CHALMERS
UNIVERSITY OF TECHNOLOGY

UNIVERSITY OF GOTHENBURG

MASTER'S THESIS

Iteratively Regularized Adaptive Finite Element Method for Reconstruction of Coefficients in Maxwell's System

SAMAR HOSSEINZADEGAN

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2015.

Iteratively Regularized Adaptive Finite Element Method for Reconstruction of Coefficients in Maxwell's System

Department of Mathematical Sciences
Samar Hosseinzadegan

Supervisor and Examiner : Dr. Larisa Beilina

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2015.

Acknowledgments

First and foremost I offer my sincerest gratitude to my supervisor, Larisa Beilina, who has supported me throughout my thesis with her patience, motivation, enthusiasm, and immense knowledge. I am grateful to her for introducing me to such an interesting research project. One simply could not wish for a better or friendlier supervisor.

Many thanks go to all members of the Applied Mathematics department who supported me during my studies. Specifically, I am grateful to Mohammad Asadzadeh who is really helpful in building up my theoretical foundations in partial differential equations.

Last but not the least, I would like to thank my parents Fariba and Hamid Reza, and my sisters Dorreh and Samane, for their constant love and support during my life.

Abstract

We consider iteratively regularized adaptive finite element method for reconstruction of spatially distributed dielectric permittivity and magnetic permeability functions, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$, simultaneously, using time-dependent backscattering data. These functions are unknown coefficients in Maxwell's system of equations. We formulate our problem as the coefficient inverse problem (CIP) for the full Maxwell's system. To solve our inverse problem we minimize Tikhonov regularization functional on the locally adaptively refined meshes.

In this work, we consider and compare different techniques for choosing regularization parameter in the Tikhonov functional in order to get improved solution of our inverse problem. Our goal is to choose optimized regularization parameters in the solution of our CIP. This means, we choose regularization parameters and parameters in the set up of the program such that we will get best reconstructions of functions $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ to our backscattering data of the electric field on every iteration of the optimization procedure. Our numerical work consist in the reconstruction of unknown coefficients $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$, on the adaptivity locally refined meshes. Software packages WavES [60] and PETSc [50] are used for computations of reconstructions of these functions. Simulations are done on resources at Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

1 Introduction

In this work we consider a coefficient inverse problem (CIP) for Maxwell's system in 3-D. This is continuation of the recent research on this topic presented in [15]. Both spatially distributed functions of this equation, dielectric permittivity $\varepsilon(\mathbf{x})$ and magnetic permeability $\mu(\mathbf{x})$, are of our interest and we try to reconstruct simultaneously these two functions using time-dependent backscattering data of the electric field $E(\mathbf{x},\, t)$. Here we are dealing only with a single measurement data of the electric field $E(x, t)$. This means, that we are working only with a single direction of the propagation of a plane wave.

The numerical procedure for the reconstruction of unknown coefficients, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$, is formulated as an optimization problem. To solve forward and adjoint problems in the optimization problem we use hybrid finite element/difference approach of [13].

The theoretical part related to stability results for our CIP have been proposed by Dirichlet-to-Neumann map for the case of multiply measurements or by Carleman estimates for the case of finite measurements. We refer to [16, 22, 23, 38, 47, 52] for results using Dirichlet-to-Neumann map with an infinite number of boundary observations. See also [18, 29, 37, 41, 42] for various Carleman estimation results.

As we already mentioned, in our study we work with a CIP where data are generated by a single propagation of a plane wave. The theoretical stability results for a such kind of CIPs for Maxwell's equations involving both electric $E(\mathbf{x},\, t)$ and magnectic $H(\mathbf{x},\, t)$ fields are presented in [41, 42]. Stability and uniqueness results for both coefficients $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ when we have observations of only the electric field $E(\mathbf{x})$ are presented in [15, 18].

The applications of our CIP are, for example, in airport security system, imaging of land mines [10, 11, 12], reconstructing the electromagnetic parameters in nanocomposites or artificial material [51, 53, 54], imaging of defects and their sizes in a non-destructive testing of materials and in photonic crystals [25], measurement of the moisture content [24] and drying processes [44].

The main new impact of our work compared with [15] is that we have improved reconstructions of $\varepsilon,\ \mu$ using the optimization technique of [14]. We apply a posteriori error estimates for the error in the reconstructed coefficients and in the Tikhonov functional using iterative choice of the regularization parameters appearing in this functional. Our computations show that using iteratively regularized adaptive finite element method we can significantly improve shapes and locations of reconstructed functions.

In this work, we reconstruct coefficients $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ by finding the stationary point of the a Lagrangian involving the solution of a forward equation (the state equation), a backward equation (the adjoint equation) and two equations expressing that the gradients with respect to the coefficients, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ vanish.

To do that, we seek a solution for the forward and backward equations and update both coefficients $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ at every step of our iterations in a conjugate gradient method. To solve forward and backward equations we consider numerical solution of the Maxwell's system. There exist different numerical techniques to solve the Maxwell's equations, for example, the edge elements of Nédélec [46], the node-based first-order formulation of Lee and Madsen [40], the node-based curl-curl formulation with divergence condition of Paulsen and Lynch [48] or the interior-penalty discontinuous Galerkin FEM [33]. In the finite element computations, the edge elements are the most satisfactory from the theoretical point of view [43] since they automatically satisfy the divergence free condition. On the other hand, since the solution of a linear system is required at every time iteration and this could be time consuming we are not interested in them for time-dependent computations. Besides, explicit time stepping scheme cannot be used generally since the results from mass-lumping may not lead to a matrix with strictly positive entries on its diameter [30].

In our work, we use the stabilized domain decomposition finite element/finite difference approach developed in [9] and further extended in [13] to the case which we use, with divergence free condition of Paulsen and Lynch [48]. This condition allows to remove the spurious solutions from the computational solution during the local mesh refinement in the case when discontinuities in material are not too big [13]. In our implementations, we use the software package WavES [60] and PETSc [50] with MPI (message passing interface) due to its efficiency and convenience.

The outline of our work is following. In section 2 we describe topic of Inverse Problems and application of it in the science and engineering. In sections 3-6 we explain definitions and mathematical tools, generally theoretical part applied in our computations. In section 7 we describe the methods applied for the iterative choice of regularization parameters in the Tikhonov functional. The Maxwell's system of equations is presented in section 8. In section 9 we present the model problem for the Maxwell's equations as well as domain decomposition, forward problem and numerical methods applied for the Maxwell's system solution. We state our CIP, the optimization approach which is used for its solution and describe numerical methods applied to solution of our CIP in section 10. In section 11 we present general framework of a posteriori error estimate, formulate a posteriori error estimates for the regularized solution and for the Tikhonov functional. Adaptive algorithms for reconstruction of both coefficients are presented in section 12. Finally, we show numerical results obtained in our computations using adaptive algorithms in section 13.

2 Inverse Problems: Theory and Applications in Science and Engineering

2.1 Scattering Idea

The main goal of this master thesis is reconstruction of unknown parameters in Maxwell's equations using solution of the corresponding Inverse Problem. Here some questions arise when we come to an Inverse Problem. The first question is " What is an Inverse Problem?" and we try to answer to this question with a simple example taken in [34, 61, 62] which we present in this section .

Inverse Problems is a research area dealing with inversion of models or data which arises naturally in many branches of science and engineering where the values of some model parameters must be obtained from the observed data. This means that when we want to gain information about a physical object or system from observed measurements, we can consider a mathematical framework named Inverse Problem. Theory and applications of inverse problems have experienced a considerable growth in the last few years. Inverse problems have applications in science and mathematics, including computer vision, natural language processing, machine learning, statistics, statistical inference, geophysics, medical imaging (such as computed axial tomography and EEG/ERP), remote sensing, ocean acoustic tomography, nondestructive testing, acoustics, aerodynamics, electromagnetics, hydrological engineering, image analysis, shape design, structural dynamic modification and reconstruction, tomography and many other fields.

Potential topics include, but are not limited to:

- Inverse problems for differential and integral equations,
- Inverse problems and homogenization techniques,
- Regularization techniques,
- Statistical inverse problems,
- Numerical algorithms for inverse problems,
- Fractal-based methods in inverse problems.

Solution of Inverse Problems (IPs) is important due to providing information about a physical parameters that we cannot observe directly or it is very difficult and not reasonable for us to solve them directly.

2.2 Example of an Inverse problem

Here we show an example of Inverse Problem which arises in geophysics [62] and uses multiply measurements. Imagine a situation where we want to gain information about the structure of the wave speed g inside

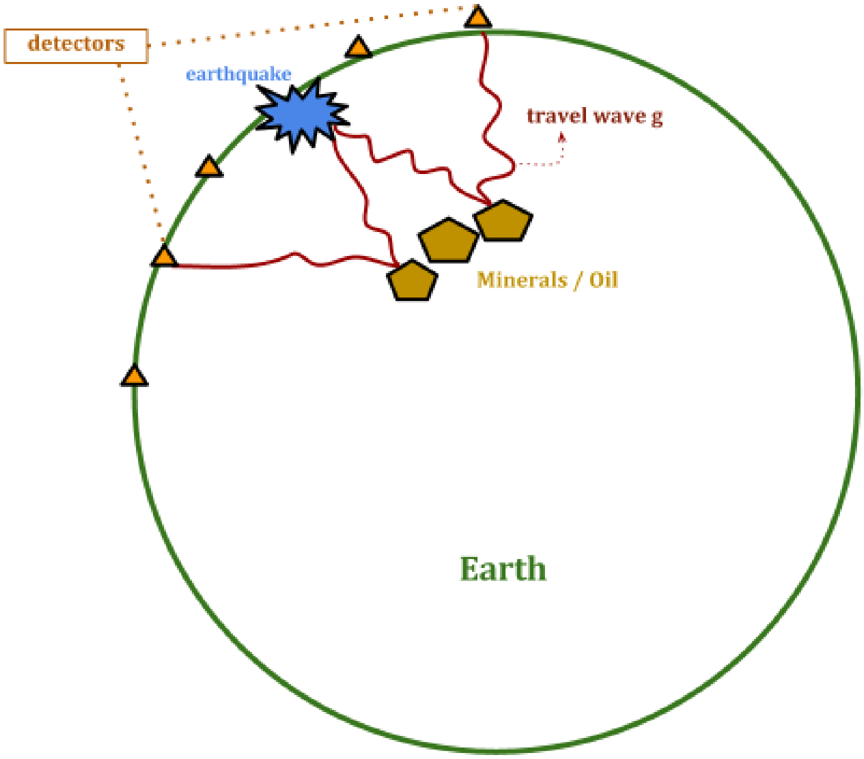


Figure 1: *Earthquake in the Earth.*

the earth Ω, from the observed seismic wave-field, denoted by Λ_g at the boundary of the earth $\partial\Omega$. We just consider Λ_g as the available measurements at the surface of the earth.

In Figure 1, we see the traveling seismic waves g produced by an earthquake inside the earth. Having knowledge about the traveling seismic waves provides information about the location of oil and minerals deposits. To gain this information, we need to perturb the boundary of the earth by an artificial explosion or by a natural earthquake. This perturbation produces waves that travels through geodesics in the metric g and hits the oil and mineral deposits and reflects back to the surface of the earth. We wait up to time T until the waves have reached the boundary. After that we are able to measure the backscattered wave-field using a seismograph.

The wave equation in the bounded domain $\Omega \in \mathbf{R}^3$ with the boundary $\partial\Omega$ illustrates the mathematical formulation of the problem described above:

$$\begin{cases} u_{tt} - \nabla \cdot (g \nabla u) = 0, & \text{in} \quad \Omega \times (0, T), \\ u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0, & \text{for} \quad \mathbf{x} \in \Omega, \\ u(\mathbf{x}, t) = f(\mathbf{x}, t), & \text{on} \quad \partial\Omega \times (0, T), \end{cases} \tag{1}$$

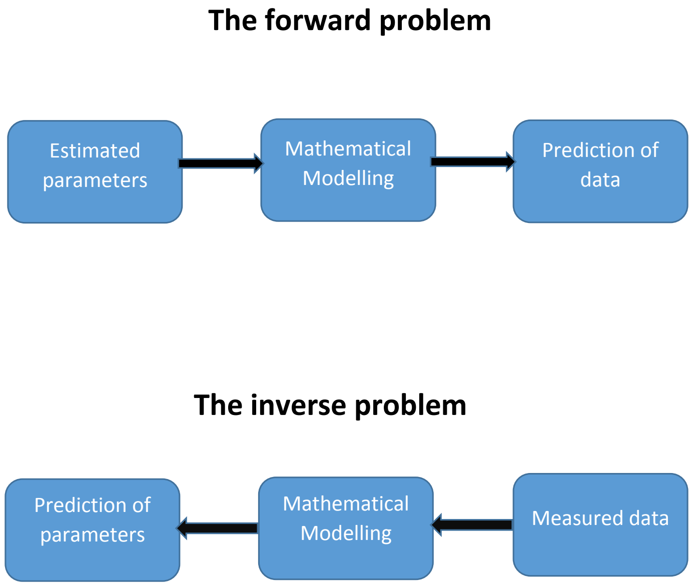


Figure 2: *Flowchart of forwards and inverse problem.*

where $u(\mathbf{x}, t)$ is the displacement in some direction of the point $\mathbf{x}$ and the time t, where T is the final time. The information of the seismic wave field is encoded in the hyperbolic Dirichlet to Neumann map

$$\Lambda_g : f \to \frac{\partial u}{\partial n}\bigg|_{\partial\Omega} \tag{2}$$

Here, n denotes the outer unit normal to $\partial\Omega$ and Λ_g maps the initial perturbation f. Dirichlet condition is applied to the recovered wave-field $\frac{\partial u}{\partial n}|_{\partial\Omega}$. Neumann condition carries within information of the inside metric of the earth.

In this example, the aim is to recover g from Λ_g via the forward problem which is an well-posed problem under regularity assumptions over g. This means that the Neumann data $\frac{\partial u}{\partial n}|_{\partial\Omega}$ can be obtained by solving the problem (1) for u with Dirichlet data f when the metric g is known.

Let A denote the $Forward\ Operator$ with $A(g) = \Lambda_g$, and this operator is well-posed and A^{-1} denote the inversion operator. The general idea behind inverse problems is to invert this operator and study the properties of this inversion, see Figure (2). Here the unknown parameter is g while the data is Λ_g. We are interested to answer the question "What can we say about metric g by having enough measurements?" $\big($ or when we have $\Lambda_g = (f, \frac{\partial u}{\partial n}|_{\partial\Omega})\big)$.

When we come to an inverse problem, some questions arise. These questions are : By having the observed data measurement,

1. **Existence:** Does any unknown parameter exist that fulfills these observations?
2. **Uniqueness:** Can the unknown parameters be reconstructed uniquely by the observed data measurements?
3. **Stability:** How the errors in the data measurements will affect reconstruction of our interest parameter?
4. **Reconstruction:** How can we reconstruct the unknown parameter from the data measurements computationally?

The forward problem is to predict what happens to the solution of the problem over time or space by having all background knowledge of the problem including mathematical model and parameters in the partial differential equations, initial conditions and boundary conditions.

However, the inverse problem should recover the missing information that provides the solution when some background knowledge of the problem including mathematical model and some data observed at the boundary of our domain are given but all parameters in partial differential equation are not clarified. On the other side, one should remember that Inverse Problems are mostly ill-posed and non-linear so we may need to consider more conditions to be able to treat them in the desired way.

3 Introduction to the Theory of Ill-posed Problems

Theory of well and ill-posed problems is completely described in the literature, see for example [5, 8, 56, 58, 59]. Below we briefly present main definitions of the theory of well and ill-posed problems.

3.1 Well-posed and Ill-posed Problems

Well-posed problem: Based on idea introduced by Jacques Hadamard, a problem is well-posed if all following conditions are satisfied:

1. Solution to the problem exists,
2. The existed solution is unique,
3. Solution is continuous with respect to initial conditions and boundary conditions.

Ill-posed problem: If at least one of the conditions (1, 2, 3) is violated, the problem is not well-posed any more and is named as an ill-posed problem.

Most of inverse problems do not fulfill Hadamard's postulates of well-posedness so one can say an inverse problem is often (almost always) ill-posed. This is because of not existing a unique solution to the problem that continuously depends on its data. The results below are well known and we briefly present them.

Well-posed and Ill-posed problems in Operator Forms:

Let A be a mapping operator on a Hilbert space H_1 into a Hilbert space H_2 such that $A : H_1 \to H_2$. Consider a Hilbert space H_1 and let $O(z^*)$ denote a neighbourhood of an element $z^* \in H_1$ and $D(A)$ is the domain of definition and $R(A)$ is the range of A.

The problem $A(z) = u$ is well-posed on the pair of Hilbert spaces H_1 and H_2 if

1. **Existence:** $\forall u \in H_2, \exists z^* \in H_1$ s.t $A(z^*) = u, i.e. R(A) = H_2,$
2. **Uniqueness:** $if\big(A(z_1) = u \wedge A(z_2) = u\big) \to z_1 = z_2$ so the inverse operator $A^{-1} : H_2 \to H_1$ exists,
3. **Stability:** for any neighbourhood $O(z^*) \subset H_1$ of the solution $z^*$to the equation $Az = u$, $\exists$ neighbourhood $O(u) \subset H_2$ of the right-hand side u such that for all $u_\delta \in O(u)$ the element $A^{-1}u_\delta = z_\delta$ belongs to the neighbourhood $O(z_\delta)$ which shows continuity of the operator A^{-1} .

Consequently, the problem $A(z) = u$ is ill-posed on the pair of Hilbert spaces H_1 and H_2 if at least one of the three well-posedness conditions does not hold. Here z^* denotes the exact solution to the problem.

3.2 The Regularity Condition

All results below are due to [7, 8] and we briefly present them here. We consider the operator equation in the following form:

$$A(z) = 0. \tag{3}$$

Here $A : H_1 \to H_2$ is a non-linear operator which acts on a real Hilbert spaces (H_1, H_2). Consider z^* as the solution of our interest which we call the exact solution. Let $L(H_1, H_2)$ denote the normed space of all linear continuous operators from H_1 into H_2. Assume that A is defined and also is Fréchet differentiable everywhere in H_1, see section (4.2).

Thus two following assumptions and sometimes weaker conditions than (4), (5) are considered here on the Fréchet derivative A' for iterative processes:

$$\|A'(z_1)\|_{L(H_1,H_2)} \leq M_1, \qquad \forall z_1 \in H_1, \tag{4}$$

$$\|A'(z_1) - A'(z_2)\|_{L(H_1,H_2)} \leq M_2\|z_2 - z_1\|, \qquad \forall z_1, z_2 \in H_1. \tag{5}$$

On weaker conditions, one may replace $\forall z_1, z_2 \in H_1$ with $\forall z_1, z_2 \in \Omega$, where Ω is a bounded subset of H_1. As an example, we may take $\Omega = B(0, R)$, a ball of a sufficiently large radius. Generally,

$$B(z_1, r) = \{z_2 \in H_1 : \|z_2 - z_1\|_{H_1} \leq r\} \tag{6}$$

We can claim that if we use $z_1, z_2 \in \Omega$ instead of $\forall z_1, z_2 \in H_1$ in (4), (5), then condition (4) with

$$M_1 = \|A'(\bar{z})\|_{L(H_1,H_2)} + M_2 \sup_{\forall z_1 \in \Omega} \|z_1 - \bar{z}\|_{H_1}, \tag{7}$$

is a result of (5) for any $\bar{z} \in \Omega$. It follows from (5) that $A'(z)$ is dependent continuously on z which maps from H_1 to $L(H_1, H_2)$. For the operator which has a continuous derivative and for operators from the class $A(M_1, M_2)$ particularly, we have

$$A(z_1 + h) - A(z_1) = \int_0^1 A'(z_1 + th)h \ \ dt, \qquad\qquad z_1, h \in H_1, \tag{8}$$

as a result of the Newton-Leibniz theorem. The modified version of the Taylor expansion's formula reads as

$$A(z_1 + h) = A(z_1) + A'(z_1)h + T(z_1, h), \qquad\qquad \forall z_1, h \in H_1. \tag{9}$$

By our assumption on $T(z_1, h)$

$$\|T(z_1, h)\|_{H_2} \leq \frac{1}{2} M_2 \|h\|^2_{H_1}, \tag{10}$$

and it follows then

$$\|A(z_1 + h) - A(z_1)\|_{H_2} \leq M_1 \|h\|_{H_1}, \qquad\qquad \forall z_1, h \in H_1, \tag{11}$$

Linearization Approach: One can choose linearization approach among all other approaches for reconstruction in iterative methods for different equations with differentiable operators. We may take an arbitrary $z_0 \in H_1$. Then we rewrite (3) in the following way

$$A(z_0 + h) = 0, \tag{12}$$

where $h = z_1 - z_0$ is an unknown. Let us consider the Linearization procedure for the operator equation (3) at the point z_0. We neglect the term $T(z, h)$ in (9) in order to get the following approximation

$$A(z + h) \approx A(z) + A'(z)h. \tag{13}$$

Equation (12) becomes linear with respect to h

$$A(z_0) + A'(z_0)h = 0, \qquad\qquad h \in H_1. \tag{14}$$

We consider an approximation to the solution of equation (14), $\hat{h} \in H_1$, when there exists a solution to the linearized equation (14). One may consider the point $\hat{z} = z_0 + \hat{h}$ as approximation of initial solution, z^*, of equation (3). The linearization procedure is based on the discarding the small term $T(z, h)$ in the right hand side of (9) when $z = z_0$. Moreover, when we want to be sure about existing a unique solution for any $A(z_0) \in H_2$, it is necessary to have a continuous invertible operator $A'(z_0) \in H_2$ together with some estimate on the norm of the inverse operator $A'(z_0)^{-1}$.

Remark: The operator $A \in L(H_1, H_2)$ is *continuous invertible* if there exists an inverse operator A^{-1} such that $A^{-1} \in L(H_2, H_1)$.

If $A'(z_0)$ is not continuous and invertible then the operator $A'(z_0)^{-1}$ is either undefined, or its domain $D(A'(z_0)^{-1})$ does not coincide with the entire space H_2. This results in having unsolvable equation (14) or this equation may have infinitely many solutions. If $A'(z_0)^{-1} \notin L(H_2, H_1)$ then even for $A(z_0) \in D(A'(z_0)^{-1})$ the solution to (14) can not be dependent continuously on $A(z_0)$. So even small errors in the problem, which are not avoidable in computational process, will results in significant changes of the results, or may turn (14) into an unsolvable equation.

Regularity and Irregularity conditions: Assume $A'(z_0)$ as a continuous invertible,i.e.,

$$A'(z_0)^{-1} \in L(H_2, H_1). \tag{15}$$

Then A is called *regular* at the point z_0. The equation (3) is *regular in a neighbourhood* $O(z^*, \delta)$ *of the solution* z^*, if there exists a neighbourhood $O(z^*, \delta)$, such that for any $z \in O(z^*, \delta)$ the operator $A'(z^{-1})$ exists and belongs to $L(H_1, H_2)$. If not, equation (3) is called *irregular in a neighbourhood of* z^*. So the irregularity of (3) in a neighbourhood of z^* means that there exists an arbitrary point z, close to z^*, where either the operator $A'(z)^{-1}$ is not defined, or $A'(z) \notin L(H_2, H_1)$.

