagation*, 44(10) :1375–1385, 1996. (Cited on page 2.)

[36] Charbel Farhat, Antonini Macedo, Michel Lesoinne, Francois-Xavier Roux, Frédéric Magoulés, and Armel de La Bourdonnaie. Two-level domain decomposition methods with lagrange multipliers for the fast iterative solution of acoustic scattering problems. *Computer methods in applied mechanics and engineering*, 184(2) :213–239, 2000. (Cited on pages 2, 19, 21, 25, 26, 27, 29, 31, 34, 35 and 126.)

[37] Charbel Farhat, Antonini Macedo, and Michel Lesoinne. A two-level domain decomposition method for the iterative solution of high frequency exterior Helmholtz problems. *Numerich Mathematik*, 85(2) :283–308, 2000. (Cited on pages 2, 19, 21, 25, 26, 27 and 126.)

[38] CT Wolfe, U Navsariwala, and Stephen D Gedney. A parallel finite-element tearing and interconnecting algorithm for solution of the vector wave equation with pml absorbing medium. *IEEE Transactions on Antennas and Propagation*, 48(2) :278–284, 2000. (Cited on pages 2, 19 and 22.)

[39] Charbel Farhat, Michel Lesoinne, Patrick LeTallec, Kendall Pierson, and Daniel Rixen. FETI-DP : a dual–primal unified FETI method-part i : A faster alternative to the two-level FETI method. *International journal for numerical methods in engineering*, 50(7) :1523–1544, 2001. (Cited on pages 2, 14, 19, 21, 23, 24, 35, 72 and 77.)

[40] Charbel Farhat and Francois-Xavier Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering*, 32(6) :1205–1227, 1991. (Cited on pages 2, 21 and 22.)

[41] Yujia Li and Jian-Ming Jin. A vector dual-primal finite element tearing and interconnecting method for solving 3-d large-scale electromagnetic problems. *IEEE Transactions on Antennas and Propagation*, 54(10) :3000–3009, 2006. (Cited on pages 2, 14, 16, 19, 21, 23, 24, 31 and 72.)

[42] Charbel Farhat, Philip Avery, Radek Tezaur, and Jing Li. FETI-DPh : a dual-primal domain decomposition method for acoustic scattering. *Journal of Computational Acoustics*, 13(03) :499–524, 2005. (Cited on pages 2, 19, 21, 23 and 35.)

[43] Yu-Jia Li and Jian-Ming Jin. A new dual-primal domain decomposition approach for finite element simulation of 3-d large-scale electromagnetic problems. *IEEE Transactions on Antennas and Propagation*, 55(10) :2803–2810, 2007. (Cited on pages 2, 14, 16, 19, 25, 31, 32, 33, 35, 46, 71, 72 and 125.)

[44] Zhen Peng and Jin-Fa Lee. Non-conformal domain decomposition method with second-order transmission conditions for time-harmonic electromagnetics. *Journal of Computational Physics*, 229(16) :5615–5629, 2010. (Cited on pages 2 and 71.)

[45] Yassine Boubendir, Xavier Antoine, and Christophe Geuzaine. A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation. *Journal of Computational Physics*, 231(2) :262–280, 2012. (Cited on pages 2, 14, 19, 71, 75 and 126.)

[46] Y. Boubendir, X. Antoine, and C. Geuzaine. A non-overlapping quasi-optimized Schwarz domain decomposition algorithm for the Helmholtz equation. In R. Bank, M. Holst, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XX*, number 91 in Lecture Notes in Computational Science and Engineering, pages 519–526. Springer, 2013. (Cited on pages 2 and 75.)

[47] Bruno Stupfel. Absorbing boundary conditions on arbitrary boundaries for the scalar and vector wave equations. *IEEE Transactions on Antennas and Propagation*, 42(6) :773–780, 1994. (Cited on page 2.)

[48] Yu-Jia Li and Jian-Ming Jin. Implementation of the second-order abc in the FETI-DPEM method for 3d EM problems. *IEEE Transactions on Antennas and Propagation*, 56(8) :2765–2769, 2008. (Cited on pages 2, 32, 33, 35, 69 and 125.)