Linearization Procedure: Consider for some $m > 0$

$$\|A'(z_0)^{-1}\|_{L(H_2,H_1)} \leq m < \infty. \tag{16}$$

By considering (8), (10), (14), and (16), we can estimate the error $\|\hat{z} - z^*\|_{H_1}$ in terms of $\|z_0 - z^*\|_{H_1}$. Also by (14) we get

$$\hat{h} = -A'(z_0)^{-1}A(z_0), \tag{17}$$

then

$$\hat{z} = z_0 - A'(z_0)^{-1}A(z_0) \tag{18}$$

By subtracting z^* from (18) then using the fact that $A(z^*) = 0$, taking the norm of it we get:

$$\|\hat{z} - z^*\|_{H_1} = \|z_0 - z^* - A'(z_0)^{-1}\big(A(z_0) - A(z^*)\big)\|_{H_1}. \tag{19}$$

Then (9) and (10) provide,

$$A(z^*) - A(z_0) = A'(z_0)(z^* - z_0) + T(z_0, z^* - z_0), \tag{20}$$

where

$$\|T(z_0, z^* - z_0)\|_{H_2} \leq \frac{1}{2} M_2 \|z_0 - z^*\|^2_{H_1}.$$

By substituting (20) into (19) we get:

$$\begin{aligned} \|\hat{z} - z^*\|_{H_1} &= \|z_0 - z^* - A'(z_0)^{-1}\big[- A'(z_0)(z^* - z_0) - T(z_0, z^* - z_0)\big]\|_{H_1} \\ &\leq \|z_0 - z^* - A'(z_0)^{-1}[-A'(z_0)(z^* - z_0) - \frac{1}{2} M_2 (z^* - z_0)\big]\|_{H_1} \\ &= \|z_0 - z^* - m\big[- A'(z_0))(z^* - z_0) + \frac{1}{2} M_2 (z^* - z_0)\big]\|_{H_1}. \end{aligned} \tag{21}$$

Now we have:

$$\|\hat{z} - z^*\|_{H_1} \leq \frac{1}{2} m M_2 \|z_0 - z^*\|_{H_1}. \tag{22}$$

Now we can claim from (22) that $\hat{z}$ the solution of linearized equation (14), $\hat{z}$, is closer to the solution of original equation, z^*, than the initial element z_0, which shows that z_0 is not that far from z^*. If we repeat this linearization procedure at the point $z_1 = \hat{z}$, then we will get an iterative process.

4 Tikhonov's Scheme

In this section we will describe the Tikhonov's method for the solution of ill-posed problems. We present Tikhonov's scheme according to [7].

4.1 The Tikhonov Functional

Let us consider the operator equation:

$$A(z) = u. \tag{23}$$

We assume that the right hand side of (23) is given with the small parameter $\delta \in (0, 1)$ which characterizes the level of the error in data. Let u^* be the perfect noiseless right hand side of (23) which corresponds to the exact solution z^* such that

$$A\left(z^*\right) = u^*, \quad \|u - u^*\| \leq \delta. \tag{24}$$

Introduce the operator $F : H \to L_2$ such that

$$F(z) := Az - u. \tag{25}$$

Hence $Az^* - u^* = 0$. Since $\|u - u^*\| \leq \delta$ then

$$||F(z^*)|| \leq \delta. \tag{26}$$

The operator $\|F\|^2$ has the Frechét derivative at every point $\alpha \in H$. Although H is a finite dimensional space, where all norms are equivalent, we have used in our computations $L_2(\Omega)$ norm in the regularizing term of the Tikhonov functional.

Recall that the problem of solution of the operator equation (23) is a classical ill-posed problem [56] since the operator A^{-1} may not be compact. Thus, we assume that there exists the exact solution z^* to our inverse problem but we never will get this solution in computations. Because of that we call by the *regularized solution* z_γ some approximation of the unknown exact solution z^* which is satisfied to the requirements of closeness to the exact solution z^* and stability with respect to the small errors of the right-hand side u of equation (23).

We use Tikhonov regularization algorithm of [7, 56] which is based on the minimization of the Tikhonov functional. Thus, to find regularized solution z_γ of equation (23), we minimize the Tikhonov regularization functional $\Psi_\gamma(z)$

$$\Psi_\gamma\left(z\right) = \frac{1}{2}\left\|F(z)\right\|_{L_2}^2 + \frac{\gamma}{2}\left\|z - z_0\right\|_{L_2}^2, \tag{27}$$

$$\Psi_\gamma : H \to \mathbb{R}, \quad z_0 \in H,$$

where $\gamma = \gamma\left(\alpha(\delta)\right) := \gamma(\delta) > 0$ is a small regularization parameter. The choice of the point z_0 and the regularization parameter γ depends on the concrete minimization problem. Usually z_0 is a good first approximation for the exact solution z^*. To make it more clear, assume that (3) has a solution and let $\mathbf{X}^*(A)$ contain its solutions. then the set of minimizers of functional

$$\Psi(z) = \frac{1}{2}\|A(z)\|_{H_2}^2 \ \text{ for } z \in H_1, \tag{28}$$

coincides with $\mathbf{X}^*(A)$. It is clear that the The Tikhonov functional (27) does not necessarily have a global minimizer. However for any $\gamma > 0$ and for any $\alpha > 0$ there exists an element $z_\gamma^\alpha \in H_1$, such that $\Psi_\gamma(z_\gamma^\alpha)$ differs from the infimum by not more than α, i.e.,

$$\inf_{z \in H_1} \Psi_\gamma(z) \le \Psi_\gamma(z_\gamma^\alpha) \le \inf_{z \in H_1} \Psi_\gamma(z) + \alpha. \tag{29}$$

Suppose $\alpha = \alpha(\gamma)$ depends on the regularization parameter γ in such a way that

$$\lim_{\gamma \to 0} \frac{\alpha(\gamma)}{\gamma} = 0. \tag{30}$$

Take some $z^* \in \mathbf{X}^*(A)$. Recall, that by our assumption, $\mathbf{X}^*(A) \neq \emptyset$. By the definition of z_γ^α,

$$\Psi_\gamma(z_\gamma^\alpha) \le \Psi_\gamma(z^*) + \alpha(\gamma). \tag{31}$$

Therefore, by (27), (28) and (31) we have

$$\Psi(z_\gamma^\alpha) + \frac{\gamma}{2}\|z_\gamma^\alpha - z_0\|_{H_1}^2 \le \frac{\gamma}{2}\|z^* - z_0\|_{H_1}^2 + \alpha(\gamma) \tag{32}$$

From (32) follows that every part in the left hand side is less than right hand side :

$$\Psi(z_\gamma^\alpha) \leq \frac{\gamma}{2}\|z^* - z_0\|_{H_1}^2 + \alpha(\gamma), \tag{33}$$

$$\|z_\gamma^\alpha - z_0\|_{H_1}^2 \leq \|z^* - z_0\|_{H_1}^2 + \frac{2\alpha(\gamma)}{\gamma}. \tag{34}$$

For simplification the set of elements $\{z_\gamma^{\alpha(\gamma)}\}_{\gamma\in(0,\gamma_0]}(\gamma_0 > 0)$ is called a sequence from now. By (30)

$$\lim_{\gamma\to 0} \alpha(\gamma) = 0.$$

Hence, (33) indicates that the sequence $\{z_\gamma^{\alpha(\gamma)}\}$ is *minimizing* for the functional Ψ as γ tends to zero, i.e.,

$$\lim_{\gamma\to 0} \Psi(z_\gamma^{\alpha(\gamma)}) = \inf_{z\in H_1} \Psi(z) = 0.$$

4.2 The Fréchet Derivative and the convexity of the Tikhonov Functional

In our description of the iterative regularization we follow [7, 8].

We consider a general form of the Tikhonov functional (27) to present the convexity property of this functional. Let H_1,H_2, H be three Hilbert spaces, $H \subseteq H_1$ as a set, the norm in H is stronger than the norm in H_1 and $\overline{H} = H_1$, where the closure is understood in the norm of H_1. We denote scalar products and norms in these spaces as

$$(\cdot,\cdot)\,,\|\cdot\| \text{ for } H_1$$

$$(\cdot,\cdot)_2\,,\|\cdot\|_2 \text{ for } H_2$$

$$\text{and } [\cdot,\cdot]\,,[\cdot] \text{ for } H.$$

Let $A : H_1 \to W_2$ be a bounded linear operator. Our goal is to find the function $z \in H$ which minimizes the Tikhonov functional

$$\Psi_\gamma(z) : H \to \mathbb{R}, \tag{35}$$

$$\Psi_\gamma(z) = \frac{1}{2}\|Az - u\|_2^2 + \frac{\gamma}{2}[z - z_0]^2\,, u \in H_2; z, z_0 \in H, \tag{36}$$

where $\gamma \in (0, 1)$ is the regularization parameter.

To do that we search for a stationary point of the above functional with respect to z satisfying $\forall b \in H$

$$\Psi'_\gamma(z)(b) = 0. \tag{37}$$

where $\Psi'_\gamma(z)$ is the Fréchet derivative of the functional (36) and $\Psi'_\gamma(z)(b)$ means that Ψ'_γ acts on b.

The following lemma is well known [7] and we present it for the case when the operator $A : L_2 \to L_2$. In the case when $A : H^1 \to L_2$ the explicit derivation of the Fréchet derivative of the functional (36) is technically more difficult because of the presence of H^1 norm in the regularization term of the functional (36), and we omit here explicit presentation of the Fréchet derivative of the functional (36) in this case.

Lemma 1. *Let $A : L_2 \to L_2$ be a bounded linear operator. Then the Fréchet derivative of the functional (35) is*

$$\Psi'_\gamma\left(z\right)\left(b\right) = \left(A^*Az - A^*u, b\right) + \gamma\left[z - z_0, b\right], \forall b \in H. \tag{38}$$

Lemma 2 is also well known since $A : H_1 \to H_2$ is a bounded linear operator. We again formulate this lemma only for our specific case, referring to [58] for a more general case. The situation is naturally more complicated for a nonlinear operator, and we refer to [6] for this case.

Lemma 2. *Let the operator $A : H_1 \to H_2$ be a bounded linear operator which has the Fréchet derivative of the functional (35). Then the functional $\Psi_\gamma\left(z\right)$ is strongly convex on the space H and*

$$\left(\Psi'_\gamma\left(z_1\right) - \Psi'_\gamma\left(z_2\right), z_1 - z_2\right) \geq \gamma[z_1 - z_2]^2, \forall z_1, z_2 \in H.$$

It is known from the theory of convex optimization that Lemma 2 implies existence and uniqueness of the global minimizer $z_\gamma \in H$ of the functional Ψ_α defined in (35) such that

$$\Psi_\gamma(z_\gamma) = \inf_{z \in H} \Psi_\gamma(z).$$

It is well known that the operator F is Lipschitz continuous if

$$\|F(z_1) - F(z_2)\| \leq ||A|| \cdot \|z_1 - z_2\| \;\; \forall z_1, z_2 \in H.$$

We also have introduced new constant $M_2 = M_2\left(||A||, \gamma\right) = const. > 0$ [7] such that

$$\left\|\Psi'_\gamma\left(z_1\right) - \Psi'_\gamma\left(z_2\right)\right\| \leq M_2\left\|z_1 - z_2\right\|, \forall z_1, z_2 \in H. \tag{39}$$

5 Optimisation Methods

To find minimum of the Tikhonov functional we can use any gradient-like method, for example, gradient method (GM) or conjugate gradient method (CGM). We briefly present usual gradient method and it's modification, CGM, which improves speed of computations. We refer to [7, 8, 31, 56] for analysis of these methods.

Let (z_1, z_2) denotes scalar product in the Euclidean space E. Let z^* be the perfect noiseless solution such that $Az^* = u^*$, where A is the matrix in (3). From (3) also follows that $u^* = 0$. Let $\tilde{z}$ be measured permittivity which corresponds to measured functions k, β such that $A\tilde{z} = \tilde{u}$.

We assume that the level of the noise δ in data is such that

$$||\tilde{u} - u^*|| \leq \delta. \tag{40}$$

Since $u^* = 0$ then the equation above means that

$$||\tilde{u}|| \leq \delta. \tag{41}$$

Here we present briefly the gradient method and conjugate gradient method to minimize the quadratic functional

$$f(z) = \frac{1}{2}(Az, z) - (\tilde{u}, z), \tag{42}$$

where A is the positive definite matrix. It is easy to see that $f'(z) = Az - \tilde{u}, f''(z) = A$. When A is positive definite then $\lambda_{min}(A) > 0$ and quadratic functional is strictly convex. If $\lambda_{min}(A) = 0$ the functional is convex, but not strictly convex and z_0 is an initial approximation to z.

5.1 Gradient Method (GM)

The minimizing sequence in a gradient method is constructed via the following formulas:

$$\begin{aligned} h_k &= -f'(z_k) \\ z_{k+1} &= z_k + \alpha_k h_k, \\ \alpha_k &= argmin\{f(z_k + \alpha h_k) : \alpha \geq 0\}. \end{aligned} \tag{43}$$

5.2 Conjugate Gradient Method (CGM)

The minimizing sequence in a conjugate gradient method is constructed via the following formulas:

$$\begin{aligned} h_0 &= -f'(z_0), \\ z_{k+1} &= z_k + \alpha_k h_k, \\ \alpha_k &= argmin\{f(z_k + \alpha h_k) : \alpha \geq 0\}, \\ h_k &= -f'(z_k) + \beta_{k-1} h_{k-1}. \end{aligned} \tag{44}$$

In the case when β_{k-1} is determined as

$$\beta_{k-1} = \begin{cases} \frac{||f'(z_k)||^2}{||f'(z_{k-1})||^2} & \text{if } k \neq 1, n+1, 2n+1, ..., \\ 0 & \text{otherwise,} \end{cases} \tag{45}$$

then the method is called Fletcher-Reeves method.

If β_{k-1} is determined as

$$\beta_{k-1} = \begin{cases} \frac{(f'(z_k), f'(z_k) - f'(z_{k-1}))}{||f'(z_{k-1})||^2} & \text{if} k \neq 1, n+1, 2n+1, ..., \\ 0 & \text{otherwise,} \end{cases} \tag{46}$$

then the method is called Polack-Ribiere method.

6 Principle of Iterative Regularization

In our description of the iterative regularization we follow [7, 8].

6.1 The idea of iterative regularization

Let us consider an arbitrary sequence of the regularization parameters $\{\gamma\}$ such that $\gamma_n > 0$ and $\gamma_n \to 0$ as $n \to 0$. We are able to find an approximate minimum of a strongly convex functional Ψ_{γ_n} with an accuracy $\alpha_n = \alpha(\gamma_n)$ where α_n is a positive number. We may take some finite number $M_{\gamma_n}$of steps of the chosen method. We consider the point, $z_{\gamma_n}^{\alpha_n}$, gained by these assumptions. Using (33) and (34) for this $z_{\gamma_n}^{\alpha_n}$ we get :

$$\Psi(z_{\gamma_n}^{\alpha_n}) \leq \frac{\gamma_n}{2}\|z^* - z_0\|_{H_1}^2 + \alpha_n,\ \Psi(z) = \frac{1}{2}\|A(z)\|_{H_2}^2, \tag{47}$$

$$\|z_{\gamma_n}^{\alpha_n} - z_0\|_{H_1}^2 \leq \|z^* - z_0\|_{H_1}^2 + \frac{2\alpha_n}{\gamma_n}. \tag{48}$$

Here assume that

$$\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \gamma_n = \lim_{n\to} \frac{\alpha_n}{\gamma_n} = 0. \tag{49}$$

We assume that the sequence $\{z_{\gamma_n}^{\alpha_n}\}$ converges weakly in H_1 to some point z^*. Since a continuous convex functional Ψ is weakly lower semi-continuous, therefore

$$\Psi(z^*) \leq \liminf_{n\to\infty} \Psi(z_{\gamma_n}^{\alpha_n}).$$

(47) and (49) result in $\Psi(z^*) = 0$, i.e., $z^* \in \mathbf{X}^*(A)$. Considering arguments of section 4 and using (48) and (49), we get

$$\lim_{n\to\infty} \|z_{\gamma_n}^{\alpha_n} - z_{z_0}^*\|_{H_1} = 0. \tag{50}$$

where $z_{z_0}^*$ denotes the nearest point to z_0 in $\mathbf{X}^*$. So the sequence $\{z_{\gamma_n}^{\alpha_n}\}$ minimizes the discrepancy functional that converges strongly to $z_{z_0}^* \in \mathbf{X}^*(A)$ when we have convexity and closeness property for $\mathbf{X}^*(A)$. Also if we assume that $\mathbf{X}^*(A) \neq \emptyset$, we saw that M_{γ_n} is rapidly growing as $\gamma_n, \alpha_n \to 0$. Practically, the mentioned approach is not that good because of being done in two steps and also due to being uncertain about choice of M_{γ_n}. We want here to expand a special one-step iterative process which guarantees convergence of the generated sequence z_n to the $z_{z_0}^*$. How we do reconstruction of such processes is coming below. We consider a basic scheme for minimizing the strongly convex functional. Besides, we take an arbitrary sequence such that

$$0 < \gamma_{n+1} \leq \gamma_n,\ n = 0, 1, ...,\ \lim_{n\to\infty} \gamma_n = 0. \tag{51}$$

We suppose the point z_n has already been constructed. Then the next point could be obtained by fixing a number of steps (independent of n) of the basic iterative method applied to the minimization problem for the Tikhonov functional Ψ_{γ_n} with the initial point z_n. After doing this, by repeating the process with $\gamma = \gamma_{n+1}$ and with initial point z_{n+1}, etc. This algorithm for the reconstruction of the convergent minimizing sequences for the discrepancy Ψ is called the *iterative regularization principle* [9]. To do such a implementation, we need to

1 Choose a basic method, acceptable for finding a global minimizer of an arbitrary strongly convex functional, and to fix a number of iterations that one has to perform at every step of the main iterative process; in most cases, one iteration is performed;

2 Choose a sequence of the regularization parameters$\{\gamma_n\}$, satisfying conditions (51);

3 Adjust the inner parameters of the basic method (step size, etc) and the regularization parameter $\gamma_n, n = 0, 1, ...$

6.2 An illustrative example of the equations with a convex discrepancy functional

Consider a gradient method with a constant step applied to a functional Ψ_γ for a fixed regularization parameter $\gamma > 0$, as an example. This method takes the form

$$z_{n+1} = z_n - \alpha \Psi'_\gamma(z_n),\ n = 0, 1, ... \text{ and } z_0 \in H_1. \tag{52}$$

According to (38) we have the following necessary condition for minimum:

$$\Psi'_\gamma(z_\gamma) := \Psi'(z_\gamma) + \gamma(z_\gamma - z_0) = 0. \tag{53}$$

By subtracting global minimizer z_γ from both side of (52), taking norms, then taking into account equations (53) and (48) we get :

$$\begin{cases} \|z_{n+1} - z_\gamma\|^2_{H_1} = \|z_n - z_\gamma - \alpha\big(\Psi'_\gamma(z_n) - \Psi'_\gamma(z_\gamma)\big)\|^2_{H_1} \\ \qquad\qquad\qquad = \|z_n - z_\gamma\|^2_{H_1} - 2\alpha\big((\Psi'_\gamma(z_n) - \Psi'_\gamma(z_\gamma), z_n - z_\gamma\big)_{H_1} + \alpha^2 \|\Psi'_\gamma(z_n) - \Psi'_\gamma(z_\gamma)\big)\|^2_{H_1}. \end{cases} \tag{54}$$

In (54) we have used the fact that $\Psi'_\gamma(z_\gamma) = 0$.

Recalling that

$$\Psi'(z) = A'^*(z)A(z), \quad z \in H_1. \tag{55}$$

Taking into account (53) and (55), then using (4) and (5), adding and subtracting the term $\|A'^*(z_n)\| \cdot \|A(z_\gamma)\|$ we have :

$$
\begin{aligned}
\|\Psi'_\gamma(z_n) - \Psi'_\gamma(z_\gamma)\big)\|_{H_1} &\le \|\Psi'(z_n) - \Psi'(z_\gamma\|_{H_1} + \gamma\|z_n - z_\gamma\|_{H_1} \\
&\le \|A'^*(z_n)\|_{L(H_2,H_1)}\|A(z_n) - A(z_\gamma)\|_{H_2} \\
&\quad + \|A'^*(z_n) - A'^*(z_\gamma)\|_{L(H_2,H_1)}\|A(z_\gamma)\|_{H_2} \\
&\quad + \gamma\|z_n - z_\gamma\|_{H_1} \\
&\le \big(M_1^2 + M_2\|A(z_\gamma)\|_{H_2} + \gamma\big)\|z_n - z_\gamma\|_{H_1}.
\end{aligned}
\tag{56}
$$

By definition of $\Psi(z_\gamma)$ in (36), then using (29) and (32) we obtain :

$$
\Psi_\gamma(z_\gamma) = \frac{1}{2}\|A(z_\gamma)\|_{H_2}^2 + \frac{\gamma}{2}\|z_\gamma - z_0\|_{H_1}^2 \le \Psi(z_{z_0}^*) = \frac{\gamma}{2}\|z_{z_0}^* - z_0\|_{H_1}^2.
$$

Now taking square root of the above inequality we get,

$$
\|A(z_\gamma)\|_{H_2} \le \sqrt{\gamma}\|z_{z_0}^* - z_0\|_{H_1}. \tag{57}
$$

By substituting (57) into (56) we have

$$
\|\Psi'_\gamma(z_n) - \Psi'_\gamma(z_\gamma)\|_{H_1} \le L\|z_n - z_\gamma\|_{H_1}, \text{where } L = M_1^2 + M_2\|z_{z_0}^* - z_0\|_{H_1}\sqrt{\gamma} + \gamma. \tag{58}
$$

We substitute (56) and (58) into (54) and using the fact that the functional Ψ_γ is strongly convex in the space H_1 for any $\gamma > 0$ which means

$$
\big(\Psi'(z_1) - \Psi'(z_2)\big)_{H_1} \ge 0 \ \ \forall z_1, z_2 \in H_1.
$$

Thus,

$$
\big(\Psi'(z_1) - \Psi'(z_2), z_1 - z_2\big)_{H_1} \ge \|z_1 - z_2\|_{H_1}^2 \ \ \forall z_1, z_2 \in H_1 \tag{59}
$$

we get

$$
\|z_{n+1} - z_\gamma\|_{H_1} \le \sqrt{1 - 2\alpha\gamma + \alpha^2 L^2}\|z_n - z_\gamma\|_{H_1}, \quad n = 0, 1, ... \tag{60}
$$

Clearly, if α is an arbitrary positive number then the expression under the radical in (60) is positive. The step size $\alpha > 0$ must be chosen to make this small expression as small as possible. To have convergence in (60) we should have

$$
1 - 2\alpha\gamma + \alpha^2 L^2 < 1.
$$

We need to have following assumption to get the above inequality

$$
0 < \alpha < \frac{2\gamma}{L^2}. \tag{61}
$$

If (61) is satisfied, then estimate (60) yields

$$\|z_n - z_\gamma\|_{H_1} \leq \|z_0 - z_\gamma\|_{H_1} q^n(\gamma), \quad n = 0, 1, ...,$$

$$q(\gamma) = \sqrt{1 - 2\alpha\gamma + \alpha^2 L^2}, \;\; q(\gamma) \in (0, 1). \tag{62}$$

Based on (62), the sequence $\{z_n\}$, defined in (52), converges to z_γ linearly. The quadratic polynomial under the radical in (60) takes its minimum with respect to α when $\alpha = \gamma L^{-2}$. The corresponding minimal value of the polynomial is $1 - \gamma^2 L^{-2}$. So if $\gamma \to 0$, then $q(\gamma)$ in (62) converges to 1 for any γ satisfying condition (61). So the regularization parameter γ is decreasing, the number of iterations in (52) must grow in order to insure the uniform with respect to γ accuracy of the approximation of z_γ by the elements z_n. If we choose

$$\alpha = \frac{\epsilon\gamma}{L^2}, \epsilon \in (0, 2),$$

then the expression under the radical in (60) takes the form

$$1 - 2\alpha\gamma + \alpha^2 L^2 = 1 - C(\epsilon)\gamma^2, \;\; C(\epsilon) = \frac{\epsilon(2 - L^2\epsilon)}{L^2}.$$