[49] F Ben Belgacem, A Buffa, and Y Maday. The mortar finite element method for 3d Maxwell equations : First results. *SIAM Journal on Numerical Analysis*, 39(3) :880–901, 2001. (Cited on page 3.)

[50] Yves Achdou, Caroline Japhet, Yvon Maday, and Frédéric Nataf. A new cement to glue non-conforming grids with robin interface conditions : The finite volume case. *Numerische Mathematik*, 92(4) :593–620, 2002. (Cited on pages 3 and 19.)

[51] Seung-Cheol Lee, Marinos N Vouvakis, and Jin-Fa Lee. A non-overlapping domain decomposition method with non-matching grids for modeling large finite antenna arrays. *Journal of Computational Physics*, 203(1) :1–21, 2005. (Cited on pages 3 and 19.)

[52] Marinos N Vouvakis, Zoltan Cendes, and J-F Lee. A FEM domain decomposition method for photonic and electromagnetic band gap structures. *IEEE Transactions on Antennas and Propagation*, 54(2) :721–733, 2006. (Cited on page 3.)

[53] Marinos Vouvakis, Kezhong Zhao, Seung-Mo Seo, and Jin-Fa Lee. A domain decomposition approach for non-conformal couplings between finite and boundary elements for unbounded electromagnetic problems in r3. *Journal of Computational Physics*, 225(1) :975–994, 2007. (Cited on page 3.)

[54] George Karypis and V Kumar. Metis : Family of multilevel partitioning algorithms. *wwwusers. cs. umn. edu/˜ karypis/metis/main. shtml*, 1995. (Cited on pages 3, 46, 66 and 77.)

[55] Ivan Voznyuk, Hervé Tortel, and Amélie Litman. Scattered field computation with an extended FETI-DPEM2 method. *Progress In Electromagnetics Research*, 139, 2013. (Cited on pages 3, 14, 32, 34 and 125.)

[56] Lawrence E Larsen and John H Jacobi. Microwave scattering parameter imagery of an isolated canine kidney. *Medical physics*, 6(5) :394–403, 1979. (Cited on page 3.)

[57] Luis Jofre, Mark S Hawley, Antoni Broquetas, Elias de Los Reyes, Miguel Ferrando, and Antonio R Elias-Fuste. Medical imaging with a microwave tomographic scanner. *IEEE Transactions on Biomedical Engineering*, 37(3) :303–312, 1990. (Cited on page 3.)

[58] KJ Langenberg, M Brandfass, P Fellinger, T Gurke, and T Kreutter. A unified theory of multidimensional electromagnetic vector inverse scattering within the kirchhoff or Born approximation. In *Radar Target Imaging*, pages 113–151. Springer, 1994. (Cited on page 3.)

[59] Avinash C.. Kak and Malcolm Slaney. *Principles of computerized tomographic imaging.* Society for Industrial and Applied Mathematics, 2001. (Cited on pages 3 and 119.)

[60] RE Kleinman and PM Van den Berg. A modified gradient method for two-dimensional problems in tomography. *Journal of Computational and Applied Mathematics*, 42(1) :17–35, 1992. (Cited on page 3.)

[61] Anthony J Devaney, Edwin A Marengo, and Fred K Gruber. Time-reversal-based imaging and inverse scattering of multiply scattering point targets. *The Journal of the Acoustical Society of America*, 118(5) :3129–3138, 2005. (Cited on page 3.)

[62] Margaret Cheney. The linear sampling method and the music algorithm. *Inverse Problems*, 17(4) :591, 2001. (Cited on page 3.)

[63] David Colton and Peter Monk. A linear sampling method for the detection of leukemia using microwaves. *SIAM Journal on Applied Mathematics*, 58(3) :926–941, 1998. (Cited on page 3.)

[64] David Colton, Houssem Haddar, and Michele Piana. The linear sampling method in inverse electromagnetic scattering theory. *Inverse problems*, 19(6) :S105, 2003. (Cited on page 3.)

[65] Andreas Kirsch. The music-algorithm and the factorization method in inverse scattering theory for inhomogeneous media. *Inverse Problems*, 18(4) :1025, 2002. (Cited on page 3.)

[66] Claire Prada, Sebastien Manneville, Dimitri Spoliansky, and Mathias Fink. Decomposition of the time reversal operator : Detection and selective focusing on two scatterers. *The Journal of the Acoustical Society of America*, 99(4) :2067–2076, 1996. (Cited on page 3.)

[67] Hervé Tortel, Gilles Micolau, and Marc Saillard. Decomposition of the time reversal operator for electromagnetic scattering. *Journal of Electromagnetic Waves and Applications*, 13(5) :687–719, 1999. (Cited on page 3.)

[68] Xiao-Yun Zhang, Hervé Tortel, Amélie Litman, and Jean-Michel Geffrin. An extended-dort method and its application in a cavity configuration. *Inverse Problems*, 28(11) :115008, 2012. (Cited on page 3.)

[69] Peter M Van Den Berg and Ralph E Kleinman. A contrast source inversion method. *Inverse problems*, 13(6) :1607, 1997. (Cited on page 4.)

[70] Peter M van den Berg, AL Van Broekhoven, and Aria Abubakar. Extended contrast source inversion. *Inverse Problems*, 15(5) :1325, 1999. (Cited on page 4.)

[71] Aria Abubakar, Peter M Van den Berg, and Jordi J Mallorqui. Imaging of biomedical data using a multiplicative regularized contrast source inversion method. *IEEE Transactions on Microwave Theory and Techniques*, 50(7) :1761–1771, 2002. (Cited on page 4.)

[72] Aria Abubakar, Tarek M Habashy, Peter M van den Berg, and Dries Gisolf. The diagonalized contrast source approach : an inversion method beyond the Born approximation. *Inverse problems*, 21(2) :685, 2005. (Cited on page 4.)

[73] WC Chew and YM Wang. Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method. *IEEE Transactions on Medical Imaging*, 9(2) :218–225, 1990. (Cited on page 4.)

[74] Nadine Joachimowicz, Christian Pichot, and J-P Hugonin. Inverse scattering : An iterative numerical method for electromagnetic imaging. *IEEE Transactions on Antennas and Propagation*, 39(12) :1742–1753, 1991. (Cited on page 4.)

[75] Ann Franchois and Christian Pichot. Microwave imaging-complex permittivity reconstruction with a levenberg-marquardt method. *IEEE Transactions on Antennas and Propagation*, 45(2) :203–215, 1997. (Cited on page 4.)

[76] Anton G Tijhuis, Kamal Belkebir, Amelie CS Litman, and Bastiaan P de Hon. Theoretical and computational aspects of 2-d inverse profiling. *IEEE Transactions on Geoscience and Remote Sensing*, 39(6) :1316–1330, 2001. (Cited on page 4.)

[77] Ann Franchois and AG Tijhuis. A quasi-newton reconstruction algorithm for a complex microwave imaging scanner environment. *Radio Science*, 38(2) :VIC–12, 2003. (Cited on pages 4 and 96.)

[78] Line Garnero, Ann Franchois, J-P Hugonin, Christian Pichot, and N Joachimowicz. Microwave imaging-complex permittivity reconstruction-by simulated annealing. *IEEE Transactions on Microwave Theory and Techniques*, 39(11) :1801–1807, 1991. (Cited on page 4.)

[79] Salvatore Caorsi, Andrea Massa, and Matteo Pastorino. A computational technique based on a real-coded genetic algorithm for microwave imaging purposes. *IEEE Transactions on Geoscience and Remote Sensing*, 38(4) :1697–1708, 2000. (Cited on page 4.)

[80] JM Geffrin and Pierre Sabouroux. Continuing with the fresnel database : experimental setup and improvements in 3d scattering measurements. *Inverse Problems*, 25(2) :024001, 2009. (Cited on pages 4, 64, 81, 83, 96 and 113.)