Thus, estimate (60) can be written as

$$\|z_{n+1} - z_\gamma\|_{H_1} \leq \sqrt{1 - C(\epsilon)\gamma^2} \| \, z_n - z_\gamma\|_{H_1}.$$

6.3 The iteratively regularized gradient method

In this section we present a general scheme for the analysis of the convergence of iteratively regularized algorithms. We show how we use gradient descent method with a constant step as the basic method to obtain the *iteratively regularized gradient method.*

The iteratively regularized gradient method is

$$\begin{cases} z_{n+1} = z_n - \alpha_n \Psi(\gamma_n)(z_n) & \text{where } z_0 \in H_1, \\ \qquad\;\; = z_n - \alpha_n\big(A'^*(z_n)A\big(z_n\big) + \gamma_n(z_n - z_0)\big), & \end{cases} \tag{63}$$

where $\alpha_n > 0$ shows the step size, which is an inner parameter of the gradient method, has to be chosen as a function of the regularization parameter γ_n. Consider an arbitrary sequence, $\{\gamma_n\}$ which has properties of (51). Moreover,

$$\lim_{n \to \infty} \|z_{\gamma_n} - z^*_{z_0}\|_{H_1} = 0. \tag{64}$$

By considering (64) and

$$\lim_{n \to \infty} \|z_n - z_{\gamma_n}\|_{H_1} = 0. \tag{65}$$

We can say the sequence $\{z_n\}$ converges to $z^*_{z_0}$ as $n \to \infty$.
By considering

$$\eta_n = \|z_n - z_{\gamma_n}\|_{H_1},$$

for the value of the η_n a certain difference inequality is fulfilled. For obtaining this inequality, we need to have an estimate on the norm of the difference between the elements z_{γ_1} and z_{γ_2}, the global minimizes of Tikhonov's functional Ψ_{γ_1} and Ψ_{γ_2} with two different regularization parameters γ_1 and $\gamma_2 > 0$. These elements satisfy the operator equations

$$\Psi'_{\gamma_1}(z_{\gamma_1}) = 0, \;\; \Psi'_{\gamma_2}(z_{\gamma_2}) = 0. \tag{66}$$

Here we explain in more details. Consider

$$\Psi'(z_{\gamma_1}) + \gamma_1(z_{\gamma_1} - z_0) = 0, \tag{67}$$

$$\Psi'(z_{\gamma_2}) + \gamma_2(z_{\gamma_2} - z_0) = 0. \tag{68}$$

Subtracting last two identities from each other, we obtain

$$(\gamma_1 - \gamma_2)(z_{\gamma_2} - z_0) = \Psi'(z_{\gamma_2}) - \Psi'(z_{\gamma_1}) + \gamma_1(z_{\gamma_2} - z_{\gamma_1}).$$

If we multiply both sides of this equality by $(z_{\gamma_2} - z_{\gamma_1})$ and then by considering the convexity of the functional Ψ, we have

$$\big(\Psi'(z_\beta) - \Psi'(z_{\gamma_1}), z_\beta - z_{\gamma_1}\big)_{H_1} \geq 0,$$

then we achieve

$$(\gamma_1 - \gamma_2)(z_{\gamma_2} - z_0, z_{\gamma_2} - z_{\gamma_1})_{H_1} \geq \gamma_1 \|z_{\gamma_2} - z_{\gamma_1}\|^2_{H_1}. \tag{69}$$

By considering the Cauchy-Schwartz inequality for the left hand side of the (69) and dividing the inequality (69) by $\|z_{\gamma_2} - z_{\gamma_1}\|_{H_1}$, we obtain

$$\|z_{\gamma_2} - z_{\gamma_1}\|_{H_1} \leq \frac{|\gamma_1 - \gamma_2|}{\gamma_1}\|z_{\gamma_2} - z_0\|_{H_1} \quad \forall \gamma_1, \gamma_2 > 0. \tag{70}$$

Let $\alpha(\gamma)$ be zero in (34). Then

$$\|z_{\gamma_2} - z_0\|_{H_1} \leq \|z^*_{z_0} - z_0\|_{H_1} \forall \gamma_1 > 0.$$

So (70) gives

$$\|z_{\gamma_2} - z_{\gamma_1}\|_{H_1} \leq \frac{|\gamma_1 - \gamma_2\|}{\gamma_1}\|z^*_{z_0} - z_0\|_{H_1} \quad \forall \gamma_1, \gamma_2 > 0. \tag{71}$$

Subtracting $z_{\gamma_{n+1}}$ from both sides of (63) we get :

$$z_{n+1} - z_{\gamma_{n+1}} = \big(z_n - z_{\gamma_n} - \alpha_n \Psi'_{\gamma_n}(z_n)\big) + (z_{\gamma_n} - z_{\gamma_{n+1}}) = I_1 + I_2. \tag{72}$$

To estimate I_1 we take norm of it and apply (54) with $z_\gamma := z_{\gamma_n}$ and then use (60) to get following estimate:

$$\|z_n - z_{\gamma_n} - \alpha_n \Psi'_{\gamma_n}(z_n)\|_{H_1} \leq \sqrt{1 - 2\alpha_n\gamma_n + \alpha_n^2 L^2}.\|z_n - z_{\gamma_n}\|, \tag{73}$$

To estimate I_2 we apply (71) to get:

$$\|z_{\gamma_n} - z_{\gamma_{n+1}}\|_{H_1} \leq \frac{\gamma_n - \gamma_{n+1}}{\gamma_n}\|z^*_{z_0} - z_0\|. \tag{74}$$

After we combine both estimates for I_1 and I_2 to obtain:

$$\eta_{n+1} = \|z_{n+1} - z_{\gamma_{n+1}}\|_{H_1} \leq \|z_n - z_{\gamma_n} - \alpha_n \Psi'_{\gamma_n}(z_n)\|_{H_1} + \|z_{\gamma_n} - z_{\gamma_{n+1}}\|_{H_1}, \tag{75}$$

$$\leq \sqrt{1 - 2\alpha_n\gamma_n + \alpha_n^2 L^2}\ \eta_n + \frac{\gamma_n - \gamma_{n+1}}{\gamma_n}\|z^*_{z_0} - z_0\|_{H_1}, \tag{76}$$

Here based on (58) and (51) we have,

$$L = M_1^2 + M_2\|z^*_{z_0} - z_0\|_{H_1}\sqrt{\gamma_0} + \gamma_0.$$

Any arbitrary sequence of positive step size $\{\alpha_n\}$ guarantees existence of the inequality (76) .
We can be sure that $\eta_n \to 0$, if we guarantee that the coefficient with η_n in (76) is less than 1 for sufficiently large enough n. To doing so we need to have

$$\lim \frac{\alpha_n}{\gamma_n} = 0. \tag{77}$$

We get following estimate

$$\sqrt{1 - 2\alpha_n\gamma_n + \alpha_n^2 L^2} \leq 1 - 2\alpha_n\gamma_n$$

since

$$\sqrt{1 - 2\alpha_n\gamma_n + \alpha_n^2 L^2} = \sqrt{1 - \alpha_n\gamma_n\big(2 - \frac{\alpha_n}{\gamma_n}L^2\big)} \tag{78}$$

and also with help of (77) for enough large $n \in \mathbf{N}$. So if we take a large number n, for the sequence$\{\eta_n\}$ the inequality

$$\eta_{n+1} \leq \big(1 - 2\alpha_n\gamma_n\big)\eta_n + \frac{\gamma_n - \gamma_{n+1}}{\gamma_n}\|z^*_{z_0} - z_0\|_{H_1} \tag{79}$$

is satisfied. Moreover, we can say that (51) and (77) provide

$$\lim_{n\to\infty} \alpha_n = 0. \tag{80}$$

Now we need following Lemma.

Lemma. [59] *Let the sequence of non-negative numbers* $\{\omega_n\}$ *be such that*

$$\omega_{n+1} \leq (1 - a_n)\omega_n + b_n, \qquad n = 0, 1, ..., \tag{81}$$

where $\{a_n\}, \{b_n\}$ *satisfy the assumptions*

$$0 < a_n \leq 1, \quad b_n \geq 0, \quad n = 0, 1, ...,$$

$$\sum\nolimits_{n=0}^{\infty} a_n = \infty, \qquad \lim_{n\to\infty} \frac{b_n}{a_n} = 0. \tag{82}$$

Then

$$\lim_{n\to\infty} \omega_n = 0. \tag{83}$$

We can apply the Lemma above to (79). Then for the case of (79) we have should choose a_n and b_n such that

$$a_n := \gamma_n \alpha_n, \quad b_n := \frac{\gamma_n - \gamma_{n+1}}{\gamma_n}. \tag{84}$$

We should have that (82) is fulfilled, or following properties should be satisfied.

$$\sum\nolimits_{n=0}^{\infty} \gamma_n \alpha_n = \infty, \quad \lim_{n\to\infty} \frac{\gamma_n - \gamma_{n+1}}{\gamma_n^2 \alpha_n} = 0. \tag{85}$$

We may take sequences $\{\gamma_n\}$ and $\{\alpha_n\}$ such that the conditions (51), (77) and (85) are fulfilled. In this case these sequences should slowly converge to zero. As an example, we may take

$$\gamma_n = \frac{\gamma_0}{(n+1)^p}, \; \alpha_n = \frac{\alpha_0}{(n+1)^{p+q}}, \quad n = 0, 1, ..., \tag{86}$$

here $\alpha_0, \gamma_0 > 0, \; p \in (0,1)$ and $q > 0$. Therefore assumptions (51) and (77) are satisfied. The first condition of (85) holds if

$$2p + q \leq 1. \tag{87}$$

Obviously by having inequality (87), the second condition in (85) holds. Besides, the following convergence rate guarantees for the sequence $\{\eta_n\}$ in this case:

$$\eta_n = \mathbf{O}\big(\frac{1}{n^{1-2p-q}}\big). \tag{88}$$

Now we can claim that by considering (51), (77) and (85), process (63) converges $z^*_{z_0}$, i.e., the identity

$$\lim_{n\to\infty} \|z_n - z^*_{z_0}\| = 0$$

holds, for any choice of the initial approximation $z_0 \in H_1$.

7 Methods for the Choice of Regularization Parameters in the Tikhonov Functional

In section 6.3 we described in details that we should choose iterative regularizations parameters, γ_1, γ_2, in the Tikhonov functional (138) such that conditions (51), (77) and (85) are fulfilled with $\gamma := \gamma_1$ or $\gamma := \gamma_2$. As an example, we may take sequences such that

$$\gamma_n = \frac{\gamma_0}{(n+1)^p}, \; \alpha_n = \frac{\alpha_0}{(n+1)^{p+q}}, \quad n = 0, 1, ..., \tag{89}$$

here $\alpha_0, \gamma_0 > 0, p \in (0,1)$ and $q > 0$. So one method for choosing iterative regularization parameters, γ_1, γ_2, in the Tikhonov functional is sequences such that (86) holds.

Regularization independent of the noise level (Heuristic): quasi-optimality [3]

The quasi-optimality criterion chooses the regularization parameter in inverse problems without taking into account the noise level. This rule works remarkably well in practice, although in [2] is shown that there are always counter examples with very poor performance.

The classical quasi-optimality criterion is applied to the regularized solutions, z_γ, based on Tikhonov's method and proposes to choose positive regularization parameters, γ_1, γ_2, with $\gamma := \gamma_1$ or $\gamma := \gamma_2$ in the Tikhonov functional such that

$$\min_{\gamma_n} \|\gamma_n \frac{dz_{\gamma_n}}{d\gamma_n}\| \to 0. \tag{90}$$

Using reparametrization with $\gamma_n = Cq^n$ for $q \in (0,1), n \in \mathbf{R}$, the condition (90) is fulfilled. In numerical tests we use the following second rule for iterative computation of the regularization parameters γ_1, γ_2:

$$\gamma_n := Cq^n, n = 1, 2, ..., \tag{91}$$

where n is the number of iteration in conjugate gradient method. Then condition (90) reduces to the following criterion:

$$\min_{n} \|z_{Cq^n} - z_{Cq^{n+1}}\| \to 0. \tag{92}$$

8 System of Maxwell's Equations

We consider the Maxwell's equations in an inhomogeneous isotropic medium in Ω with a piecewise smooth boundary $\partial\Omega$, where the bounded domain $\Omega \subset \mathbb{R}^d, \; d = 2, 3$. We define $\Omega_T := \Omega \times (0,\, T), \partial\Omega_T := \partial\Omega \times (0,\, T), T > 0$. Electromagnetic equations then are :

$$\begin{cases} \partial_t D - \nabla \times H(\mathbf{x},\, t) = 0, & (\mathbf{x},\, t) \in \Omega \times (0,\, T), \\ \partial_t B + \nabla \times E(\mathbf{x},\, t) = 0, & (\mathbf{x},\, t) \in \Omega \times (0,\, T), \\ D(\mathbf{x}, t) = \varepsilon E(\mathbf{x},\, t), \quad B(\mathbf{x},\, t) = \mu H(\mathbf{x},\, t), & \\ E(\mathbf{x},\, 0) = E_0(\mathbf{x}), \quad H(\mathbf{x},\, 0) = H_0(\mathbf{x}), & \\ \nabla \cdot D(\mathbf{x},\, t) = 0, \quad \nabla \cdot B(\mathbf{x},\, t) = 0, & (\mathbf{x},\, t) \in \partial\Omega \times (0,\, T), \\ n \times D(\mathbf{x},\, t) = 0, \quad n \times B(\mathbf{x},\, t) = 0, & (\mathbf{x},\, t) \in \partial\Omega \times (0,\, T). \end{cases} \tag{93}$$

Here $E(\mathbf{x},\, t)$ and $H(\mathbf{x},\, t)$ are the three-dimensional electric and magnetic fields of the time t and the space variable $\mathbf{x} = (x_1, x_2, x_3)$, while $D\,(\mathbf{x})$ and $B\,(\mathbf{x})$ are the electric and magnetic inductions, respectively.

The dielectric permittivity and magnetic permeability functions, $\varepsilon\,(\mathbf{x})$ and $\mu\,(\mathbf{x})$, depend only on $\mathbf{x} \in \Omega$ and $n = n(\mathbf{x})$ presents the outward normal vector on $\partial\Omega$.

Our goal is to reconstruct the coefficients $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ simultaneously using only backscattered electric field in the system (93). In (93) functions $\varepsilon(\mathbf{x}) = \varepsilon_{\mathrm{r}}(\mathbf{x})\varepsilon_0$, $\mu(\mathbf{x}) = \mu_{\mathrm{r}}(\mathbf{x})\mu_0$ and $\varepsilon_0 \approx 8.85 \times 10^{-12}(F/m), \mu_0 \approx 4\pi 10^{-7}(H/m)$ are the electric permittivity and magnetic permeability of vacuum and ε_r and μ_r are dimensionless relative electric permittivity and relative magnetic permeability, respectively. Since the units of both parameters, $\varepsilon(\mathbf{x})$ and ε_0 are the same and equals to Farad/meter, it results in dimensionless form of $\varepsilon_{\mathrm{r}}(\mathbf{x})$. In the same way, $\mu_{\mathrm{r}}(\mathbf{x})$ is dimensionless. We consider only a finite number of observations of the electric field on the boundary $\partial\Omega$ of the bounded domain Ω.

For solving the problem (93) numerically we apply the domain decomposition FEM/FDM method of [13]. In doing so we decompose Ω into two subregions, such that $\Omega = \Omega_{\mathrm{FEM}} \cup \Omega_{\mathrm{FDM}}$, $\Omega_{\mathrm{FEM}} \cap \Omega_{\mathrm{FDM}} = \emptyset$, see Figure 4 for example of these domains. We apply a finite element method and a finite difference method to the sub-domains, Ω_{FEM} and Ω_{FDM} , respectively, with first order absorbing boundary conditions, see Figure 3. The boundary $\partial\Omega$ is such that $\partial\Omega = \partial_1\Omega \cup \partial_2\Omega \cup \partial_3\Omega$ where $\partial_1\Omega$ and $\partial_2\Omega$ denote the front and back sides of the domain Ω, respectively, and $\partial_3\Omega$ is the union of the left, right, top and bottom sides of the domain.

We assume that functions ε and μ of equation (93) are such that

$$
\begin{aligned}
&\varepsilon(\mathbf{x}) \geq 1, \mu(\mathbf{x}) \geq 1 \text{ for } \mathbf{x} \in \Omega_{\mathrm{FEM}}, \\
&\varepsilon(\mathbf{x}) = 1, \mu(\mathbf{x}) = 1 \text{ for } \mathbf{x} \in \Omega_{\mathrm{FDM}}.
\end{aligned}
$$

We also assume that functions ε and μ should be determined only inside the finite element domain and are known inside of finite difference domain such that,

$$
\varepsilon \in C^2\left(\mathbb{R}^3\right), \quad \varepsilon(\mathbf{x}) \in [1,\, b_1] \text{ for } \mathbf{x} \in \mathbb{R}^3, \quad \varepsilon(\mathbf{x}) = 1 \text{ for } \mathbf{x} \in \mathbb{R}^3 \setminus \Omega_{FEM}, \tag{94}
$$

$$
\mu \in C^2\left(\mathbb{R}^3\right), \quad \mu(\mathbf{x}) \in [1,\, b_2] \text{ for } \mathbf{x} \in \mathbb{R}^3, \quad \mu(\mathbf{x}) = 1 \text{ for } \mathbf{x} \in \mathbb{R}^3 \setminus \Omega_{FEM}. \tag{95}
$$

Here, b_1 and b_2 are constants strictly larger than 1 and are chosen experimentally in the similar way with [12, 39, 57] and we have priori knowledge about them.

By eliminating B and D from (93) we obtain the model problem for the electric field E only with the perfectly conducting boundary conditions at the boundary $\partial\Omega$ such that

$$
\begin{aligned}
&\varepsilon(\mathbf{x})\frac{\partial^2 E(\mathbf{x},\, t)}{\partial t^2} + \nabla \times \left(\mu^{-1}(\mathbf{x})\nabla \times E(\mathbf{x},\, t)\right) = 0, && (\mathbf{x},\, t) \in \Omega \times (0,\, T), && (96)\\
&\nabla \cdot \left(\varepsilon(\mathbf{x})E(\mathbf{x})\right) = 0, && (\mathbf{x},\, t) \in \Omega \times (0,\, T), && (97)\\
&E(\mathbf{x},\, 0) = f_1(\mathbf{x}), \quad E_t(\mathbf{x},\, 0) = f_2(\mathbf{x}), && \mathbf{x} \in \Omega, && (98)\\
&n \times E = 0 && (\mathbf{x},\, t) \in \partial\Omega \times (0,\, T). && (99)
\end{aligned}
$$

The similar equation can be obtained also for H. Here we assume that

$$f_1 \in H^1(\Omega), f_2 \in L_2(\Omega).$$

For solving the equations (96)–(99) numerically, we use the finite difference method on a structured mesh with constant coefficients $\varepsilon = \hat{\varepsilon} = 1$ and $\mu = \hat{\mu} = 1$ in finite difference domain. In our computations all materials with values of $\varepsilon > 10$ are treated as metals and we call ε as "appearing dielectric constant", see [12, 57] for more information and explanation.

9 The Model Problem

For stabilization of the finite element solution using standard piecewise continuous functions, we enforce the divergence condition and add a Coulomb-type gauge condition [1, 45, 48] to (96) with $0 \leq s \leq 1$:

$$\begin{aligned}
&\varepsilon(\mathbf{x}) \frac{\partial^2 E(\mathbf{x},\, t)}{\partial t^2} + \nabla \times \left(\mu^{-1}(\mathbf{x}) \nabla \times E(\mathbf{x},\, t)\right) \\
&\qquad\qquad - s \nabla \Big(\nabla \cdot \left(\varepsilon(\mathbf{x}) E(\mathbf{x},\, t) \right) \Big) = 0, && (\mathbf{x},\, t) \in \Omega \times (0,\, T), && (100)\\
&E(\mathbf{x},\, 0) = f_1(\mathbf{x}), \quad E_t(\mathbf{x},\, 0) = f_2(\mathbf{x}), && \mathbf{x} \in \Omega, && (101)\\
&n \times E = 0 && (\mathbf{x},\, t) \in \partial\Omega \times (0,\, T). && (102)
\end{aligned}$$

In our computations we apply the Neumann boundary conditions to the left and right hand side of a domain and first order absorbing boundary conditions [32] at the rest of boundaries as well as homogeneous initial conditions. Then the forward problem used in our computations is

$$\begin{aligned}
&\varepsilon(\mathbf{x}) \frac{\partial^2 E(\mathbf{x},\, t)}{\partial t^2} + \nabla \times \left(\mu^{-1}(\mathbf{x}) \nabla \times E(\mathbf{x},\, t)\right) \\
&\qquad\qquad - s \nabla \Big(\nabla \cdot \left(\varepsilon(\mathbf{x}) E(\mathbf{x},\, t) \right) \Big) = 0, && (\mathbf{x},\, t) \in \Omega \times (0,\, T), && (103)\\
&E(\mathbf{x},\, 0) = 0, \quad E_t(\mathbf{x},\, 0) = 0, && \mathbf{x} \in \Omega, && (104)\\
&\partial_n E\,(\mathbf{x},\, t) = (0,\, f(t),\, 0), && (\mathbf{x},\, t) \in \partial_1\Omega \times (0,\, t_1]\,, && (105)\\
&\partial_n E(\mathbf{x},\, t) = -\partial_t E(\mathbf{x},\, t), && (\mathbf{x},\, t) \in \partial_1\Omega \times (t_1,\, T)\,, && (106)\\
&\partial_n E(\mathbf{x},\, t) = -\partial_t E(\mathbf{x},\, t), && (\mathbf{x},\, t) \in \partial_2\Omega \times (0,\, T), && (107)\\
&\partial_n E(\mathbf{x},\, t) = 0, && (\mathbf{x},\, t) \in \partial_3\Omega \times (0,\, T). && (108)
\end{aligned}$$

Due to application of Neumann boundary conditions, we are allowed to have assumption of infinite structure in lateral directions and thus, considering the CIPs of the reconstruction of unknown parameters ε and μ in a waveguide. Numerical solution of the forward problem (103)-(108) approximates well the solution of the original Maxwell's equations for $s = 1$, check [13].