[81] Amélie Litman, Lorenzo Crocco, et al. Testing inversion algorithms againts experimental data : 3d targets. *Inverse Problems*, 25(2) :1–20, 2009. (Cited on pages 4, 96, 102 and 113.)

[82] Roger F Harrington. Time-harmonic electromagnetic fields. 1961. (Cited on page 5.)

[83] Jean G Van Bladel. *Electromagnetic fields*, volume 19. John Wiley & Sons, 2007. (Cited on page 5.)

[84] Magdy F Iskander. *Electromagnetic fields and waves.* Waveland Press, 2013. (Cited on page 5.)

[85] Thomas BA Senior and John Leonidas Volakis. *Approximate boundary conditions in electromagnetics.* Number 41. Iet, 1995. (Cited on pages 5 and 8.)

[86] Bevan B Baker and Edward Thomas Copson. *The mathematical theory of Huygens' principle*, volume 329. American Mathematical Soc., 2003. (Cited on page 5.)

[87] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2) :185–200, 1994. (Cited on pages 5, 10 and 64.)

[88] Julius Adams Stratton. *Electromagnetic theory*, volume 33. John Wiley & Sons, 2007. (Cited on page 5.)

[89] John Kraus. *Electromagnetics.* McGraw-Hill, 1992. (Cited on page 5.)

[90] Nathan Ida. *Engineering electromagnetics.* Springer, 2004. (Cited on page 8.)

[91] Arnold Sommerfeld. *Partial differential equations in physics*, volume 1. Academic Press, 1949. (Cited on page 8.)

[92] Weng Cho Chew and William H Weedon. A 3d perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and optical technology letters*, 7(13) :599–604, 1994. (Cited on page 10.)

[93] Zachary S Sacks, David M Kingsland, Robert Lee, and Jin-Fa Lee. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Transactions on Antennas and Propagation*, 43(12) :1460–1463, 1995. (Cited on page 10.)

[94] Richard Courant et al. Variational methods for the solution of problems of equilibrium and vibrations. *Bull. Amer. Math. Soc*, 49(1) :1–23, 1943. (Cited on page 14.)

[95] Iain S Duff, Albert Maurice Erisman, and John Ker Reid. *Direct methods for sparse matrices.* Clarendon Press, Oxford, 1986. (Cited on page 14.)

[96] Ming-Feng Xue and Jian-Ming Jin. Nonconformal FETI-DP methods for large-scale electromagnetic simulation. *IEEE Transactions on Antennas and Propagation*, 60(9) :4291–4305, 2012. (Cited on pages 14, 32, 35, 69 and 71.)

[97] Yu-Jia Li and Jian-Ming Jin. Parallel implementation of the FETI-DPEM algorithm for general 3d EM simulations. *Journal of Computational Physics*, 228(9) :3255–3267, 2009. (Cited on pages 14, 33, 35, 69, 71, 77, 78, 79, 90, 125 and 126.)

[98] Ted Belytschko, Yun Yun Lu, and Lei Gu. Element-free Galerkin methods. *International journal for numerical methods in engineering*, 37(2) :229–256, 1994. (Cited on page 15.)

[99] Raphaël Lencrerot. *Outils de modélisation et d'imagerie pour un scanner micro-onde : Application au contrôle de la teneur en eau d'une colonne de sol.* PhD thesis, Université Paul Cézanne-Aix-Marseille III, 2008. (Cited on pages 16 and 99.)

[100] Charbel Farhat, Po-Shu Chen, and Francois-Xavier Roux. The dual Schur complement method with well-posed local neumann problems : regularization with a perturbed lagrangian formulation. *SIAM Journal on Scientific Computing*, 14(3) :752–759, 1993. (Cited on page 19.)

[101] Armel de La Bourdonnaye, Charbel Farhat, Antonini Macedo, Frédéric Magoulès, François-Xavier Roux, et al. A non-overlapping domain decomposition method for the exterior Helmholtz problem. *Contemporary Mathematics*, 218(2) :42–66, 1998. (Cited on pages 19, 25 and 126.)

[102] Patrick Le Tallec. Domain decomposition methods in computational mechanics. *Computational mechanics advances*, 1(2) :121–220, 1994. (Cited on page 19.)

[103] Charbel Farhat, Radek Tezaur, and Jari Toivanen. A domain decomposition method for discontinuous Galerkin discretizations of Helmholtz problems with plane waves and lagrange multipliers. *International journal for numerical methods in engineering*, 78(13) :1513–1531, 2009. (Cited on pages 19 and 126.)

[104] Victorita Dolean, Martin Gander, and Luca Gerardo-Giorda. Optimized Schwarz methods for Maxwell equations. *arXiv preprint math/0610531*, 2006. (Cited on page 19.)

[105] F Brezzi LD Marini. A three-field domain decomposition method. In *Domain Decomposition Methods in Science and Engineering : The Sixth International Conference on Domain Decomposition, June 15-19, 1992, Como, Italy*, volume 157, page 27. American Mathematical Soc., 1994. (Cited on pages 19 and 27.)

[106] Y Boubendir, A Bendali, and MB Fares. Coupling of a non-overlapping domain decomposition method for a nodal finite element method with a boundary element method. *International journal for numerical methods in engineering*, 73(11) :1624–1650, 2008. (Cited on pages 20, 64, 75 and 126.)

[107] Victorita Dolean, Mohamed El Bouajaji, Martin J Gander, Stéphane Lanteri, and Ronan Perrussel. Domain decomposition methods for electromagnetic wave propagation problems in heterogeneous media and complex domains. In *Domain Decomposition Methods in Science and Engineering XIX*, pages 15–26. Springer, 2011. (Cited on page 20.)

[108] Charbel Farhat. Implicit parallel processing in structural mechanics. *Compt. Mech. Adv.*, 2 :1–124, 1994. (Cited on pages 21 and 22.)

[109] M Lesoinne and K Pierson. 44. FETI-DP : An efficient, scalable and unified dual-primal FETI method. 1999. (Cited on pages 21 and 23.)

[110] Diethard Klatte and Klaus Tammer. Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization. *Annals of Operations Research*, 27(1) :285–307, 1990. (Cited on pages 22 and 98.)

[111] Jian-Ming Jin and Douglas J Riley. *Finite element analysis of antennas and arrays*. John Wiley & Sons, 2009. (Cited on pages 24, 32, 33, 35, 36, 37, 46 and 64.)

[112] F-X Roux, F Magoules, L Series, and Y Boubendir. Approximation of optimal interface boundary conditions for two-lagrange multiplier FETI method. In *Domain Decomposition Methods in Science and Engineering*, pages 283–290. Springer, 2005. (Cited on page 26.)

[113] Ming-Lin Yang and Xin-Qing Sheng. On the finite element tearing and interconnecting method for scattering by large 3d inhomogeneous targets. *International Journal of Antennas and Propagation*, 2012, 2011. (Cited on pages 32, 33 and 72.)

[114] H. Whitney. *Geometric Integration Theory*. Princeton University Press, 1957. (Cited on page 39.)

[115] J.C. Nedelec. Mixed finite elements in R3. *Numerische Mathematik*, 35 :315–341, 1980. (Cited on page 39.)

[116] Weng Cho Chew. *Waves and fields in inhomogeneous media*, volume 522. IEEE Press, New York, 1995. (Cited on page 46.)

[117] Patrick R Amestoy, Iain S Duff, and J-Y L'Excellent. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer methods in applied mechanics and engineering*, 184(2) :501–520, 2000. (Cited on pages 46, 64, 66 and 77.)

[118] Christophe Geuzaine and Jean-François Remacle. Gmsh : A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11) :1309–1331, 2009. (Cited on page 46.)

[119] Patrick R Amestoy, Iain S Duff, Jean-Yves L'Excellent, and Jacko Koster. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23(1) :15–41, 2001. (Cited on pages 46 and 64.)