9.1 The finite element method for the model problem

In this section we formulate the finite element method for the problem (103)-(108) which for convenience of the reader we write below

$$
\begin{aligned}
&\varepsilon(\mathbf{x})\frac{\partial^2 E(\mathbf{x},\, t)}{\partial t^2} + \nabla \times \left(\mu^{-1}(\mathbf{x})\nabla \times E(\mathbf{x},\, t)\right) \\
&\qquad - s\nabla\Big(\nabla \cdot \left(\varepsilon(\mathbf{x})E(\mathbf{x},\, t)\right)\Big) = 0, && (\mathbf{x},\, t) \in \Omega_T = \Omega \times (0,\, T), && (109)\\
&E(\mathbf{x},\, 0) = 0, \quad E_t(\mathbf{x},\, 0) = 0, && \mathbf{x} \in \Omega, && (110)\\
&\partial_n E\,(\mathbf{x},\, t) = (0,\, f(t),\, 0), && (\mathbf{x},\, t) \in S_{1,1} = \partial_1\Omega \times (0,\, t_1]\,, && (111)\\
&\partial_n E(\mathbf{x},\, t) = -\partial_t E(\mathbf{x},\, t), && (\mathbf{x},\, t) \in S_{1,2} = \partial_1\Omega \times (t_1,\, T)\,, && (112)\\
&\partial_n E(\mathbf{x},\, t) = -\partial_t E(\mathbf{x},\, t), && (\mathbf{x},\, t) \in S_2 = \partial_2\Omega \times (0,\, T), && (113)\\
&\partial_n E(\mathbf{x},\, t) = 0, && (\mathbf{x},\, t) \in S_3 = \partial_3\Omega \times (0,\, T). && (114)
\end{aligned}
$$

To do that we discretize $\Omega_{FEM} \times (0,T)$ presenting by $K_h = \{K\}$ a partition of the domain Ω_{FEM} into tetrahedra K, see p.10 of [15]. Here $h = h(x)$ is a mesh function defined as $h|_K = h_K-$ the local diameter of the elements, and we let J_k be a partition of the time interval $(0,T)$ into time subintervals $J = (t_{k-1}, t_k]$ of uniform length $\tau = t_k - t_{k-1}$. Here the minimal angle condition on the K_h [19] is considered as well. We need to define the finite element trial and test spaces, W_h^E and W_h^ϕ, respectively, to formulate the finite element method and solve the state problem (103)–(108). We introduce the finite element trial space W_h^E defined by

$$W_h^E := \{w \in W_E : w|_{K\times J} \in [P_1(K) \times P_1(J)]^3,\ \forall K \in K_h,\ \forall J \in J_\tau\},$$

where $P_1(K)$ and $P_1(J)$ denote the set of piecewise linear functions on K and J, respectively. and

$$W^E = \{w \in H^1(\Omega_T)^3 : w(\cdot, 0) = 0, \text{ where } \partial_n w|_{\partial\Omega} = -\partial_t w\},$$

We also introduce the finite element test space W_h^ϕ defined by

$$W_h^\phi := \{w \in W_\phi : w|_{K\times J} \in [P_1(K) \times P_1(J)]^3,\ \forall K \in K_h,\ \forall J \in J_\tau\},$$

where

$$W^\phi = \{w \in H^1(\Omega_T)^3 : w(\cdot, T) = 0, \text{ where } \partial_n w|_{\partial\Omega} = -\partial_t w\}.$$

Hence, the finite element spaces W_h^E and W_h^ϕ consist of continuous piecewise linear functions in space and time, which satisfy homogeneous initial and first order absorbing boundary conditions. We also define the following L_2 inner products and norms:

$$((p, q)) = \int_\Omega \int_0^T pq \mathrm{d}x\mathrm{d}t, \quad \|p\|^2 = ((p, p)), (\alpha, \beta) = \int_\Omega \alpha\beta \mathrm{d}x, \quad |\alpha|^2 = (\alpha, \alpha).$$

Multiplying (109) by $\bar{\phi} \in W^{\phi}$ and integrating in space and time we get :

$$\Bigg(\Bigg(\varepsilon \frac{\partial^2 E}{\partial t^2} + \nabla \times (\mu^{-1} \nabla \times E) - s\nabla(\nabla \cdot (\varepsilon E), \bar{\phi}\Bigg)\Bigg) = \Bigg(\Bigg(\varepsilon \frac{\partial^2 E}{\partial t^2}, \bar{\phi}\Bigg)\Bigg) +$$
$$\Bigg(\Bigg(\nabla \times (\mu^{-1} \nabla \times E), \bar{\phi}\Bigg)\Bigg) - s\Bigg(\Bigg(\nabla(\nabla \cdot (\varepsilon E)), \bar{\phi}\Bigg)\Bigg) = I_1 + I_2 + I_3,\ \forall \bar{\phi} \in W^{\phi}. \tag{115}$$

We now integrate by parts (115), term I_1 in time and terms I_2, I_3 in space, to get :

$$I_1 = \Bigg(\Bigg(\varepsilon \frac{\partial^2 E}{\partial t^2}, \bar{\phi}\Bigg)\Bigg) = \Bigg(\varepsilon, \Bigg[\bar{\phi} \cdot \frac{\partial E}{\partial t}\Bigg]_0^T - \int_0^T \frac{\partial E}{\partial t} \frac{\partial \bar{\phi}}{\partial t} dt\Bigg)$$
$$= \Bigg(\varepsilon, \Bigg[\phi(\bar{t}) \frac{\partial E}{\partial t}(T) - \bar{\phi}(0) \frac{\partial E}{\partial t}(0)\Bigg] - \int_0^T \frac{\partial E}{\partial t} \frac{\partial \bar{\phi}}{\partial t} dt\Bigg) = \Bigg(\Bigg(\varepsilon \frac{\partial E}{\partial t}, \frac{\partial \bar{\phi}}{\partial t}\Bigg)\Bigg). \tag{116}$$

Since $\mu = \varepsilon = 1$ on $\partial\Omega$ and in Ω_{FDM} we have

$$I_2 = \Bigg(\nabla \times (\mu^{-1} \nabla \times E), \bar{\phi}\Bigg) = \Bigg(\bar{\phi} \cdot n \times (\mu^{-1} \nabla \times E) \mid_{\partial\Omega}\Bigg) + \Bigg(\Bigg(\mu^{-1} \nabla \times E, \nabla \times \bar{\phi}\Bigg)\Bigg) =$$
$$\Bigg(n \times \bar{\phi}, \mu^{-1} \nabla \times E\Bigg)_{S_{1,1} \cup S_{1,2} \cup S_2 \cup S_3} + \Bigg(\Bigg(\mu^{-1} \nabla \times E, \nabla \times \bar{\phi}\Bigg)\Bigg) =$$
$$\Bigg(p(t), \bar{\phi}\Bigg)_{S_{1,1}} - \Bigg(\frac{\partial E}{\partial t}, \bar{\phi}\Bigg)_{S_{1,2} \cup S_2} + \Bigg(\Bigg(\mu^{-1} \nabla \times E, \nabla \times \bar{\phi}\Bigg)\Bigg) \tag{117}$$

with $p(t) = (0, f(t), 0)$.

$$I_3 = s\Bigg(\Bigg(\nabla(\nabla \cdot (\varepsilon E), \bar{\phi}\Bigg)\Bigg) = s\Bigg[\Bigg(\bar{\phi} \cdot n \cdot (\nabla \cdot (\varepsilon E))\Bigg)_{\partial\Omega} - \Bigg(\Bigg(\nabla \cdot (\varepsilon E), \nabla \cdot \bar{\phi}\Bigg)\Bigg)\Bigg] = -s\Bigg(\Bigg(\nabla \cdot (\varepsilon E), \nabla \cdot \bar{\phi}\Bigg)\Bigg). \tag{118}$$

Collecting I_1, I_2, I_3 we get the following variational formulation :

$$\Bigg(\Bigg(\varepsilon \frac{\partial E}{\partial t}, \frac{\partial \bar{\phi}}{\partial t}\Bigg)\Bigg) + \Bigg(p(t), \bar{\phi}\Bigg)_{S_{1,1}} - \Bigg(\frac{\partial E}{\partial t}, \bar{\phi}\Bigg)_{S_{1,2} \cup S_2} + \Bigg(\Bigg(\mu^{-1} \nabla \times E, \nabla \times \bar{\phi}\Bigg)\Bigg) + s\Bigg(\Bigg(\nabla \cdot (\varepsilon E), \nabla \cdot \bar{\phi}\Bigg)\Bigg) = 0. \tag{119}$$

Also E is the weak solution of equation (109) if $E \in W^E$ and (119) holds. We approximate $E(\mathbf{x}, t)$ with $E_h(\mathbf{x}, t) \in W_h^E$ and the finite element method reads : find $E_h \in W_h^E$ such that

$$\Bigg(\Bigg(\varepsilon \frac{\partial E_h}{\partial t}, \frac{\partial \bar{\phi}}{\partial t}\Bigg)\Bigg) + \Bigg(p(t), \bar{\phi}\Bigg)_{S_{1,1}} - \Bigg(\frac{\partial E_h}{\partial t}, \bar{\phi}\Bigg)_{S_{1,2} \cup S_2} +$$
$$\Bigg(\Bigg(\mu^{-1} \nabla \times E_h, \nabla \times \bar{\phi}\Bigg)\Bigg) + s\Bigg(\nabla \cdot (\varepsilon E_h), \nabla \cdot \bar{\phi}\Bigg)\Bigg) = 0, \quad \forall \bar{\phi} \in W_h^{\phi}.$$

9.2 The explicit FEM scheme for the electric field

We expand $E_h(x, t)$ in terms of the standard continuous piecewise linear functions $\{\phi_i(x)\}_{i=1}^m$ in space and $\{\psi_k(t)\}_{k=1}^n$ in time as $E_h(x, t) = \Sigma_{k=1}^n \Sigma_{i=1}^m E_i \phi_i(x) \psi_k(t)$, where $E_i = E_{h_{i,k}}$ denote unknown coefficients at the point i of the mesh K_h and time moment k substitute this expansion in variational formulation (119) with $\bar{\phi}(x,t) = \phi_j(x)\psi_l(t)$ and obtain the following system of discrete equations

$$
\begin{aligned}
&- \sum_{K \in \Omega_{FEM}} \sum_{k,l=1}^{n} \sum_{i,j=1}^{m} E_i \int_k \varepsilon(x)\phi_i(x)\phi_j(x) \int_{t_{k-1}}^{t_{k+1}} \partial_t\psi_k(t)\partial_t\psi_l(t) dxdt \\
&- \sum_{\partial K \in \partial\Omega_{FEM}} \sum_{k,l=1}^{n} \sum_{i,j=1}^{m} E_i \int_{\partial k} \phi_i(x)\phi_j(x) \int_{t_{k-1}}^{t_{k+1}} \partial_t\psi_k(t)\psi_l(t) dSdt \\
&+ \sum_{K \in \Omega_{FEM}} \sum_{k,l=1}^{n} \sum_{i,j=1}^{m} E_i \int_k \mu^{-1}\nabla \times \phi_i(x)\nabla \times \phi_j(x) \int_{t_{k-1}}^{t_{k+1}} \psi_k(t)\psi_l(t) dxdt \\
&+s \sum_{K \in \Omega_{FEM}} \sum_{k,l=1}^{n} \sum_{i,j=1}^{m} E_i \int_k \nabla \cdot (\varepsilon\phi_i(x))\nabla \cdot \phi_j(x) \int_{t_{k-1}}^{t_{k+1}} \psi_k(t)\psi_l(t) dxdt \\
&+ \sum_{\partial K \in \partial\Omega_{FEM}} \sum_{k,l=1}^{n} \sum_{i,j=1}^{m} \int_{\partial K} \int_{t_{k-1}}^{t_{k+1}} p_k\phi_j(x)\psi_l(t) dxdt \\
&= I_1 + I_2 + I_3 + I_4 + I_5 = 0.
\end{aligned} \tag{120}
$$

Here p_k are nodes values of $p(t)$ at nodes of the time mesh J_τ.

Similarly with [13], now by the definition of piecewise linear functions in time, we are able to compute explicitly the time integrals appearing in (120). After substituting computed integrals in time in (120) we get the following linear system of equations

$$
\begin{aligned}
M(E^{k+1} - 2E^k + E^{k-1}) &= -\tau^2 K\big(\frac{1}{6}E^{k-1} + \frac{2}{3}E^k + \frac{1}{6}E^{k+1}\big) \\
-s\tau^2 C\big(\frac{1}{6}E^{k-1} + \frac{2}{3}E^k + \frac{1}{6}E^{k+1}\big) &+ \frac{1}{2}\tau M_{\partial\Omega}(E^{k+1} - E^{k-1}) - \tau^2 p^k.
\end{aligned} \tag{121}
$$

In the scheme above E^0 and E^1 are equal to zero. Here M and $M_{\partial\Omega}$ denote the block mass matrices in space, in Ω and over the boundary of Ω, respectively. K is the block stiffness matrix corresponding to the rotation term, C is the stiffness matrices corresponding to the divergence terms, and τ is the time step.

Similarly with [13] we define the mapping F_K for the reference element $\hat{K}$ such that $F_K(\hat{K}) = K$ and let $\hat{\phi}$ be the piecewise linear local basis function on the the reference element $\hat{K}$ where $\phi \circ F_K = \hat{\phi}$. The matrices entries in (121) can be computed explicitly by:

$$
\begin{aligned}
M_{i,j}^K &= (\varepsilon\phi_i \circ F_K, \phi_j \circ F_K)_K, \\
M_{i,j}^{\partial\Omega} &= (\phi_i \circ F_K, \phi_j \circ F_K)_{\partial\Omega_K}, \\
K_{i,j}^K &= (\mu^{-1}\nabla \times \varphi_i \circ F_K, \nabla \times \varphi_j \circ F_K)_K, \\
C_{i,j}^K &= (\nabla \cdot (\varepsilon\phi_i) \circ F_K), \nabla \cdot \phi_j \circ F_K)_K, \\
P_{i,j}^K &= (p_k, \phi_j)_K.
\end{aligned}
$$

We approximate M by the lumped mass matrix M^L in space to obtain an explicit scheme. In other words, the diagonal approximation is obtained by taking the row sum of M [35, 36]. We also use the mass lumping in time by replacing terms corresponding to the mass matrix in time, $\frac{E^{k-1}}{6} + \frac{2E^k}{3} + \frac{E^{k+1}}{6}$, by E^k. Then (121) becomes

$$M(E^{k+1} - 2E^k + E^{k-1}) = -\tau^2 K E^k - s\tau^2 C E^k + \frac{1}{2}\tau M_{\partial\Omega}(E^{k+1} - E^{k-1}) - \tau^2 p^k. \tag{122}$$

Next, we multiply (122) by $(M^L)^{-1}$ to obtain the following fully explicit time-stepping method to solve (109)-(114):

$$\begin{aligned} E^{k+1}\big(1 - \frac{1}{2}\tau M^{\partial\Omega}(M^L)^{-1}\big) &= 2E^k - \tau^2 (M^L)^{-1} K E^k - s\tau^2 (M^L)^{-1} C E^k \\ &\quad -\big(1 + \frac{1}{2}\tau M^{\partial\Omega}(M^L)^{-1}\big) E^{k-1} - \tau^2 (M^L)^{-1} p^k. \end{aligned} \tag{123}$$

When we apply finite element method only in Ω_{FEM} then the scheme (123) will be reduced to :

$$E^{k+1} = 2E^k - \tau^2 (M^L)^{-1} K E^k - s\tau^2 (M^L)^{-1} C E^k - E^{k-1} - \tau^2 (M^L)^{-1} p^k. \tag{124}$$

9.3 The explicit FDM scheme

As mentioned before, we set $\varepsilon(x) = \mu(x) = 1$ in Ω_{FDM} so in Ω_{FDM} we need to solve a system of vector wave equations for vector field $E = E(E_1, E_2, E_3)$:

$$\begin{aligned} &\partial_t^2 E - \Delta E = 0, && && (125) \\ &E(\mathbf{x}, 0) = 0, \quad E_t(\mathbf{x}, 0) = 0, && \mathbf{x} \in \Omega, && (126) \\ &\partial_n E(\mathbf{x}, t) = (0, f(t), 0), && (\mathbf{x}, t) \in \partial_1\Omega \times (0, t_1], && (127) \\ &\partial_n E(\mathbf{x}, t) = -\partial_t E(\mathbf{x}, t), && (\mathbf{x}, t) \in \partial_1\Omega \times (t_1, T), && (128) \\ &\partial_n E(\mathbf{x}, t) = -\partial_t E(\mathbf{x}, t), && (\mathbf{x}, t) \in \partial_2\Omega \times (0, T), && (129) \\ &\partial_n E(\mathbf{x}, t) = 0, && (\mathbf{x}, t) \in \partial_3\Omega \times (0, T). && (130) \end{aligned}$$

We gain the below explicit scheme in Ω_{FDM} by using standard finite difference discretization:

$$E_{i,j,m}^{k+1} = -\tau^2 \Delta E_{i,j,m}^k + 2E_{i,j,m}^k - E_{i,j,m}^{k-1}. \tag{131}$$

Here $E_{i,j,m}^k$ denotes the solution on the time iteration k at the discrete point (i, j, m), τ is the time step, and $\Delta E_{i,j,m}^k$ is the discrete Laplacian which can be written as

$$\begin{aligned} \Delta E_{i,j,m}^k = &\frac{E_{i+1,j,m}^k - 2E_{i,j,m}^k + E_{i-1,j,m}^k}{dx_1^2} + \frac{E_{i,j+1,m}^k - 2E_{i,j,m}^k + E_{i,j-1,m}^k}{dx_2^2} \\ &+ \frac{E_{i,j,m+1}^k - 2E_{i,j,m}^k + E_{i,j,m-1}^k}{dx_3^2}, \end{aligned} \tag{132}$$

where dx_1, dx_2, and dx_3 denote the steps of discrete finite difference meshes in x_1, x_2, x_3 directions, respectively.

To approximate absorbing boundary conditions in (128)–(130) we apply a forward finite difference approximation at the middle point. This helps to get a numerical approximation of higher order than a normal backward or forward finite difference approximation. At the left boundary of Ω_{FDM} the absorbing boundary conditions (127) can be written as

$$\frac{\partial E(x,t)}{\partial x} = \frac{\partial E(x,t)}{\partial t}. \tag{133}$$

By discretizing (133) we get:

$$\frac{E^{k+1}_{i,j,m} - E^{k}_{i,j,m}}{dt} + \frac{E^{k+1}_{i+1,j,m} - E^{k}_{i+1,j,m}}{dt} - \frac{E^{k}_{i+1,j,m} - E^{k}_{i,j,m}}{dx} - \frac{E^{k+1}_{i+1,j,m} - E^{k}_{i,j,m}}{dx} = 0.$$

The above equation can be transformed to get the explicit scheme for computation of boundary condition on the left boundary of Ω_{FDM} :

$$E^{k+1}_{i,j,m} = E^{k}_{i+1,j,m} + E^{k}_{i,j,m}\frac{dx - dt}{dx + dt} - E^{k+1}_{i+1,j,m}\frac{dx - dt}{dx + dt}.$$

Similarly we get discretizations for all other boundaries.

9.4 The domain decomposition FEM/FDM methods

We consider the data correlation between the finite element method on the unstructured part of the mesh, Ω_{FEM}, and the finite difference method on the structured part, Ω_{FDM} for the solution of the forward problem in (103)-(108). Their correlation is related to having mesh overlapping across a two-element thick layer around Ω_{FEM}, see Figure 6 for the domain decomposition.

The interior nodes of computational domain Ω belong to either of the following sets, see Figure 6.

$N_\circ$ nodes '$\circ$' interior to Ω_{FDM} that lie on the boundary of Ω_{FEM},

$N_\Diamond$ nodes '$\Diamond$' interior to Ω_{FEM} that lie on the boundary of Ω_{FDM},

N_* nodes '$*$' interior to Ω_{FEM} that are not contained in Ω_{FDM},

N_+ nodes '$+$' interior to Ω_{FDM} that are not contained in Ω_{FEM}.

Stability of the explicit scheme:

Since we use explicit domain decomposition FEM/FDM, we also need to check stability such that with our choice of the time step τ the whole scheme (131) remains stable. We use stability analysis on the structured meshes and we choose the largest time step in our computation according to CFL stability condition [9] where

$$\tau \leq \frac{\sqrt{\varepsilon\mu}}{\sqrt{\frac{1}{dx_1^2}+\frac{1}{dx_2^2}+\frac{1}{dx_3^2}}}. \tag{134}$$

$$\tag{135}$$

We assume $dx_1 = dx_2 = dx_3$, and condition (134) is transformed in 3D as

$$\tau \leq h\sqrt{\frac{\varepsilon\mu}{3}} \tag{136}$$

Algorithm for the hybrid FEM/FDM method At every time step we do following operation:

1 Compute E^{k+1} from (131) with absorbing boundary conditions (128)-(129) at $\partial\Omega$ on the structured part of the mesh Ω_{FDM} when E^k and E^{k-1} are known.

2 Compute E^{k+1} by means of explicit finite element scheme (124) on the unstructured part of the mesh Ω_{FEM} when E^k and E^{k-1} are known.

3 Use the values of the electric field E^{k+1} at nodes $N_{\diamondsuit}$, which are computed using the finite element scheme (124), as a boundary condition for the finite difference method in Ω_{FDM}.

4 Use the values of the electric field E^{k+1} at nodes $N_{\circ}$, which are computed using the finite difference scheme (131), as a boundary condition for the finite element method in Ω_{FEM}.

5 Apply swap of the solutions for the electric field in order to apply the algorithm on a new time level k.

10 Statement of Inverse Problem

In this section, we present the coefficient inverse problems (CIP). We also describe how we optimize our inverse problem by means of domain decomposition finite element/finite difference and Tikhonov functional.

10.1 Coefficient Inverse Problem (CIP)

Assume that the coefficients ε and μ satisfy (94)–(95) such that b_1, b_2 are given. Also functions ε, μ are unknown in the domain $\Omega \setminus \Omega_{FDM}$. Determine the functions $\varepsilon\left(x\right), \mu(x)$ for $x \in \Omega \backslash \Omega_{\mathrm{FDM}}$, assuming that the following function $\tilde{E}\left(x, t\right)$ is known

$$E\left(\mathbf{x}, t\right) = \tilde{E}\left(\mathbf{x}, t\right), \; \forall\left(\mathbf{x}, t\right) \in S_T. \tag{137}$$

In applications, the assumption $\varepsilon(x) = \mu(x) = 1$ for $x \in \Omega_{\mathrm{FDM}}$ means that the functions $\varepsilon(x)$ and $\mu(x)$ have a known constant value outside of the medium of interest $\Omega \backslash \Omega_{\mathrm{FDM}}$. The function $\tilde{E}(x,t)$ models time dependent measurements of the electric wave field at the backscattering boundary $\partial_1\Omega$ of the domain of interest. In practice, measurements are performed on a number of detectors, see [12, 57].

10.2 Optimization method via Tikhonov functional

We reformulate our inverse problem as an optimization problem and solve the equations (103)-(108) in a way that two functions, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$, lead us to a best fit solution to time and space domain observations $\widetilde{E}(\mathbf{x},t)$ measured at a finite number of observation points on $\partial_1\Omega$.

We introduce the Tikhonov regularization functional to solve our CIP as

$$\begin{aligned} F(E,\varepsilon,\mu) = \frac{1}{2}\int_{S_T} \big(E(\mathbf{x},\,t) - \tilde{E}(\mathbf{x},\,t)\big)^2 z_\delta(t)\,\mathrm{d}\sigma\,\mathrm{d}\mathbf{x} \\ +\frac{1}{2}\gamma_1\int_\Omega \big(\varepsilon(\mathbf{x}) - \varepsilon_0(\mathbf{x})\big)^2\,\mathrm{d}\mathbf{x} \\ +\frac{1}{2}\gamma_2\int_\Omega \big(\mu(\mathbf{x}) - \mu_0(\mathbf{x})\big)^2\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{138}$$

The regularization parameters, γ_1 and γ_2, are strictly positive constants and ε_0 and μ_0 are initial guess for permittivity and permeability functions, respectively. Moreover, $S_T := \partial_1\Omega \times (0,T)$ where $\partial_1\Omega$ is the backscattering side of the domain Ω with the time domain observations.