[120] Jean-David Benamou and Bruno Després. A domain decomposition method for the Helmholtz equation and related optimal control problems. *Journal of Computational Physics*, 136(1) :68–82, 1997. (Cited on page 53.)

[121] Timothy A Davis. Algorithm 832 : Umfpack v4. 3—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software (TOMS)*, 30(2) :196–199, 2004. (Cited on page 64.)

[122] Walter Edwin Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics*, 9(1) :17–29, 1951. (Cited on page 69.)

[123] B. Despres, P. Joly, and J. Roberts. A domain decomposition method for the harmonic Maxwell's equations. In *IMACS, International Symposium on Iterative Methods in Linear Algebra*, pages 475–484, 1992. (Cited on pages 75 and 94.)

[124] Radek Tezaur, Antonini Macedo, and Charbel Farhat. Iterative solution of large-scale acoustic scattering problems with multiple right hand-sides by a domain decomposition method with lagrange multipliers. *International Journal for Numerical Methods in Engineering*, 51(10) :1175–1193, 2001. (Cited on page 77.)

[125] Intel. *Intel Math Kernel Library*, August 2009. (Cited on page 78.)

[126] Christelle Eyraud, Jean-Michel Geffrin, Pierre Sabouroux, Patrick C Chaumet, Hervé Tortel, Hugues Giovannini, and Amélie Litman. Validation of a 3d bistatic microwave scattering measurement setup. *Radio Science*, 43(4) :2737–2746, 2008. (Cited on pages 81 and 96.)

[127] Jeremi Dauchet. *Analyse radiative des photobioréacteurs.* PhD thesis, Institut Pascal, 2012. (Cited on pages 88 and 90.)

[128] Simon R Arridge. Optical tomography in medical imaging. *Inverse problems*, 15(2) :R41, 1999. (Cited on page 96.)

[129] Abhijit J Chaudhari, Felix Darvas, James R Bading, Rex A Moats, Peter S Conti, Desmond J Smith, Simon R Cherry, and Richard M Leahy. Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging. *Physics in medicine and biology*, 50(23) :5421, 2005. (Cited on page 96.)

[130] Martin Schweiger, Simon R Arridge, and Ilkka Nissilä. Gauss-newton method for image reconstruction in diffuse optical tomography. *Physics in medicine and biology*, 50(10) :2365, 2005. (Cited on page 96.)

[131] Gildas Marin, Christophe Guerin, Sylvain Baillet, Line Garnero, and Gérard Meunier. Influence of skull anisotropy for the forward and inverse problem in eeg : simulation studies using FEM on realistic head models. *Human brain mapping*, 6(4) :250–269, 1998. (Cited on page 96.)

[132] Michael Seger, Gerald Fischer, Robert Modre, Bernd Messnarz, Friedrich Hanser, and Bernhard Tilg. Lead field computation for the electrocardiographic inverse problem-finite elements versus boundary elements. *Computer methods and programs in biomedicine*, 77(3) :241–252, 2005. (Cited on page 96.)

[133] Assad A Oberai, Nachiket H Gokhale, and Gonzalo R Feijoo. Solution of inverse problems in elasticity imaging using the adjoint method. *Inverse Problems*, 19(2) :297, 2003. (Cited on page 96.)

[134] Jia Lu, Xianlian Zhou, and Madhavan L Raghavan. Inverse elastostatic stress analysis in pre-deformed biological structures : demonstration using abdominal aortic aneurysms. *Journal of biomechanics*, 40(3) :693–696, 2007. (Cited on page 96.)

[135] Fabio Arizzi and Egidio Rizzi. Elastoplastic parameter identification by simulation of static and dynamic indentation tests. *Modelling and Simulation in Materials Science and Engineering*, 22(3) :035017, 2014. (Cited on page 96.)

[136] Guizhi Xu, Huanli Wu, Shuo Yang, Shuo Liu, Ying Li, Qingxin Yang, Weili Yan, and Mingshi Wang. 3-d electrical impedance tomography forward problem with finite element method. *IEEE Transactions on Magnetics*, 41(5) :1832–1835, 2005. (Cited on page 96.)

[137] Frank Hettlich. Fréchet derivatives in inverse obstacle scattering. *Inverse problems*, 11(2) :371, 1995. (Cited on pages 96 and 98.)

[138] CG Xie, SM Huang, MS Beck, BS Hoyle, R Thorn, C Lenn, and D Snowden. Electrical capacitance tomography for flow imaging : system model for development of image reconstruction algorithms and design of primary sensors. *IEE Proceedings G (Circuits, Devices and Systems)*, 139(1) :89–98, 1992. (Cited on page 96.)

[139] Manuchehr Soleimani, William RB Lionheart, Antony J Peyton, Xiandong Ma, and Stuart R Higson. A three-dimensional inverse finite-element method applied to experimental eddy-current imaging data. *IEEE Transactions on Magnetics*, 42(5) :1560–1567, 2006. (Cited on page 96.)

[140] Amer Zakaria, Ian Jeffrey, and Joe LoVetri. Full-vectorial parallel finite-element contrast source inversion method. *Progress In Electromagnetics Research*, 142 :463–483, 2013. (Cited on page 96.)

[141] Ioannis T Rekanos, Traianos V Yioultsis, and Theodoros D Tsiboukis. Inverse scattering using the finite-element method and a nonlinear optimization technique. *IEEE Transactions on Microwave Theory and Techniques*, 47(3) :336–344, 1999. (Cited on page 96.)

[142] W Rachowicz and A Zdunek. Application of the FEM with adaptivity for electromagnetic inverse medium scattering problems. *Computer Methods in Applied Mechanics and Engineering*, 200(29) :2337–2347, 2011. (Cited on page 96.)

[143] Christoph Schwarzbach and Eldad Haber. Finite element based inversion for time-harmonic electromagnetic problems. *Geophysical Journal International*, 193(2) :615–634, 2013. (Cited on page 96.)

[144] Amelie Litman, R Lencrerot, and Jean-Michel Geffrin. Combining spatial support information and shape-based method for tomographic imaging inside a microwave cylindrical scanner. *Inverse Problems in Science and Engineering*, 18(1) :19–34, 2010. (Cited on page 96.)

[145] R Lencrerot, Amelie Litman, Hervé Tortel, and J-M Geffrin. Measurement strategies for a confined microwave circular scanner. *Inverse Problems in Science and Engineering*, 17(6) :787–802, 2009. (Cited on pages 96 and 99.)

[146] R Lencrerot, Amelie Litman, Hervé Tortel, and Jean-Michel Geffrin. Imposing zernike representation for imaging two-dimensional targets. *Inverse Problems*, 25(3) :035012, 2009. (Cited on pages 96 and 99.)

[147] Octavien Cmielewski, Hervé Tortel, Amelie Litman, and Marc Saillard. A two-step procedure for characterizing obstacles under a rough surface from bistatic measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 45(9) :2850–2858, 2007. (Cited on page 96.)

[148] E. Attardo, G. Vecchi, and L. Crocco. A new hybrid FEM-IE inversion method for helmholtz scalar problems. In *AP-URSI*, 2013. (Cited on page 96.)

[149] Ann Franchois and AG Tijhuis. Quantitative microwave imaging in a complex environment. In *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA 01), September 10-14, 2001, Torino, Italy*, pages 519–522, 2001. (Cited on page 96.)

[150] Weng Chow Chew, YM Wang, G Otto, Dominique Lesselier, and Jean-Charles Bolomey. On the inverse source method of solving inverse scattering problems. *Inverse Problems*, 10(3) :547, 1994. (Cited on page 97.)

[151] Christelle Eyraud, A Litman, Alain Hérique, and Wlodek Kofman. Microwave imaging from experimental data within a bayesian framework with realistic random noise. *Inverse problems*, 25(2) :024005, 2009. (Cited on page 98.)

[152] Thomas John Connolly and David JN Wall. On fréchet differentiability of some nonlinear operators occurring in inverse problems : an implicit function theorem approach. *Inverse Problems*, 6(6) :949, 1990. (Cited on page 98.)