Here z_δ is a cut-off function, which is introduced to ensure that the compability conditions at $\overline{\Omega}_T \cap \{t = T\}$ for the adjoint problem (check section 4.3 of [5]) are satisfied, and δ is a strictly small positive number. We choose a function z_δ such that

$$z_\delta \in C^\infty[0,T],\; z_\delta(t) = \begin{cases} 1 & \text{for } t \in [0, T-\delta], \\ 0 & \text{for } t \in \left(T - \frac{\delta}{2}, T\right], \\ 0 < z_\delta < 1 & \text{for } t \in \left(T-\delta, T-\frac{\delta}{2}\right). \end{cases}$$

For our analysis we introduce the following spaces of real valued vector functions

$$\begin{aligned} &H_E^1 := \{w \in H^1(\Omega_T) : w(\cdot,0) = 0\}, \\ &H_\lambda^1 := \{w \in H^1(\Omega_T) : w(\cdot,T) = 0\}, \\ &U^1 = H_E^1(\Omega_T) \times H_\lambda^1(\Omega_T) \times C\left(\overline{\Omega}\right) \times C\left(\overline{\Omega}\right), \\ &U^0 = L_2(\Omega_T) \times L_2(\Omega_T) \times L_2(\Omega) \times L_2(\Omega). \end{aligned}$$

Next we introduce the Lagrangian for minimization of the functional (138)

$$
\begin{aligned}
L(u) = F(E,\varepsilon,\mu) &- \int_{\Omega_T} \varepsilon \frac{\partial \lambda}{\partial t}\frac{\partial E}{\partial t}\, dxdt + \int_{\Omega_T} (\mu^{-1}\nabla \times E)(\nabla \times \lambda)\, dxdt \\
&+ s\int_{\Omega_T} (\nabla \cdot (\varepsilon E))(\nabla \cdot \lambda)\, dxdt \\
&- \int_{S_{1,1}} \lambda p(t)\, d\sigma dt - \int_{S_{1,2}} \lambda \partial_t E\, d\sigma dt - \int_{S_2} \lambda \partial_t E\, d\sigma dt.
\end{aligned} \tag{139}
$$

Here $u = (E,\, \lambda,\, \varepsilon,\, \mu) \in U^1$ and $p(t) = (0, f(t), 0)$. Also we have assumed $S_{1,1} := \partial_1\Omega \times (0, t_1]$, $S_{1,2} := \partial_1\Omega \times (t_1, T)$, $S_2 := \partial_2\Omega \times (0,T)$, $S_3 := \partial_3\Omega \times (0,T)$.

We seek for a stationary point with respect to u such that $\forall \bar{u} = (\bar{E},\, \bar{\lambda}, \bar{\varepsilon_{\mathrm{r}}},\, \bar{\mu_{\mathrm{r}}}) \in U^1$,

$$
L'(u;\bar{u}) = 0, \tag{140}
$$

where L' denote the Jacobian of L at u. Functions E and λ depend on the $\varepsilon(\mathrm{x})$ and $\mu(\mathrm{x})$, also E and λ represent the weak form solutions of the forward and adjoint problems respectively, check [11] for details. For knowing how to obtain the Jacobian L' of the Lagrangian precisely, see section 4.8 of [5] where we assume that the variations of functions E and λ depend on variations of the coefficients ε and μ. In this work, we assume that the elements of the vector function $(E,\, \lambda, \varepsilon,\, \mu)$ in Jacobian L' of the Lagrangian are not dependent on each other. In both cases we obtain the same expressions for L', see [5] for similar conclusion.

We assume that $\lambda\,(x,T) = \partial_t\lambda\,(x,T) = 0$ and try to impose such conditions on the function λ that $L\,(E,\lambda,\varepsilon,\mu) := L\,(u) = F\,(E,\varepsilon,\mu)$. Next, we use the fact that $\lambda(x,T) = \frac{\partial\lambda}{\partial t}(x,T) = 0$ and $E(x,0) = \frac{\partial E}{\partial t}(x,0) = 0$, as well as $\mu = \varepsilon = 1$ on $\partial\Omega$, together with boundary conditions $\partial_n E = 0$ and $\partial_n \lambda = 0$ on S_3. We use the equation (140) to gain the Jacobian $L'(u)$ so we have for all $\bar{u}$,

$$
\begin{aligned}
0 = \frac{\partial L}{\partial \lambda}(u)(\bar{\lambda}) = &- \int_{\Omega_T} \varepsilon \frac{\partial \bar{\lambda}}{\partial t}\frac{\partial E}{\partial t}\, dxdt + \int_{\Omega_T} (\mu^{-1}\nabla \times E)(\nabla \times \bar{\lambda})\, dxdt \\
&+ s\int_{\Omega_T} (\nabla \cdot (\varepsilon E))(\nabla \cdot \bar{\lambda})\, dxdt \\
&- \int_{S_{1,1}} \bar{\lambda} p(t)\, d\sigma dt - \int_{S_{1,2}} \bar{\lambda} \partial_t E\, d\sigma dt \\
&- \int_{S_2} \bar{\lambda} \partial_t E\, d\sigma dt,\ \ \forall \bar{\lambda} \in H^1_{\lambda}(\Omega_T),
\end{aligned} \tag{141}
$$

$$
\begin{aligned}
0 = \frac{\partial L}{\partial E}(u)(\bar{E}) = &\int_{S_T} (E - \tilde{E})\, \bar{E}\, z_\delta\, dxdt \\
&- \int_{\Omega_T} \varepsilon \frac{\partial \lambda}{\partial t}\frac{\partial \bar{E}}{\partial t}\, dxdt + \int_{\Omega_T} (\mu^{-1}\nabla \times \lambda)(\nabla \times \bar{E})\, dxdt \\
&+ s\int_{\Omega_T} (\nabla \cdot \lambda)(\nabla \cdot (\varepsilon \bar{E}))\, dxdt,\ \ \forall \bar{E} \in H^1_E(\Omega_T).
\end{aligned} \tag{142}
$$

Finally, we obtain two equations that express that the gradients with respect to ε and μ vanish:

$$\begin{aligned} 0 = \frac{\partial L}{\partial \varepsilon}(u)(\bar{\varepsilon}) = &-\int_{\Omega_T} \frac{\partial \lambda}{\partial t}\frac{\partial E}{\partial t}\bar{\varepsilon}\, dxdt + s\int_{\Omega_T} (\nabla \cdot E)(\nabla \cdot \lambda)\bar{\varepsilon}\, dxdt \\ &+ \gamma_1 \int_{\Omega} (\varepsilon - \varepsilon_0)\bar{\varepsilon}\, dx,\ x \in \Omega, \end{aligned} \tag{143}$$

$$0 = \frac{\partial L}{\partial \mu}(u)(\bar{\mu}) = -\int_{\Omega_T} \frac{1}{\mu^2}\ \nabla \times E\ \nabla \times \lambda\ \bar{\mu}\ dxdt + \gamma_2 \int_{\Omega} (\mu - \mu_0)\ \bar{\mu}\ dx,\ x \in \Omega. \tag{144}$$

The equation (141) is the weak formulation of the state equation and the equation (142) is the weak formulation of the following adjoint problem

$$\begin{aligned} \varepsilon\frac{\partial^2 \lambda}{\partial t^2} + \nabla \times (\mu^{-1}\nabla \times \lambda) - s\varepsilon\nabla(\nabla \cdot \lambda) &= -(E - \tilde{E})z_\delta,\ x \in S_T, \\ \lambda(\cdot, T) &= \frac{\partial \lambda}{\partial t}(\cdot, T) = 0, \\ \partial_n \lambda &= 0,\ \text{on } S_3. \end{aligned} \tag{145}$$

10.3 Finite element method for CIP

For convenience of reader we repeat the definitions of the trial and test spaces from section 9. We discretize $\Omega_{FEM} \times (0, T)$ presenting by $K_h = \{K\}$ a partition of the domain Ω_{FEM} into tetrahedra K. Here $h = h(x)$ is a mesh function defined as $h|_K = h_K-$ the local diameter of the elements, and we let J_k be a partition of the time interval $(0, T)$ into time subintervals $J = (t_{k-1}, t_k]$ of uniform length $\tau = t_k - t_{k-1}$. Here he minimal angle condition on the K_h [19] is considered as well.

We need to define the finite element spaces V_h, W_h^E and W_h^λ to formulate the finite element method and solve the the adjoint problem. We introduce the finite element trial space W_h^E for the electric field E for the case when functions $\varepsilon(x), \mu(x)$ are smooth as

$$W_h^E := \{w \in H_E^1 : w|_{K \times J} \in [P_1(K) \times P_1(J)]^3,\ \forall K \in K_h,\ \forall J \in J_\tau\},$$

where $P_1(K)$ and $P_1(J)$ denote the set of linear functions on K and J, respectively. We also introduce the finite element test space W_h^λ defined by

$$W_h^\lambda := \{w \in H_\lambda^1 : w|_{K \times J} \in [P_1(K) \times P_1(J)]^3,\ \forall K \in K_h,\ \forall J \in J_\tau\}.$$

Hence, the finite element spaces W_h^E and W_h^λ consist of continuous piecewise linear functions in space and time.

The space of piecewise constant functions $V_h \subset L_2(\Omega)$ is used to approximate functions ε and μ, where

$$V_h := \{u \in L_2(\Omega) : u|_K \in P_0(K), \forall K \in K_h\},$$

and $P_0(K)$ is the piecewise constant function on K.

In a general case we allow functions $\varepsilon(x), \mu(x)$ to be discontinuous. Let S be the internal face of the non-empty intersection of the boundaries of two neighbouring elements K^+ and K^-. We denote the jump of the function v_h computed from the two neighbouring elements K^+ and K^- sharing the common side S as

$$[v_h] = v_h^+ - v_h^-, \tag{146}$$

and the jump of the normal component v_h across the side S as

$$[[v_h]] = v_h^+ \cdot n^+ + v_h^- \cdot n^-, \tag{147}$$

where n^+, n^- is the unit outward normal on S^+, S^-, respectively.

Thus, for a general case when functions $\varepsilon(x), \mu(x)$ are discontinuous we also introduce the discontinuous finite element space W_h as

$$W_h = \big\{ v(x) \in H^1(\Omega) : \; v|_K \in DP_1(K) \; \forall K \in K_h \big\}, \tag{148}$$

where $DP_1(K)$ denotes the discontinuous piecewise linear function on K. The finite element space W_h is constructed such that $W_h \subset H^1(\Omega)$.

Let P_h be the $L_2(\Omega)$ orthogonal projection. We define by f_h^I the standard nodal interpolant [28] of f into the space of continuous piecewise-linear functions on the mesh K_h. Then by one of properties of the orthogonal projection

$$\|f - P_h f\|_{L_2(\Omega)} \leq \left\| f - f_h^I \right\|_{L_2(\Omega)}. \tag{149}$$

It follows from [55] that

$$\|f - P_h f\|_{L_2(\Omega)} \leq C_I h \, \| \, f \|_{H^1(\Omega)}, \forall f \in H^1(\Omega), \tag{150}$$

where $C_I = C_I\left(\Omega\right)$ is positive constant depending only on the domain Ω.

Next, we define $U_h = W_h^E \times W_h^\lambda \times V_h \times V_h$. When functions $\varepsilon(x), \mu(x)$ are discontinuous, we introduce the space $DU_h = W_h^E \times W_h^\lambda \times W_h \times W_h$. Usually $\dim U_h < \infty$ and $U_h \subset U^1$ as a set and we consider U_h as a discrete analogue of the space U^1. We introduce the same norm in U_h as the one in U^0, $\|\cdot\|_{U_h} := \|\cdot\|_{U^0}$. This means that in finite dimensional spaces all norms are equivalent and in our computations we compute coefficients in the space V_h. The finite element method for the case of continuous functions $\varepsilon(x), \mu(x)$ reads: Find $u_h \in U_h$, such that

$$L'(u_h)(\bar{u}) = 0 \; \forall \bar{u} \in U_h.$$

The finite element method for the case of discontinuous functions $\varepsilon(x), \mu(x)$ reads: Find $u_h \in DU_h$, such that

$$L'(u_h)(\bar{u}) = 0 \; \forall \bar{u} \in DU_h. \tag{151}$$

10.4 Fully discrete scheme for the adjoint problem

We expand $\lambda(x,t)$ in terms of the standard continuous piecewise linear functions $\{\phi_i\}_{i=1}^m$ in space and $\{\psi_i\}_{i=1}^n$ in time and substitute them into (145), similar to section 9.2, to obtain the following system of linear equation:

$$M(\lambda^{k+1} - 2\lambda^k + \lambda^{k-1}) = -\tau^2 S^k - \tau^2 K\lambda^k - s\tau^2 D\lambda^k. \tag{152}$$

with the initial condition $\lambda(\cdot,T) = \frac{\partial\lambda}{\partial t}(\cdot,T) = 0$.

Here, M is the block mass matrix in space, K is the block stiffness matrix corresponding to the rotation term, D is the stiffness matrices corresponding to the divergence term, S^k is the load vector at time level t_k, λ^k denote the nodal values of $\lambda(\cdot,t_k)$, τ is the time step. Similarly with section 9.2 we get the explicit formulas for the entries in system (145) at each element K as:

$$\begin{aligned}
M_{i,j}^K &= (\varepsilon\, \varphi_i \circ F_K, \varphi_j \circ F_K)_K, \\
K_{i,j}^K &= (\mu^{-1}\nabla \times \varphi_i \circ F_K, \nabla \times \varphi_j \circ F_K)_K, \\
D_{i,j}^K &= (\varepsilon\nabla \cdot \varphi_i \circ F_K, \nabla \cdot \varphi_j \circ F_K)_K, \\
S_{j,m}^K &= (E - \tilde{E}, \varphi_j \circ F_K)_K,
\end{aligned} \tag{153}$$

where $(\cdot,\cdot)_K$ denotes the $L_2(K)$ scalar product. To obtain an explicit scheme, we approximate M with the lumped mass matrix M^L (for further details, see [26]). Next, we multiply (153) with $(M^L)^{-1}$ and get the following explicit method:

$$\lambda^{k-1} = -\tau^2(M^L)^{-1}S^k + 2\lambda^k - \tau^2(M^L)^{-1}K\lambda^k - s\tau^2(M^L)^{-1}D\lambda^k - \lambda^{k+1}. \tag{154}$$

11 General Framework of a Posteriori Error Estimate

In this section we briefly present a posteriori error estimates for three kinds of errors as they are described in [4]:

- For the error $|L(u) - L(u_h)|$ in the Lagrangian (139),
- For the error $|F(\varepsilon,\mu) - F(\varepsilon_h,\mu_h)|$ in the Tikhonov functional (138),
- For the errors $|\varepsilon - \varepsilon_h|$ and $|\mu - \mu_h|$ in the regularized solutions of this functional ε,μ.

Here, $u_h, \varepsilon_h, \mu_h$ are finite element approximations of the functions u, ε, μ, respectively. A posteriori error estimate in the Lagrangian was already derived in [14] for the case when only the function $\varepsilon(x)$ in system (109) is unknown. In [20, 21] were derived a posteriori error estimate in the Lagrangian which corresponds to modified system (109). A posteriori error in the Lagrangian (139) can be derived straightforward from a posteriori error estimate presented in [14] and thus, all details of this derivation are not presented here.

However, to make clear how a posteriori errors in the Lagrangian and in the Tikhonov functional can be obtained, we present general framework for them. First we note that

$$\begin{aligned} F(\varepsilon,\mu)-F(\varepsilon_h,\mu_h) &= F'_\varepsilon(\varepsilon_h,\mu_h)(\varepsilon-\varepsilon_h)+F'_\mu(\varepsilon_h,\mu_h)(\mu-\mu_h)+R(\varepsilon,\varepsilon_h)+R(\mu,\mu_h),\\ L(u)-L(u_h) &= L'(u_h)(u-u_h)+R(u,u_h), \end{aligned} \tag{155}$$

where $R(\varepsilon,\varepsilon_h), R(\mu,\mu_h), R(u,u_h)$, are remainders of the second order. We assume that (ε_h,μ_h) are located in the small neighborhood of the regularized solutions (ε,μ), correspondingly. Thus, since the terms $R(u,u_h), R(\varepsilon,\varepsilon_h), R(\mu,\mu_h)$ are of the second order then they will be small and we can neglect them in (155).

We now use the splitting

$$\begin{aligned} u-u_h &= (u-u_h^I)+(u_h^I-u_h),\\ \varepsilon-\varepsilon_h &= (\varepsilon-\varepsilon_h^I)+(\varepsilon_h^I-\varepsilon_h),\\ \mu-\mu_h &= (\mu-\mu_h^I)+(\mu_h^I-\mu_h), \end{aligned} \tag{156}$$

together with the Galerkin orthogonality principle

$$\begin{aligned} L'(u_h)(\bar{u}) &= 0 \ \ \forall \bar{u}\in U_h,\\ F'(z_h)(b) &= 0 \ \ \forall b\in V_h \ \text{ or } \ \forall b\in W_h, \end{aligned} \tag{157}$$

insert (156) into (155) and get the following error representations:

$$\begin{aligned} L(u)-L(u_h) &\approx L'(u_h)(u-u_h^I),\\ F(\varepsilon,\mu)-F(\varepsilon_h,\mu_h) &\approx F'_\varepsilon(\varepsilon_h,\mu_h)(\varepsilon-\varepsilon_h^I)+F'_\mu(\varepsilon_h,\mu_h)(\mu-\mu_h^I). \end{aligned} \tag{158}$$

In (156), (158) functions $u_h^I\in U_h$ and $\varepsilon_h^I,\mu_h^I\in V_h$ or $\varepsilon_h^I,\mu_h^I\in W_h$ denote the interpolants of u,ε,μ, correspondingly.

Using (158) we conclude that a posteriori error estimate in the Lagrangian involves the derivative of the Lagrangian $L'(u_h)$ which we define as a residual, multiplied by weights $u-u_h^I$. Similarly, a posteriori error estimate in the Tikhonov functional involves the derivatives of the Tikhonov functional $F'_\varepsilon(\varepsilon_h,\mu_h)$ and $F'_\mu(\varepsilon_h,\mu_h)$ which represents residuals, multiplied by weights $\varepsilon-\varepsilon_h^I$ and $\mu-\mu_h^I$, correspondingly.

To derive the errors $|\varepsilon-\varepsilon_h|$ and $|\mu-\mu_h|$ in the regularized solutions ε,μ of the functional (138) we will use the convexity property of the Tikhonov functional together with the interpolation property (150). We now make both error estimates more explicit.

11.1 A posteriori error estimate in the regularized solution

In this section we formulate theorem for a posteriori error estimates $|\varepsilon-\varepsilon_h|$ and $|\mu-\mu_h|$ in the regularized solution (ε,μ) of the functional (138) which was derived in [4]. We define the scalar product $(\cdot,\cdot)_{L_2}$ as $(\cdot,\cdot)$, as well as define the norm $\|\cdot,\cdot\|_{L_2}$ as $\|\cdot,\cdot\|$.

Theorem

Assume the Tikhonov functional (138) is strongly convex in the neighbourhood $V_{(\gamma_1,\gamma_2)(\delta)}\left(\varepsilon^*,\mu^*\right)$. Let $(\varepsilon_h,\mu_h)\in W_h$ be a finite element approximations of the regularized solution $(\varepsilon,\mu)\in H^1(\Omega)$ on the finite element mesh K_h. Then there exists a constant D defined as $D=D\left(M_1,M_2\right)=const.>0$ and $z^*=(\varepsilon^*,\mu^*)$ defied as exact solution to our exact function such that

$$\left\|F'\left(z_1\right)-F'\left(z_2\right)\right\|\leq D\left\|z_1-z_2\right\|,\forall z_1,z_2\in V_1(z^*). \tag{159}$$

such that the following a posteriori error estimates hold

$$\begin{aligned}\|\varepsilon-\varepsilon_h\|&\leq\frac{D}{\alpha_1}C_I\left(||h\varepsilon_h||+\|[\varepsilon_h]\|\right)=\frac{2D}{\delta^{2\nu_1}}C_I\left(||h\varepsilon_h||+\|[\varepsilon_h]\|\right)\ \forall\varepsilon_h\in W_h,\\ \|\mu-\mu_h\|&\leq\frac{D}{\alpha_2}C_I\left(\|h\mu_h\|+\|[\mu_h]\|\right)=\frac{2D}{\delta^{2\nu_2}}C_I\left(\|h\mu_h\|+\|[\mu_h]\|\right)\ \forall\mu_h\in W_h.\end{aligned} \tag{160}$$

In the case when $\varepsilon_h,\mu_h\in V_h$ we have a posteriori error estimate

$$\begin{aligned}\|\varepsilon-\varepsilon_h\|&\leq\frac{D}{\alpha_1}C_I\left\|h\varepsilon_h\right\|_{L_2(\Omega)}=\frac{2D}{\delta^{2\nu_1}}C_I\left\|h\varepsilon_h\right\|_{L_2(\Omega)},\\ \|\mu-\mu_h\|&\leq\frac{D}{\alpha_2}C_I\left\|h\mu_h\right\|_{L_2(\Omega)}=\frac{D}{\delta^{2\nu_2}}C_I\left\|h\mu_h\right\|_{L_2(\Omega)}.\end{aligned}$$

Proof: For the proof see [4].

11.2 A posteriori error estimate for the Tikhonov functional

In next theorem we present a posteriori error estimate for the error in the Tikhonov functional (138) on the finite element mesh K_h which was derived in [4].

Theorem

Suppose that there exists minimizer $(\varepsilon,\mu)\in H^1(\Omega)$ of the Tikhonov functional (138) on the mesh K_h. Suppose also that there exists finite element approximation (ε_h,μ_h) of (ε,μ) of $F(\varepsilon,\mu)$ on the set W_h and mesh K_h with the mesh function h. Then the following approximate a posteriori error estimate for the error $e=|F(\varepsilon,\mu)-F(\varepsilon_h,\mu_h)|$ in the Tikhonov functional (138) holds

$$\begin{aligned}e=|F(\varepsilon,\mu)-F(\varepsilon_h,\mu_h)|\leq C_I(&\|F'_\varepsilon(\varepsilon_h,\mu_h)\|\left(||h\varepsilon_h||+\|[\varepsilon_h]\|\right)\\&+\left\|F'_\mu(\varepsilon_h,\mu_h)\right\|\left(||h\mu_h||+\|[\mu_h]\|\right)).\end{aligned} \tag{161}$$

In the case when the finite element approximation $z_h\in V_h$ we have following a posteriori error estimate

$$e=|F(\varepsilon,\mu)-F(\varepsilon_h,\mu_h)|\leq C_I\left(\|F'_\varepsilon(\varepsilon_h,\mu_h)\|\,||h\varepsilon_h||+\left\|F'_\mu(\varepsilon_h,\mu_h)\right\|\,||h\mu_h||\right). \tag{162}$$

Proof: For the proof see [4].

12 The Adaptive Algorithms

In our adaptive algorithms for the mesh refinement we apply the ideas of [17], Theorem 5.1 and the criterion of Remark 5.2 of [14]. In Algorithm 1 we use constant regularization parameters in the optimization procedure while in Algorithm 2 we choose these parameters iteratively. Functions $E_h\left(x,t,\varepsilon_h^m,\mu_h^m\right),\lambda_h\left(x,t,\varepsilon_h^m,\mu_h^m\right)$ are computed by solving the state and adjoint problems with $\varepsilon:=\varepsilon_h^m$ and $\mu:=\mu_h^m$. We iteratively update approximations ε_h^m and μ_h^m of the functions ε_h and μ_h, respectively , where m is the number of iteration in our optimization procedure.