[153] Amélie Litman and Kamal Belkebir. Two-dimensional inverse profiling problem using phaseless data. *JOSA A*, 23(11) :2737–2746, 2006. (Cited on page 99.)

[154] A Van Den Bos. Complex gradient and hessian. In *Vision, Image and Signal Processing, IEE Proceedings-*, volume 141, pages 380–383. IET, 1994. (Cited on page 100.)

[155] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1-3) :503–528, 1989. (Cited on page 100.)

[156] Roger Fletcher. *Practical methods of optimization.* John Wiley & Sons, 2013. (Cited on page 100.)

[157] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5) :1190–1208, 1995. (Cited on page 100.)

[158] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. L-bfgs-b : a limited memory fortran code for solving bound constrained optimization problems. *Dept. of Electrical Engineering and Computer Science, Northwestern Univ., TR NAM-11, Evanston, IL*, 1994. (Cited on page 100.)

[159] Christelle Eyraud, Jean-Michel Geffrin, Amélie Litman, and Jean-Pierre Spinelli. A large 3d target with small inner details : A difficult cocktail for imaging purposes without a priori knowledge on the scatterers geometry. *Radio Science*, 47(1), 2012. (Cited on page 102.)

[160] WC Chew and JH Lin. A frequency-hopping approach for microwave imaging of large inhomogeneous bodies. *IEEE Microwave and Guided Wave Letters*, 5(12) :439–441, 1995. (Cited on page 104.)

[161] Ovidio M Bucci and Giorgio Franceschetti. On the degrees of freedom of scattered fields. *IEEE Transactions on Antennas and Propagation*, 37(7) :918–926, 1989. (Cited on page 104.)

[162] Ovidio Mario Bucci, Lorenzo Crocco, Tommaso Isernia, and Vito Pascazio. Subsurface inverse scattering problems : quantifying, qualifying, and achieving the available information. *IEEE Transactions on Geoscience and Remote Sensing, Year = 2001, Number = 11, Pages = 2527–2538, Volume = 39, Owner = voznyuk, Publisher = IEEE, Timestamp = 2014.06.30.* (Cited on page 104.)

[163] ZQ Peng and AG Tijhuis. Transient scattering by a lossy dielectric cylinder : Marching-on-in-frequency approach. *Journal of electromagnetic waves and applications*, 7(5) :739–763, 1993. (Cited on pages 109 and 126.)

[164] Martin J Gander, Frédéric Magoules, and Frédéric Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 24(1) :38–60, 2002. (Cited on page 126.)

[165] Abderrahmane Bendali, Yassine Boubendir, and M Fares. A FETI-like domain decomposition method for coupling finite elements and boundary elements in large-size problems of acoustic scattering. *Computers & structures*, 85(9) :526–535, 2007. (Cited on page 126.)

[166] Yassine Boubendir. An analysis of the BEM-FEM non-overlapping domain decomposition method for a scattering problem. *Journal of computational and applied mathematics*, 204(2) :282–291, 2007. (Cited on page 126.)

[167] Ovidio M Bucci, Claudio Gennarelli, and Catello Savarese. Representation of electromagnetic fields over arbitrary surfaces by a finite and nonredundant number of samples. *IEEE Transactions on Antennas and Propagation*, 46(3) :351–359, 1998. (Cited on page 126.)

[168] Jürgen De Zaeytijd, Ann Franchois, and J-M Geffrin. A new value picking regularization strategy—application to the 3-d electromagnetic inverse scattering problem. *IEEE Transactions on Antennas and Propagation*, 57(4) :1133–1149, 2009. (Cited on page 126.)

[169] Amelie Litman, Dominique Lesselier, and Fadil Santosa. Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set. *Inverse problems*, 14(3) :685, 1998. (Cited on page 126.)

[170] Amelie Litman. Reconstruction by level sets of n-ary scattering obstacles. *Inverse Problems*, 21(6) :S131, 2005. (Cited on page 126.)