Algorithm 1 : Adaptive Algorithm

Step 0. Choose the mesh K_h in Ω and time partition J of the time interval $(0,T)$. Start with the initial approximations $\varepsilon_h^0=\varepsilon_0$ and $\mu_h^0=\mu_0$ and which are located in the small neighbourhood of the exact solution (ε^*,μ^*), see [7] for details, then compute the sequences of ε_h^m,μ_h^m via the following steps:

Step 1. Compute the approximate solutions $E_h\left(x,t,\varepsilon_h^m,\mu_h^m\right)$ and $\lambda_h\left(x,t,\varepsilon_h^m,\mu_h^m\right)$ of state (103) and adjoint (145) problems on K_h and J, using domain decomposition FEM/FDM method described in section 9.4.

Step 2. Update the coefficient $\varepsilon_h:=\varepsilon_h^{m+1}$ and $\mu_h:=\mu_h^{m+1}$ on K_h and J using the conjugate gradient method:

$$\varepsilon_h^{m+1}(\mathbf{x}):=\varepsilon_h^m(\mathbf{x})+\alpha_1 d_1^m(\mathbf{x}),$$

$$\mu_h^{m+1}(\mathbf{x}):=\mu_h^m(\mathbf{x})+\alpha_2 d_2^m(\mathbf{x}),$$

where $\alpha_i, i=1,2,$ are step-sizes in the gradient update [49] and

$$d_1^m(x)=-g_1^m(\mathbf{x})+\beta_1^m d_1^{m-1}(\mathbf{x}),$$

$$d_2^m(x)=-g_2^m(\mathbf{x})+\beta_2^m d_2^{m-1}(\mathbf{x}),$$

with

$$\beta_1^m=\frac{||g_1^m(\mathbf{x})||^2}{||g_1^{m-1}(\mathbf{x})||^2},$$

$$\beta_2^m=\frac{||g_2^m(\mathbf{x})||^2}{||g_2^{m-1}(\mathbf{x})||^2},$$

where $d_1^0(x)=-g_1^0(\mathbf{x}),d_2^0(x)=-g_2^0(\mathbf{x})$.

Here, functions $g_1^m(\mathbf{x})$ and $g_2^m(\mathbf{x})$ are defined as :

$$\begin{aligned} g_1^m(\mathbf{x}) = &-\int_0^T \frac{\partial \lambda_h^m}{\partial t}(\mathbf{x},\,t)\cdot\frac{\partial E_h^m}{\partial t}(\mathbf{x},\,t)\,\mathrm{d}t \\ &+\xi\int_0^T \nabla\cdot E_h^m(\mathbf{x},\,t)\nabla\cdot\lambda_h^m(\mathbf{x},\,t)\,\mathrm{d}t+\gamma_1({\varepsilon_h}^m(\mathbf{x})-\varepsilon_0(\mathbf{x})), \end{aligned} \tag{163}$$

$$g_2^m(\mathbf{x}) = -\int_0^T \frac{1}{(\mu_h^m)^2} \, \nabla \times E_h^m(\mathbf{x}, \, t) \, \nabla \times \lambda_h^m \, (\mathbf{x}, \, t) \, \mathrm{d}t + \gamma_2(\mu_h^m(\mathbf{x}) - \mu_0(\mathbf{x})). \tag{164}$$

Step 3. Stop updating the coefficient ε_h^m and set $\varepsilon_h := \varepsilon_h^{m+1}$, $M := m + 1$ if either $||g_1^m||_{L_2(\Omega)} \leq \theta$ or norms $||\varepsilon_h^m||_{L_2(\Omega)}$ are stabilized. Here, θ is the tolerance in m updates of the gradient method.

Step 4. Stop updating the coefficient μ_h^m and set $\mu_h := \mu_h^{m+1}$, $M := m + 1$ if either $||g_2^m(\mathbf{x})||_{L_2(\Omega)} \leq \theta$ or norms $||\mu_h^m||_{L_2(\Omega)}$ are stabilized. Otherwise set $m := m + 1$ and go to step 1.

Step 5. Compute g_1^M and g_2^M via (165). Refine the mesh at all grid points $\mathbf{x}$ where

$$|g_1^M(\mathbf{x}) + g_2^M(\mathbf{x})| \geq \beta_1 \max_{\mathbf{x} \in \overline{\Omega}} |g_1^M(\mathbf{x}) + g_2^M(\mathbf{x})|.$$

Here the tolerance number $\beta_1 \in (0, \, 1)$ is chosen by the user.

Step 6. Construct a new mesh K_h in Ω and a new partition J_k of the time interval $(0, \, T)$. On J_k the new time step τ should be chosen in such a way that the CFL condition (134) is satisfied. Interpolate the initial guess ε_0 and μ_0 from the previous mesh to the new mesh. Next, return to step 1 at $m = 1$ and perform all above steps on the new mesh. Stop mesh refinements if norms defined in steps 3 and 4 either increase or stabilize, compared with the previous mesh.

Algorithm 2 : Iteratively regularized adaptive algorithm

Step 0. Choose the mesh K_h in Ω and time partition J of the time interval $(0, T)$. Start with the initial approximations $\varepsilon_h^0 = \varepsilon_0$ and $\mu_h^0 = \mu_0$ and which are located in the small neighbourhood of the exact solution (ε^*, μ^*), see [7] for details, then compute the sequences of ε_h^m, μ_h^m via the following steps:

Step 1. Compute the approximate solutions $E_h\left(x, t, \varepsilon_h^m, \mu_h^m\right)$ and $\lambda_h\left(x, t, \varepsilon_h^m, \mu_h^m\right)$ of state (103) and adjoint (145) problems on K_h and J, using domain decomposition FEM/FDM method described in section 9.4.

Step 2. Update the coefficient $\varepsilon_h := \varepsilon_h^{m+1}$ and $\mu_h := \mu_h^{m+1}$ on K_h and J using the conjugate gradient method:

$$\varepsilon_h^{m+1}(\mathbf{x}) := \varepsilon_h^m(\mathbf{x}) + \alpha_1 d_1^m(\mathbf{x}),$$

$$\mu_h^{m+1}(\mathbf{x}) := \mu_h^m(\mathbf{x}) + \alpha_2 d_2^m(\mathbf{x}),$$

where $\alpha_i, i = 1, 2$, are step-sizes in the gradient update [49] and

$$d_1^m(x) = -g_1^m(\mathbf{x}) + \beta_1^m d_1^{m-1}(\mathbf{x}),$$

$$d_2^m(x) = -g_2^m(\mathbf{x}) + \beta_2^m d_2^{m-1}(\mathbf{x}),$$

with

$$\beta_1^m = \frac{||g_1^m(\mathbf{x})||^2}{||g_1^{m-1}(\mathbf{x})||^2},$$
$$\beta_2^m = \frac{||g_2^m(\mathbf{x})||^2}{||g_2^{m-1}(\mathbf{x})||^2},$$

where $d_1^0(x) = -g_1^0(\mathbf{x}), d_2^0(x) = -g_2^0(\mathbf{x})$.

Here, functions $g_1^m(\mathbf{x})$ and $g_2^m(\mathbf{x})$ are defined as :

$$\begin{aligned} g_1^m(\mathbf{x}) = &-\int_0^T \frac{\partial \lambda_h^m}{\partial t}(\mathbf{x},\, t) \cdot \frac{\partial E_h^m}{\partial t}(\mathbf{x},\, t)\, \mathrm{d}t \\ &+ \xi \int_0^T \nabla \cdot E_h^m(\mathbf{x},\, t) \nabla \cdot \lambda_h^m(\mathbf{x},\, t)\, \mathrm{d}t + \gamma_1^m({\varepsilon_h}^m(\mathbf{x}) - \varepsilon_0(\mathbf{x})), \end{aligned} \tag{165}$$

$$g_2^m(\mathbf{x}) = -\int_0^T \frac{1}{(\mu_h^m)^2}\, \nabla \times E_h^m(\mathbf{x},\, t)\, \nabla \times \lambda_h^m\, (\mathbf{x},\, t)\, \mathrm{d}t + \gamma_2^m(\mu_h^m(\mathbf{x}) - \mu_0(\mathbf{x})). \tag{166}$$

where functions $E_h\left(x, t, \varepsilon_h^m, \mu_h^m\right), \lambda_h\left(x, t, \varepsilon_h^m, \mu_h^m\right)$ are computed by solving the state and adjoint problems with $\varepsilon := \varepsilon_h^m$ and $\mu := \mu_h^m$. We iteratively update approximations ε_h^m and μ_h^m of the functions ε_h and μ_h, respectively , where m is the number of iteration in our optimization procedure. Also here,γ_1^m and γ_2^m are iteratively chosen regularization parameters by rules of section 6.3.

Step 3. Stop updating the coefficient ε_h^m and set $\varepsilon_h := \varepsilon_h^{m+1}$, $M := m+1$ if either $||g_1^m||_{L_2(\Omega)} \leq \theta$ or norms $||\varepsilon_h^m||_{L_2(\Omega)}$ are stabilized. Here, θ is the tolerance in m updates of the gradient method.

Step 4. Stop updating the coefficient μ_h^m and set $\mu_h := \mu_h^{m+1}$, $M := m+1$ if either $||g_2^m(\mathbf{x})||_{L_2(\Omega)} \leq \theta$ or norms $||\mu_h^m||_{L_2(\Omega)}$ are stabilized. Otherwise set $m := m+1$ and go to step 1.

Step 5. Compute g_1^M and g_2^M via (165). Refine the mesh at all grid points $\mathbf{x}$ where

$$|g_1^M(\mathbf{x}) + g_2^M(\mathbf{x})| \geq \beta_1 \max_{\mathbf{x} \in \overline{\Omega}} |g_1^M(\mathbf{x}) + g_2^M(\mathbf{x})|.$$

Here the tolerance number $\beta_1 \in (0,\, 1)$ is chosen by the user.

Step 6. Construct a new mesh K_h in Ω and a new partition J_k of the time interval $(0,\, T)$. On J_k the new time step τ should be chosen in such a way that the CFL condition (134) is satisfied. Interpolate the initial guess ε_0 and μ_0 from the previous mesh to the new mesh. Next, return to step 1 at $m = 1$ and perform all above steps on the new mesh. Stop mesh refinements if norms defined in steps 3 and 4 either increase or stabilize, compared with the previous mesh.

In step 2 of these algorithms the parameters α_1 and α_2 can be computed by a line search procedure, see, e.g. [49].

13 Numerical Studies

In this section we discuss the numerical simulation of the reconstruction of two unknown functions, ε and μ, inside a domain Ω_{FEM} using the adaptive algorithm presented in section 12 as well as using hybrid FEM/FDM of section 9.4. We set values of two functions $\varepsilon = \mu = 1$ inside of Ω_{FDM}.

Our goal is to reconstruct dielectric permittivity and magnetic permeability, ε and μ. In our numerical experience we work with relative dielectric permittivity and magnetic permeability, ε_r and μ_r, which are dimensionless and are defined as $\varepsilon = \varepsilon_r\varepsilon_0,\ \mu = \mu_r\mu_0$. Here ε_r and μ_r can take values in the following intervals:

$$\varepsilon_r(target) \in (1, 13),$$
$$\mu_r(target) \in (1, 2).$$

In all of our computations we initialize only one component E_2 of the electrical field $E(E_1, E_2, E_3)$ as the boundary condition in (96)–(99) on S_T [33]. Initial conditions are considered as zero. We used modified version of the stabilized domain decomposition method of [13] which was implemented using the software packages WavES [60] with two functions ε_r and μ_r.

13.1 Specific aims

We have following aims in our computations.

1 ***Simulation of the forward problem*** (109)**-**(114) ***in order to obtain backscattering data.*** Before reconstruction of two unknown relative coefficients, ε_r and μ_r, from backscattering data simultaneously, we need to simulate these data. To do this we generate geometry with exact values of ε_r and μ_r and compute the forward problem with the known exact values in them. Then we save computed solution $E(\mathbf{x}, t)$ at the observation points at S_T at the backscattered side $\partial\Omega_1$. We work only with these backscattering data to solve our CIP.

2 ***Reconstruction of 12 targets.***

The second goal is to reconstruct targets in a finite element domain Ω_{FEM} using backscattering data of item 1, see Figure 4.

3 ***Checking stability with respect to the frequency.***

The third aim is to check if computations of our IP are stable with respect to the wave length. The wave length can be computed by $\lambda = \frac{2\pi c}{\omega}$, where $c = \frac{1}{\sqrt{\varepsilon\mu}}$ is the wave speed and ω is frequency in the initialized plane wave. We choose frequency in the interval $[20, 50]$, $\omega \in [20, 50]$, and then check how the Maxwell's system response to our fixed parameters when frequency is changing.

4 ***Checking stability with respect to the noise level.***

The fourth aim is to find the interval for the noise level which makes our computations stable. This means that we search for a noise interval on what we can get stable reconstruction of ε and μ such that this noise does not destroys the reconstruction.

5 ***Checking the impact of the choice of the regularization parameters***

Finally, we reconstruct ε_r and μ_r using our 2 adaptive algorithm of section 12. In Algorithm 1 we consider regularization parameters as fixed numbers. Algorithm 2 is based on iteratively computed regularization parameters as described in section 12. Then we compare results obtained via these both algorithms.

To achieve our goals, we split the entire domain into two geometries, Ω_{FEM} and Ω_{FDM}, such that $\Omega = \Omega_{FEM} \cup \Omega_{FDM}$, see Figure 4. Next, we consider dimensionless spatial variables $x' = x/(1m)$ to get the dimensionless computational domain form of the domain Ω_{FEM}

$$\Omega_{FEM} = \{\mathbf{x} = (x_1, x_2, x_3) \in (-3.2, 3.2) \times (-0.6, 0.6) \times (-0.3, 0.3)\}$$

The dimensionless size of our computational domain Ω for the forward problem is

$$\Omega = \{\mathbf{x} = (x_1, x_2, x_3) \in (-3.4, 3.4) \times (-0.8, 0.8) \times (-0.4, 0.4)\}.$$

The space mesh in Ω_{FEM} and Ω_{FDM} consists of tetrahedra and cubes, respectively. We take mesh size, $h = 0.1$, in the optimization algorithm in our geometries in the hybrid FEM/FDM method, as well as in the overlapping regions between FEM and FDM domains. While doing computational tests, the penalty factor s in the explicit FEM scheme (123) is equal to one in Ω_{FEM}. We note that due to using the domain decomposition method as well as conditions (94)–(95), the Maxwell's system in Ω_{FDM} transforms to the wave equation

$$\begin{aligned}
\frac{\partial^2 E}{\partial t^2} - \triangle E &= 0, \text{ in } \Omega_{FDM} \times (0, T), \\
E(x, 0) = f_0(x), \quad E_t(x, 0) &= 0 \text{ in } \Omega_{FDM}, \\
\partial_n E(x, t) &= (0, f(t), 0), \text{ on } \partial\Omega_1 \times (0, t_1], \\
\partial_n E(x, t) &= -\partial_t E(x, t), \text{ on } \partial\Omega_1 \times (t_1, T), \\
\partial_n E(x, t) &= -\partial_t E(x, t), \text{ on } \partial\Omega_2 \times (0, T), \\
\partial_n E(x, t) &= 0, \text{ on } \partial\Omega_3 \times (0, T).
\end{aligned} \tag{167}$$

We initialize only the second component of the electric field E_2 as a plane wave $f(t)$ in Ω_{FDM} in time $T = [0, 3]$ such that

$$f(t) = \begin{cases} sin(\omega t), & \text{if } t \in (0, \frac{2\pi}{\omega}), \\ 0, & \text{if } t > \frac{2\pi}{\omega}, \end{cases} \tag{168}$$

while two other components of electric field E_2, E_3 are initialized as zero. This means that for the solution of forward problem via hybrid FEM/FDM method, it is required to solve equation (167) in Ω_{FDM}, and in Ω_{FEM} we have to solve

$$\begin{aligned} \varepsilon \frac{\partial^2 E}{\partial t^2} + \nabla \times (\mu^{-1} \nabla \times E) - s \nabla (\nabla \cdot (\varepsilon E)) = 0, \text{ in } \Omega_{\text{FEM}}, \\ E(\mathbf{x}, 0) = 0, \quad E_t(\mathbf{x}, 0) = 0 \text{ in } \Omega_{\text{FEM}}, \\ E(\mathbf{x}, t)|_{\partial \Omega_{\text{FEM}}} = E(\mathbf{x}, t)|_{\partial \Omega_{\text{FDM}_I}}. \end{aligned}$$

Here, $\partial \Omega_{\text{FDM}_I}$ is internal boundary of the domain Ω_{FDM}, and $\partial \Omega_{\text{FEM}}$ is the boundary of the domain Ω_{FEM}. Similarly, in Ω_{FDM} the adjoint problem transforms to the wave equation

$$\begin{aligned} \frac{\partial^2 \lambda}{\partial t^2} - \triangle \lambda = -(E - \tilde{E}) z_\delta, \text{ in } S_T, \\ \lambda(x, T) = 0, \quad \lambda_t(x, T) = 0 \text{ in } \Omega, \\ \partial_n \lambda(x, t) = 0 \text{ on } S_3. \end{aligned} \tag{169}$$

When solving the adjoint problem via hybrid FEM/FDM method in Ω_{FDM} we solve the problem (169) and in Ω_{FEM} we have to solve

$$\begin{aligned} \varepsilon \frac{\partial^2 \lambda}{\partial t^2} + \nabla \times (\mu^{-1} \nabla \times \lambda) - s \nabla (\nabla \cdot (\varepsilon \lambda)) = 0, \text{ in } \Omega_{\text{FEM}}, \\ \lambda(\mathbf{x}, T) = 0, \quad \lambda_t(\mathbf{x}, T) = 0 \text{ in } \Omega_{\text{FEM}}, \\ \lambda(\mathbf{x}, t)|_{\partial \Omega_{\text{FEM}}} = \lambda(\mathbf{x}, t)|_{\partial \Omega_{\text{FDM}_I}}. \end{aligned}$$

We define exact functions to be reconstructed $\varepsilon_r = 12$ and $\mu_r = 2$ inside of all inclusions, see Figure 5, and $\varepsilon_r = \mu_r = 1$ at all other points of computational domain Ω_{FEM}. We choose in our computations the time step $\tau = 0.006$ which satisfies the CFL condition [27] and perform computation in time interval $[0, 3]$.

13.2 Generation of backscattered data

To generate backscattered data $\tilde{E}$ at the observation points at S_T in (137) we solve the forward problem (109)–(114), with function $f(t)$ given by (168) in the time interval $T = [0, 3]$ with the exact values of the parameters $\varepsilon_r = 12, \mu_r = 2$ inside inclusions of the Figure 5, and $\varepsilon_r = \mu_r = 1$ everywhere else in Ω, see Figures 7–12. We avoid the variational crime in our tests since the data were generated on a locally refined mesh where inclusions were presented. However, we use our optimization algorithm on a different structured mesh with the mesh size $h = 0.1$ at the overlapping nodes between FEM/FDM domain, see section

9.4 for description of these nodes.

We present the isosurfaces of the simulated exact solution of the initialized components $E_1(\mathbf{x},t)$ and $E_2(\mathbf{x},t)$ of the electric field in the forward problem (96)–(99) with $\omega \in [20,50]$ at different times in Figures 14–18. We observe at these figures the backscattering wave field of the two components, $E_1(\mathbf{x},t)$ and $E_2(\mathbf{x},t)$.

13.3 Reconstructions

To start our optimization algorithm we need to have initial guess values $\varepsilon_0,\ \mu_0$ in all points of Ω. We let $\varepsilon_0 = 1, \mu_0 = 1$ in all points of Ω. This initial guesses are chosen similarly to [14, 15] since our computations show that these initial guesses reconstruct both functions ε_r, μ_r good , see also [17] for similar observations. Such choice of the initial guess values means that we start the conjugate gradient method from the homogeneous domain.

To do computations the minimal and maximal values of the functions $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$ belong to the following sets of desired acceptable parameters

$$
\begin{aligned}
C_{\varepsilon_r} &\in \{\varepsilon_r \in C(\bar{\Omega}) | 1 \leq \varepsilon_r(\mathbf{x}) \leq 13\}, \\
C_{\mu_r} &\in \{\mu_r \in C(\bar{\Omega}) | 1 \leq \mu_r(\mathbf{x}) \leq 2\}.
\end{aligned}
$$

The solution of the inverse problem requires to be regularized since different coefficients can correspond to similar wave reflection data on $\partial_1\Omega$. To do so, we need two different regularization parameters $\gamma_1 = 0.01$ and $\gamma_2 = 0.7$ in (138) for the computation of inverse problem. Chosen values for these regularization parameters are based on our computational experience of our previous work [15] on this topic. We choose the regularization parameters in a way that the values of the regularization parameters give the smallest reconstruction relative L_2 errors, e_ε and e_μ where $e_\varepsilon = \frac{\|\varepsilon - \varepsilon_h\|}{\|\varepsilon\|}$ and $e_\mu = \frac{\|\mu - \mu_h\|}{\|\mu\|}$. Here ε, μ are the exact values of the coefficients and ε_h, μ_h are computed ones. Now, we consider regularization parameters as fixed constants but in our final step of our computations, regularization parameters are chosen iteratively. The tolerance θ in our adaptive algorithm is set to $\theta = 10^{-6}$. To get images of Figures 19–39, we assume that the function $\varepsilon^n(\mathbf{x})$ and $\mu^{\mathrm{l}}(\mathbf{x})$ are our reconstructions gained by the adaptive algorithm where n and l are number of iterations in conjugate gradient method when we have stopped to compute $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$. Then we let

$$
\widetilde{\varepsilon}^n(x) = \begin{cases} \varepsilon^n(x) & \text{if } \varepsilon^n(x) > 0.63 \max\limits_{\Omega} \varepsilon^n(x), \\ 1, & \text{otherwise.} \end{cases}
$$

and

$$
\widetilde{\mu}^l(x) = \begin{cases} \mu^l(x) & \text{if } \mu^l(x) > 0.5 \max\limits_{\Omega} \mu^l(x), \\ 1, & \text{otherwise.} \end{cases}
$$

The simultaneous reconstruction of ε_r and μ_r for the frequency $\omega \in [20, 50]$ and for different level of additive noise is presented in Figures 19–39.

13.4 Stability results with respect to the frequency

To achieve our second goal, checking stability of our system with respect to ω, we need to generate the backscattering data for various frequencies from interval $[20, 50]$ with the additive noise as well as with parameters defined in our computations of Inverse problem which are fixed numbers. We introduce additive noise to our simulated data as

$$E_{obs} = E_{obs}(1 + p\alpha), \tag{170}$$

where E_{obs} are computed values of $E(\mathbf{x}, t)$ with known values of ε_r and μ_r via hybrid FEM/FDM method, $\alpha \in [-1, 1]$ is randomly chosen constants and p is the noise level.
We consider the following test cases for the generation of the backscattering data:

i) Fixed additive noise 3%, $p = 0.03$ in (170) with frequency $\omega \in [20, 50]$;

ii) Fixed additive noise 7%, $p = 0.07$ in (170) with frequency $\omega \in [20, 50]$.

Using a posteriori error estimates of section 11.2, we expect to get the better reconstruction of our coefficients and consequently improved images of reconstructed functions after refinement in our finite element domain. Figures 19–39 show reconstruction results of $\varepsilon_r,\ \mu_r$ obtained via the adaptive algorithm 1.

Numerical Results: Figure 19 presents images of computed coefficients when noise is low and regularization parameters are fixed constants. After the first iteration in conjugate gradient method, none of the coefficients reach their maximum values considered in (94)–(95).

In next step, see Figure 20, still we do not have any refinement in the geometry at the last iteration in the conjugate gradient method. This Figure shows the reconstructed functions, $\varepsilon_r,\ \mu_r$, such that $\max\limits_{\Omega_{FEM}} \varepsilon_r = 13$, $\max\limits_{\Omega_{FEM}} \mu_r = 2.06$. Figure 23 presents the final result of our adaptive algorithm 1 on a 5 times locally refined mesh in Ω_{FEM}.

Figure 24 and 32 show that the computed images of the reconstructed function ε_r on the coarse mesh in Ω_{FEM} are not good in almost half of the finite element domain while function μ_r is reconstructed properly. These computations are done with the noise level $p = 0.03$ and $p = 0.07$ in (170) and for the frequency $\omega = 30$ in (168).

Figure 24 and 32 show that we do not have good reconstruction of ε_r after the first iteration in the conjugate gradient method on the coarse mesh. Meanwhile, Figures 25–32 present that the image of function μ_r on the 5 times adaptively refined mesh is reconstructed properly with $\max\limits_{\Omega_{FEM}} \mu_r = 1.567$. Figure 26 presents reconstruction of functions ε_r and μ_r for the frequency $\omega = 40$ in (168) and additive noise $p = 0.03$ in (170) on a coarse mesh. We observe good reconstruction of locations of all inclusions and values for ε_r and μ_r. However, shapes in z direction should be improved.

Figure 27 shows that on the 5 times adaptively refined mesh in Ω_{FEM} we reconstruct both functions $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$ properly. Comparing Figures 26 and 27, we can see that adaptive algorithm can significantly improve shapes of our inclusions in all x_1, x_2, x_3 directions. Figure 28 presents reconstruction of functions ε_r and μ_r on a coarse mesh such that $\max\limits_{\Omega_{FEM}} \varepsilon_r = 13$, $\max\limits_{\Omega_{FEM}} \mu_r = 1.65$. These computations were performed with frequency $\omega = 50$ and noise level $p = 0.03$.

13.5 Stability results with respect to the noise level

Analyzing the results obtained in the previous section, we can claim that the appropriate interval of frequencies where we obtain stable reconstructions is interval $\omega \in [40, 50]$. Now we choose a fixed frequency $\omega = 40$ in (168) and check how our system responds to the different noise level. In this section we check stability of reconstructions functions ε_r and μ_r for the frequency $\omega = 40$ and for the noise level $12\%, 17\%, 25\%$ and 35% in backscattering data. We have already tested in the previous section smaller additive noise such that $p = 0.03, 0.07$ in (170). The obtained results are shown in Figures 40–47.

Numerical Results: We fix frequency $\omega = 40$ in (168) and check behaviour of the reconstructed functions, ε_r and μ_r, for a higher additive noise 12% and 17% such that $p = 0.12$ and $p = 0.17$ in (170). Using figures 41 and 43 we observe that the maximum of values for ε_r and μ_r is not small.

Figures 40–41 show that the functions ε_r and μ_r have a small error in the computed maximum values after the first iteration in the conjugate gradient method on a coarse mesh. The obtained results are compatible with our anticipation based on the theoretical observations.

Figure 43 presents that we are able to reconstruct functions, ε_r and μ_r, with a higher noise 17% with $\max\limits_{\Omega_{FEM}} \varepsilon_r = 12.95$, $\max\limits_{\Omega_{FEM}} \mu_r = 1.56$. In the next step we increase the additive noise from 17% to 25% using frequency $\omega = 40$ to check reconstructions of functions ε_r and μ_r on a coarse mesh. In these computations both functions reach their maximum values. Figure 44 presents the obtained results.

We continue our investigations through increasing the additive noise. We let additive noise $p = 0.35$ in (170) and frequency $\omega = 40$ in (168). Figure 46 shows obtained reconstructions on a coarse mesh.

Figure 47 presents that even with a higher additive noise, 35%, we can reconstruct both functions ε_r and μ_r for the frequency $\omega = 40$ on the 5 times adaptively refined mesh. This figure shows that we can not see clearly reconstructions for μ_r since about $\frac{1}{6}$ part of the finite element domain, top left side of the Ω_{FEM}, is not reconstructed properly.

13.6 Application of different techniques for the choice of regularization parameter

In the computations of this section we apply adaptive algorithm 2 with iterative choice of the regularization parameters in the Tikhonov functional (138). We consider two methods for choosing iterative regularization parameters, γ_1, γ_2, in the Tikhonov functional (138). Up to now, we have presented results based on the optimal choice of the regularization parameters, γ_1, γ_2, which were chosen in a computational efficient way. Now we choose our regularization parameters iteratively based on the theory given in section 6 specifically rules given by (89) and (91). Using rule 2 given by the (91) we test our adaptive algorithms choosing the regularization parameter in the Tikhonov functional without taking into account the noise level. Then we compare the results obtained by considering the iterative regularization parameters, γ_1, γ_2, with the results obtained by using fixed regularization parameters. Finally, we fix frequency $\omega = 40$ in (168) and test our adaptive algorithms for different additive noise level in all methods. The obtained results of our computations are given in Figures 50–57.

Numerical Results: In Figure 50, we want to check how choice of regularization parameters γ_1, γ_2, affects the obtained reconstruction of functions, ε_r and μ_r. Our computations show that for the frequency $\omega = 40$ in (168) and additive noise $p = 0.17$ the reconstruction of ε_r is not sensitive to the choice of the method for computation of the regularization parameters via rule 2 given by (89). Figure 51 presents that the reconstruction of μ_r is more sensitive to chosen method for computation of regularization parameters. One should note that although both figures look similar to each other, the iteratively computed regularization parameters γ_1, γ_2 via rule 1 given by (89) affects the maximum value of the reconstructed function μ_r, see Table 3. We can conclude that this method has enormous impact on the reconstruction of function μ_r in Ω_{FEM}.

Figure 57 shows that the iteratively computed regularization parameters, $\gamma_1,\ \gamma_2$, using the rule 1 given by (89) in the adaptive algorithms helps us to maintain the proper reconstruction of μ_r while the additive noise in data is high and $p = 0.35$ in (170), see Table 3.

Now we test the adaptive algorithm 2 with iteratively computed regularization parameters γ_1, γ_2 using the rule 2 given by (91). Figure 60 shows that using Algorithm 2 for the frequency $\omega = 40$ with $p = 0.17$ in data in (170) the reconstruction of ε_r is not too sensitive to the iteratively computed regularization parameters via rule 2 given by (91).

Figure 61 presents that reconstruction of μ_r is more sensitive to the iterative choice of the regularization parameters γ_1, γ_2, computed via rule (91). One should note that although both figures look similar to each other, when applying rule 2 given by (91) the maximum value of the reconstructed function μ_r is higher, see Table 3. We can conclude that the choice of regularization parameters has enormous impact on the reconstruction of μ_r in Ω_{FEM}.

Figure 64 shows that for the frequency $\omega = 40$ and $p = 0.35$ in data the reconstruction of ε_r is not sensitive to the method for the computation of the regularization parameters. Figure 65 shows that using the iteratively computed regularization parameters, γ_1 and γ_2, via the rule 2 given by (91) helps us to maintain the proper reconstruction of function μ_r with a high additive noise $p = 0.35$.

14 Discussion and Conclusion

In this work we have used time-dependent backscattering data for reconstruction of both coefficients, $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$. Our goal was to improve reconstructions of ε_r and μ_r obtained by optimization approach in [15] using iteratively regularized adaptive finite element method. In the theoretical part we presented different strategies for the choice of regularization parameters in the Tikhonov functional (138). We described main framework for derivation of a posteriori error estimates for CIPs and formulated two adaptive algorithms. In the numerical part of this work we tested adaptive algorithm 1 with two different noise levels, 3% and 7%, for different frequencies ω on the interval $[20, 50]$. We also performed tests with fixed frequency $\omega = 40$ in (168) and observed simultaneously reconstruction of both coefficients, $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$, with higher noise levels 12%, 17%, 25% and 35%. We tested adaptive algorithm 2 with iterative choice of regularization parameters, γ_1 and γ_2, in the Tikhonov functional (138). First part of our numerical simulations is done using adaptive algorithm 1 when regularization parameters in the Tikhonov functional is chosen optimally based on experimental results. The second part is done using adaptive algorithm 2 with iterative choice of regularization parameters computed via (89) and (91). We have used two different rules in this algorithm for choosing regularization parameters. An important aspect of this part is that we are able to reconstruct simultaneously both coefficients, $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$, with high noise level such as 35 % and we have significant improvement in shapes and locations of all inclusions to be reconstructed. We conclude that we can get significant improvement in obtained images using adaptive algorithms with different values for computation of regularization parameters in the Tikhonov functional. So we can claim that iteratively regularized adaptive FEM is powerful tool for the reconstruction of both functions, $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$, and their locations and shapes accurately.

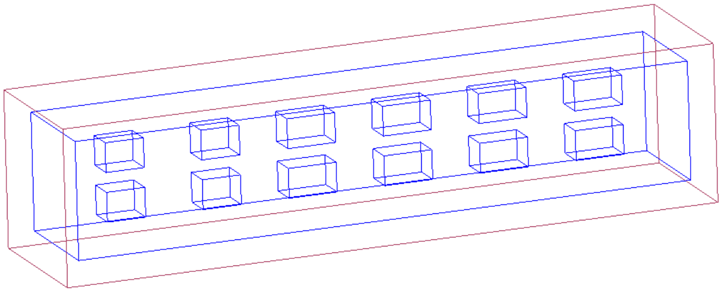


Figure 3: *Hybrid domain,* $\Omega = \Omega_{FEM} \cup \Omega_{FDM}$.

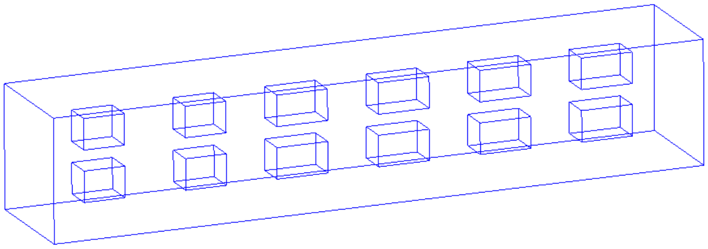


Figure 4: *The finite element domain* Ω_{FEM} *with inclusions to be reconstructed.*

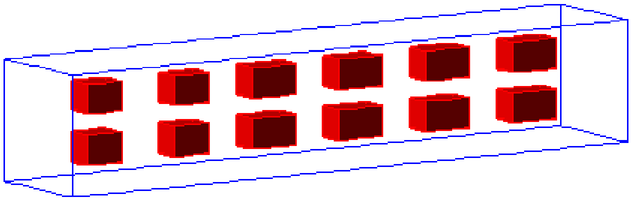
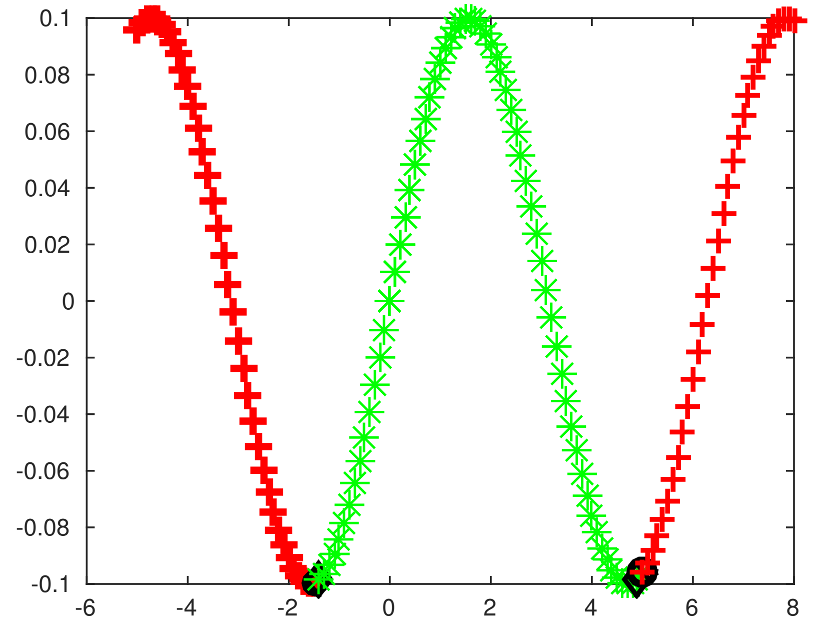
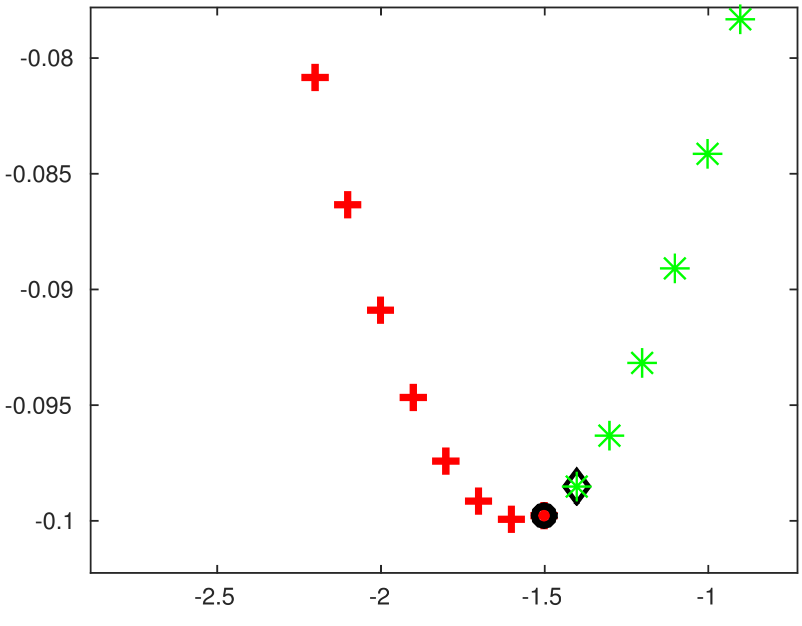


Figure 5: *The FEM domain* Ω_{FEM} *with inclusions, the exact values of functions* $\varepsilon_r(\mathbf{x})$ *and* $\mu_r(\mathbf{x})$ *are:* $\varepsilon_r(\mathbf{x}) = 12, \mu_r(\mathbf{x}) = 2$ *inside the twelve small inclusions, and* $\varepsilon_r(\mathbf{x})$ *and* $\mu_r(\mathbf{x})$ *are known everywhere else in* Ω_{FEM}.

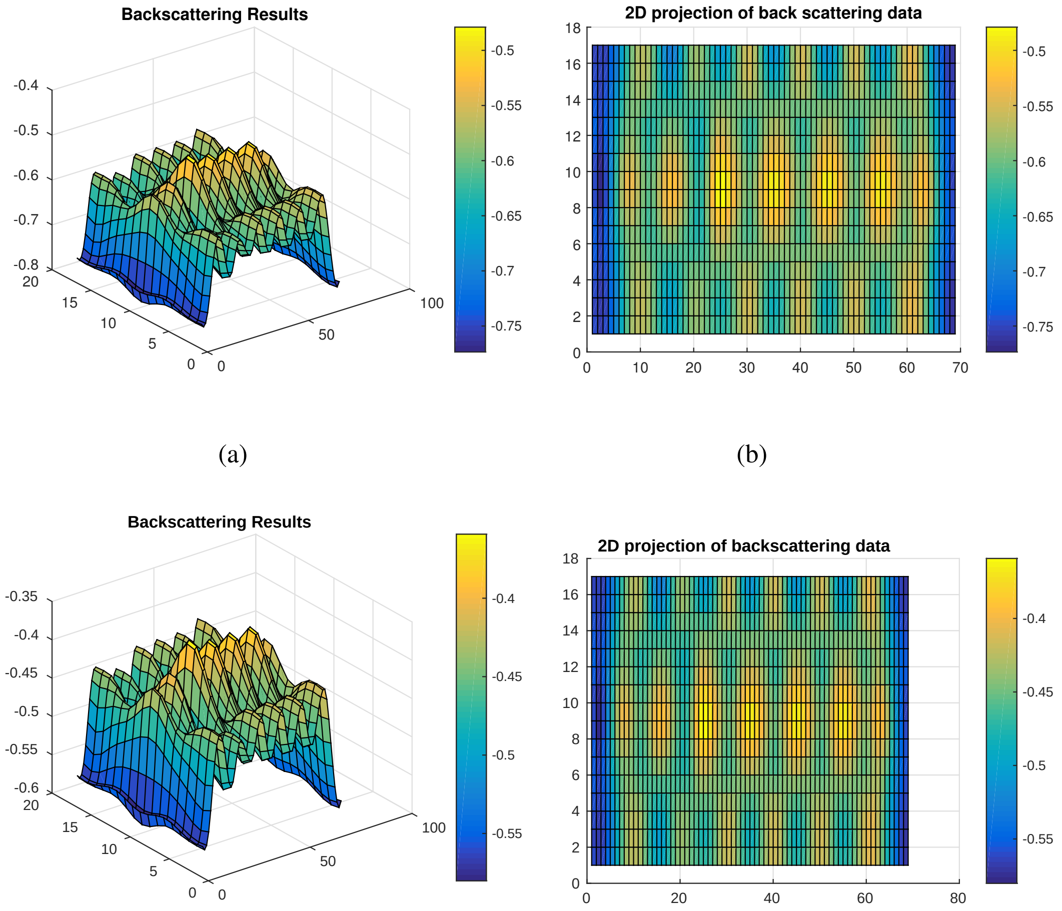


(a) Hybrid FEM/FDM.



(b) Hybrid FEM/FDM , zoomed view at the common boundary.

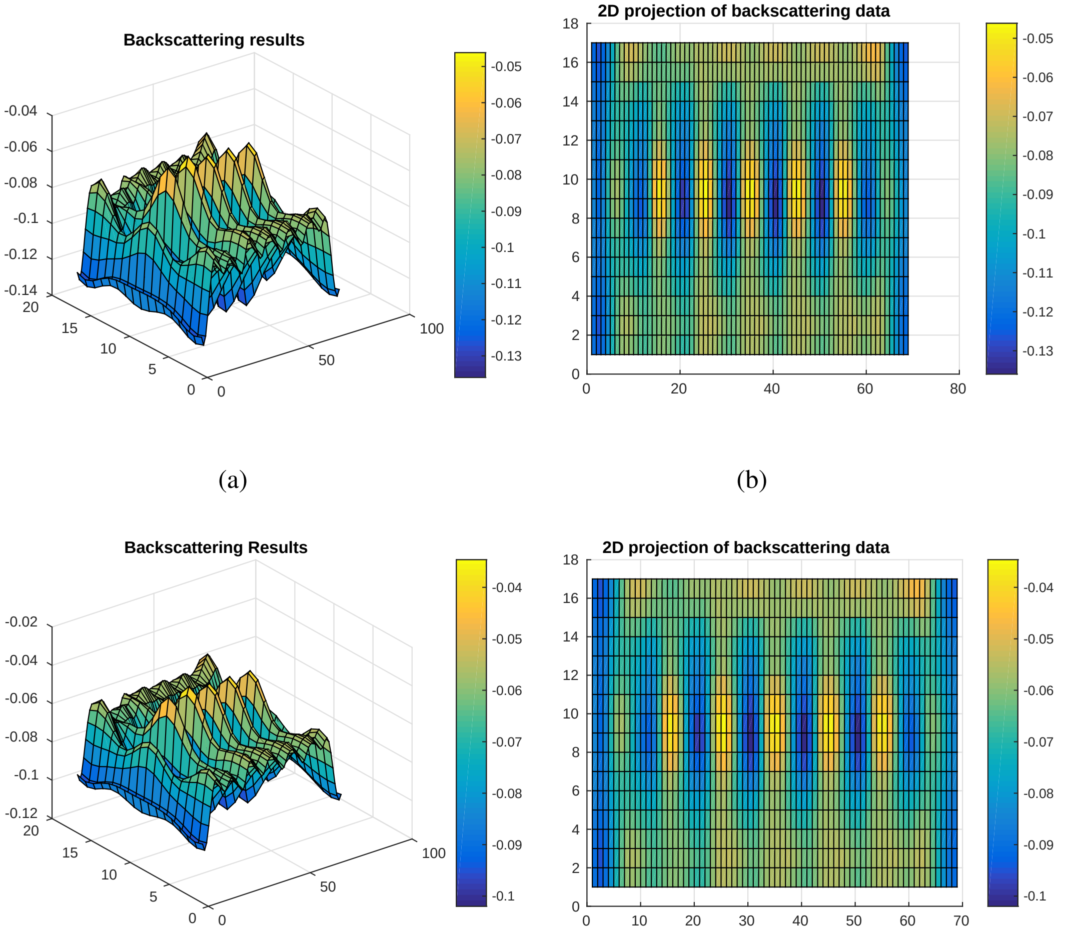
Figure 6: *Domain decomposition: hybrid FEM/FDM.*



(a) (b)

(c) (d)

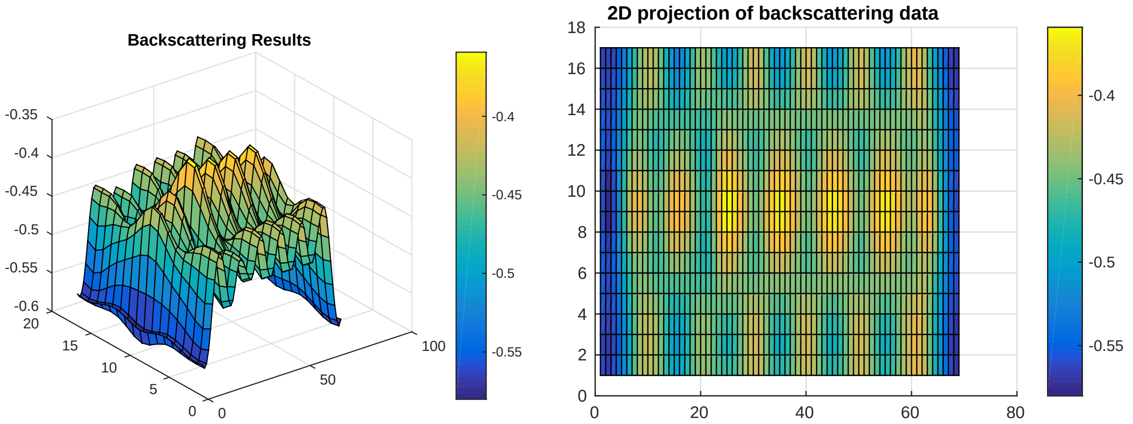
Figure 7: *Backscattering data of the one component of the electric field* $E(\mathbf{x}, t)$, $E_2(x, t)$, *frequency* $\omega = 20$ *in* (168)*:* ***(a)*** *we show backscattered data with 5% of additive noise;* ***(b)*** *data of (a) in xy projection;* ***(c)*** *backscattered data with 40% of additive noise;* ***(d)*** *data of (c) in xy projection.*



(a) (b)

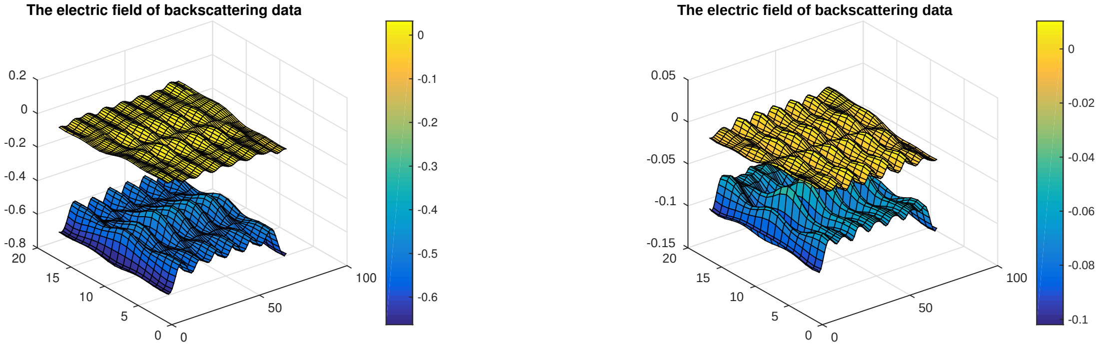
(c) (d)

Figure 8: *Backscattering data of the one component of the electric field* $E(\mathbf{x},t)$, $E_2(x,t)$, *frequency* $\omega = 40$ *in* (168)*:* ***(a)*** *we show backscattered data with 5% of additive noise;* ***(b)*** *data of (a) in xy projection;* ***(c)*** *backscattered data with 40% of additive noise;* ***(d)*** *data of (c) in xy projection.*



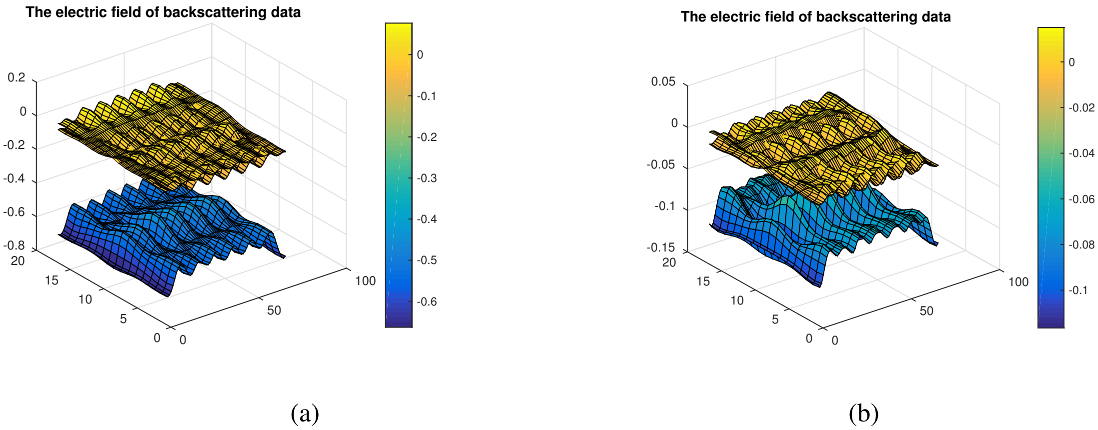
(a) (b)

Figure 9: ***(a)*** *Backscattering data of the component $E_2(\mathbf{x}, t)$, frequency $\omega = 20$ in* (168) *with 5% of additive noise in data:* ***(b)*** *Backscattering data of a) in xy projection.*



(a) (b)

Figure 10: *Comparison between computed E_2 (below) and E_1 (on top) components of the backscattered electric field with 20% additive noise:* ***(a)*** *Frequency $\omega = 20$ in* (168) ***(b)*** *Frequency $\omega = 40$ in* (168).



(a) (b)

Figure 11: *Comparison between computed E_2 (below) and E_1 and E_3 (on top) components of the backscattering electric field with 20% additive noise in data: **(a)** Frequency $\omega = 20$ in (168). **(b)** Frequency $\omega = 40$ in (168).*

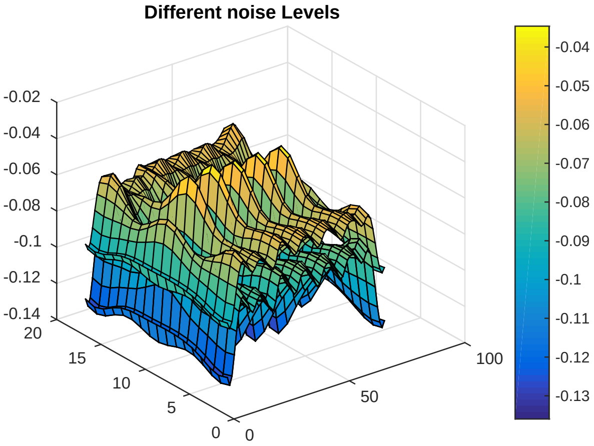
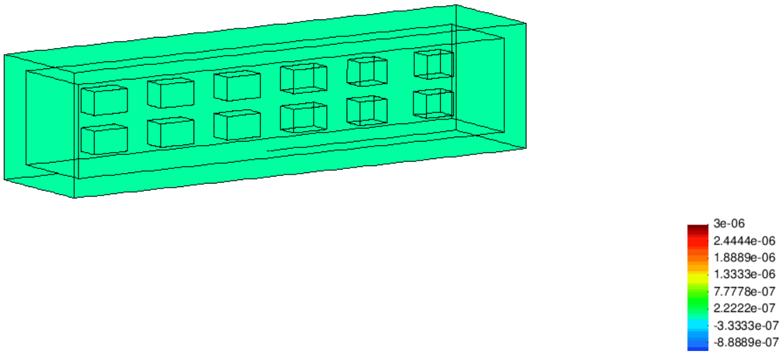
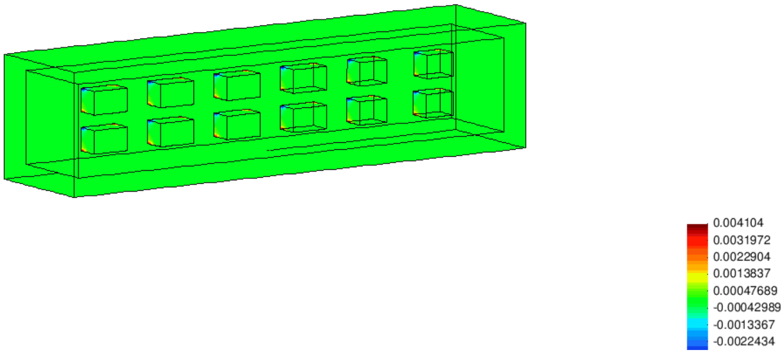


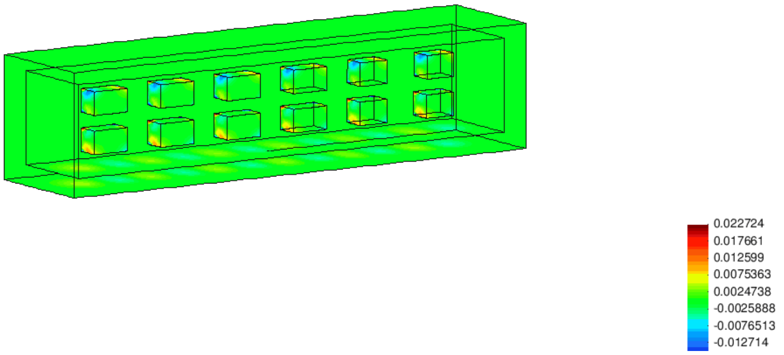
Figure 12: *Comparison of backscattering data of the component $E_2(\mathbf{x}, t)$ for the frequency $\omega = 40$ in (168) with 5% and 40% additive noise.*



(a) t=0.3 s.

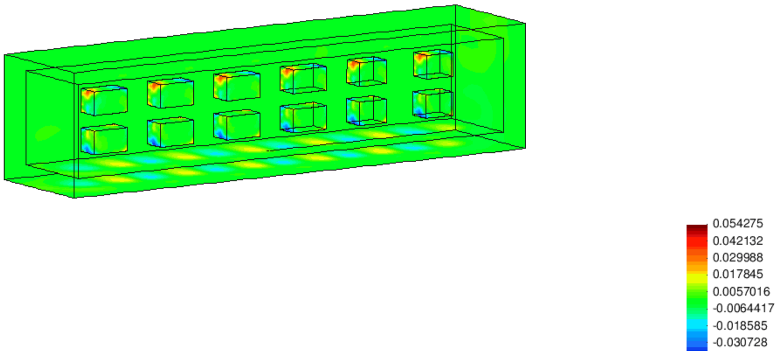


(b) t=0.6 s.

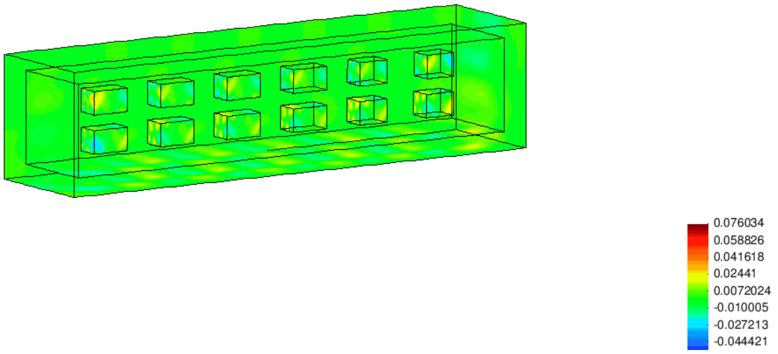


(c) t=0.9 s.

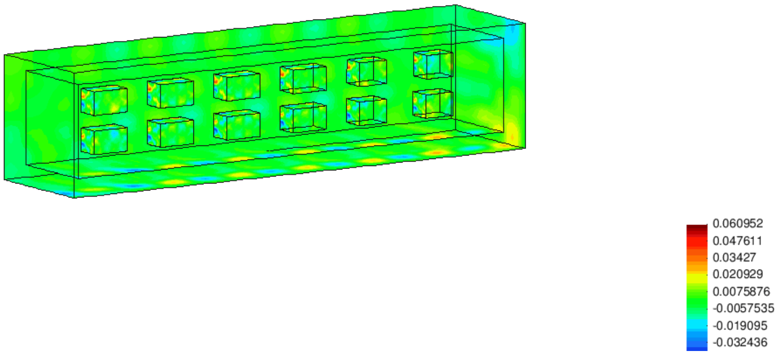
Figure 13: *Isosurface of the simulated exact solution of the component $E_1(x,t)$ for the forward problem at different times. The snapshot times are:* ***(a)*** *0.3 s,* ***(b)*** *0.6 s and* ***(c)*** *0.9 s, respectively.*



(a) t=1.2 s.

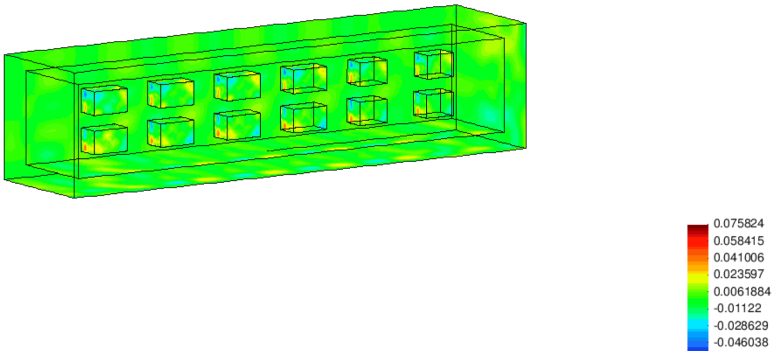


(b) t=1.5 s.

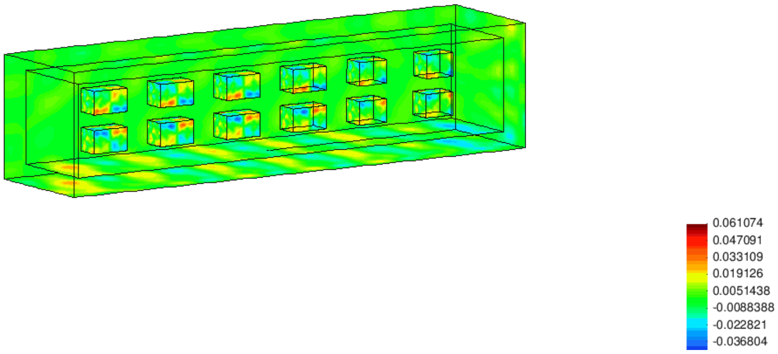


(c) t=1.8.

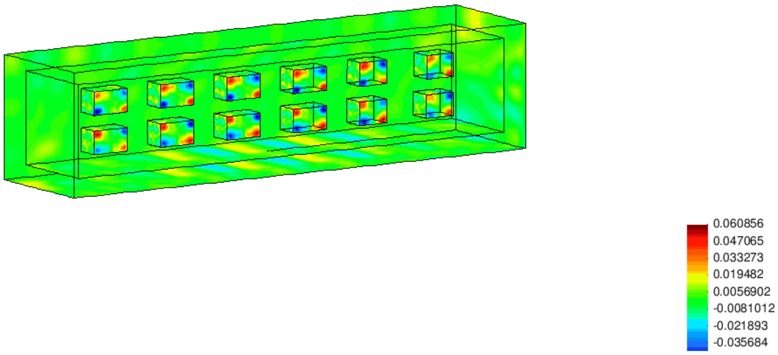
Figure 14: *Isosurface of the simulated exact solution of the component* $E_1(x,t)$ *for the forward problem at different times. The snapshot times are:* ***(a)*** *1.2 s,* ***(b)*** *1.5 s and* ***(c)*** *1.8 s, respectively.*



(a) t= 2.1 s.

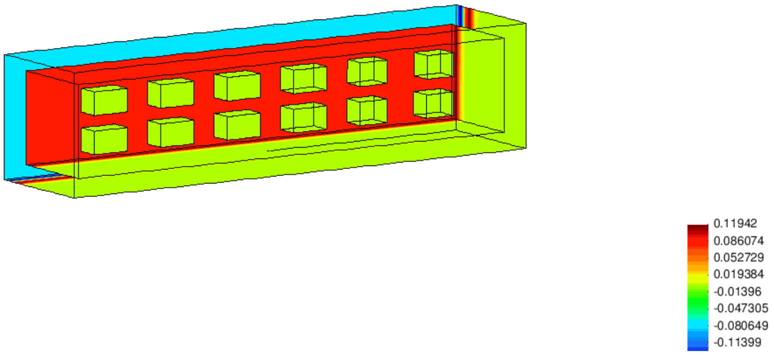


(b) t=2.4 s.

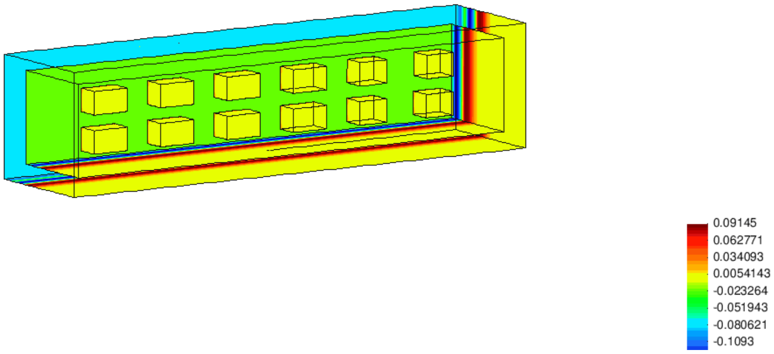


(c) t=2.7 s.

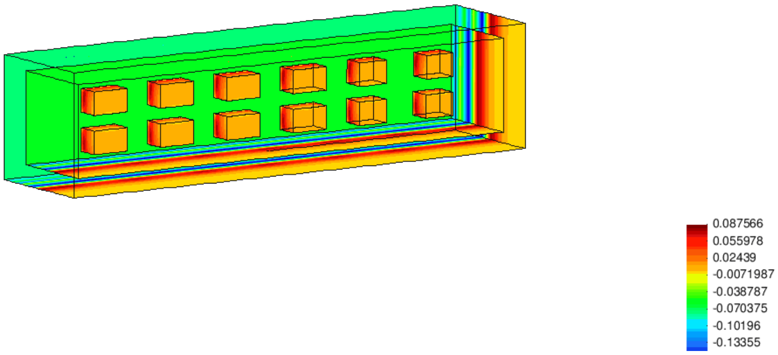
Figure 15: *Isosurface of the simulated exact solution of the component* $E_1(x,t)$ *for the forward problem at different times. The snapshot times are:* ***(a)*** *2.1 s,* ***(b)*** *2.4 s and* ***(c)*** *2.7 s, respectively.*



(a) t=0.3 s.

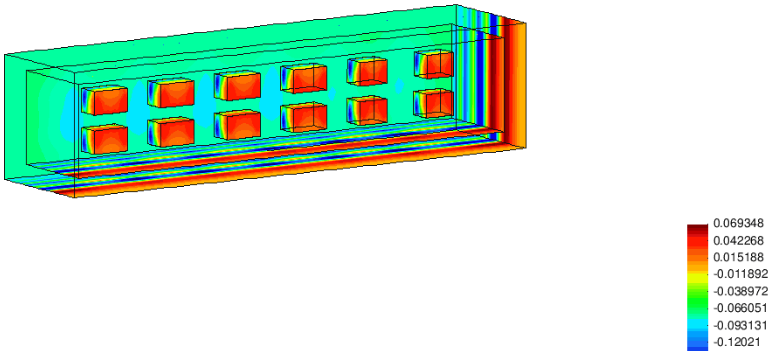


(b) t=0.6 s.

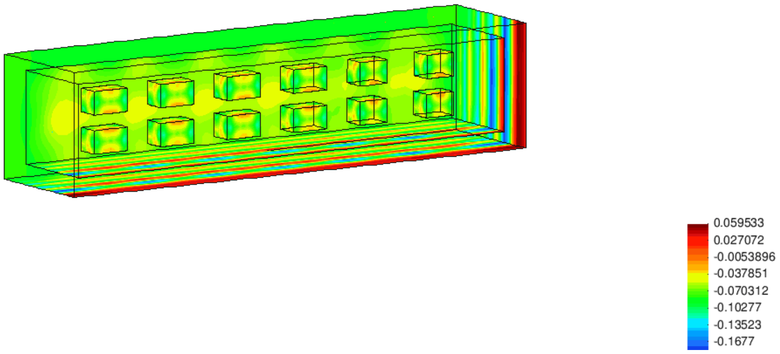


(c) t=0.9 s.

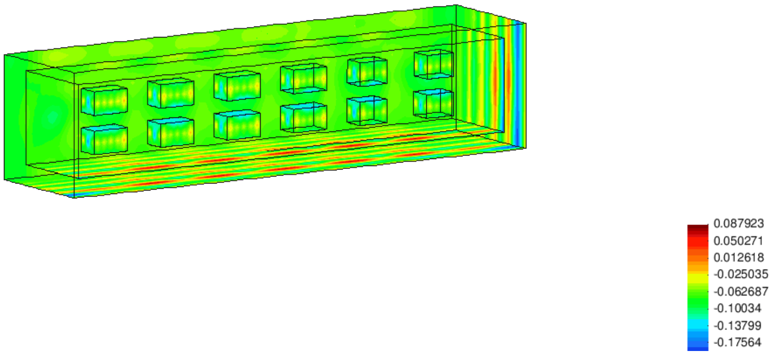
Figure 16: *Isosurface of the simulated exact solution of the component $E_2(x,t)$ for the forward problem at different times. The snapshot times are:* ***(a)*** *0.3 s,* ***(b)*** *0.6 s and* ***(c)*** *0.9 s, respectively.*



(a) t=1.2 s.

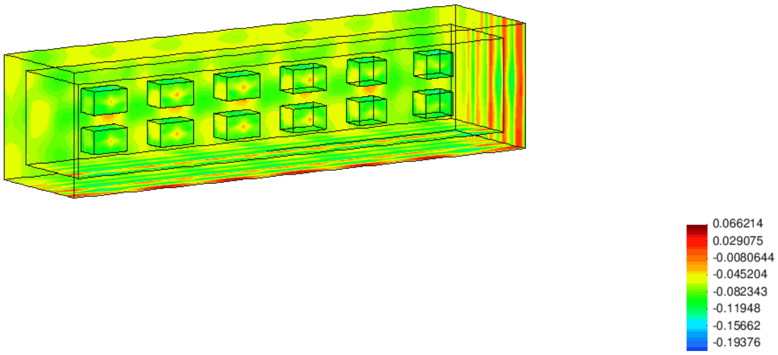


(b) t=1.5 s.

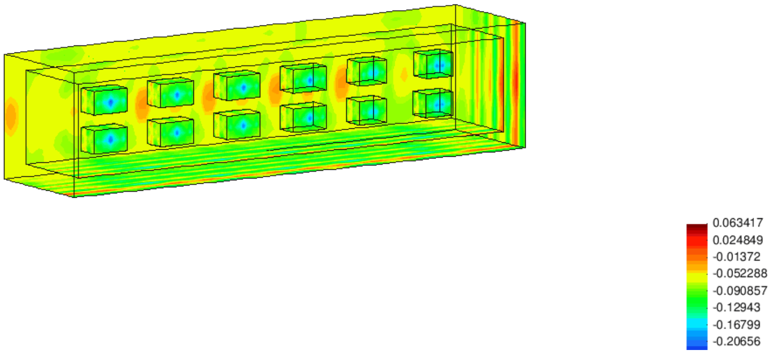


(c) t=1.8 s.

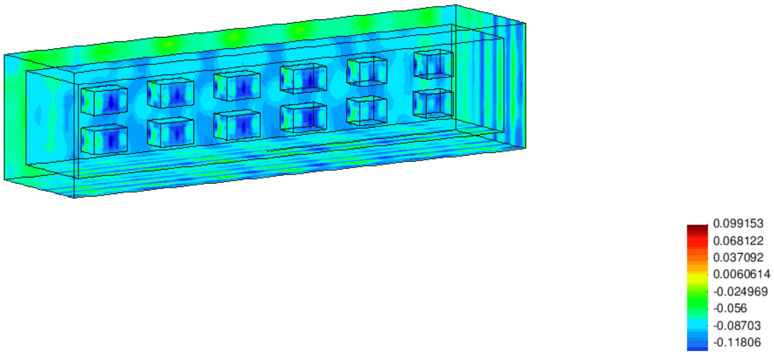
Figure 17: *Isosurface of the simulated exact solution of the component $E_2(x,t)$ for the forward problem at different times. The snapshot times are:* ***(a)*** *1.2 s,* ***(b)*** *1.5 s and* ***(c)*** *1.8 s, respectively.*



(a) t= 2.1 s.



(b) t= 2.4 s.



(c) t= 2.7 s.

Figure 18: *Isosurface of the simulated exact solution of the component $E_2(x,t)$ for the forward problem at different times. The snapshot times are:* ***(a)*** *2.1 s,* ***(b)*** *2.4 s and* ***(c)*** *2.7 s, respectively.*

Frequency ω		coarse mesh	1 ref. mesh	2 ref. mesh	3 ref. mesh	4 ref. mesh	5 ref. mesh
20	# nodes	10958	11061	11299	12121	14007	18758
	# elements	55296	55752		61440	72508	96378
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.92
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.52
30	# nodes	10958	11272	12115	13575	16763	25797
	# elements	55292	57018	60875	69450	885381	131932
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.98
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.56
40	# nodes	10958	11001	11091	11408	12156	13952
	# elements	55296	55392	55872	57544	61794	71276
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.97
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.565
50	# nodes	10958	11052	11901	12291	14017	18594
	# elements	55296	55693	56680	60228	72504	95554
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.96
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	1.65	1.65	1.65	1.65	1.65	1.464

Table 1: *Computed values of* $\varepsilon_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \varepsilon_r$ *and* $\mu_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \mu_r$ *on the adaptively refined meshes. Here regularization parameters,* γ_1, γ_2*, are fixed constants. Computations of CIP are done with the noise* $p = 0.03$ *in* (170).

Frequency ω		coarse mesh	1 ref. mesh	2 ref. mesh	3 ref. mesh	4 ref. mesh	5 ref. mesh
20	# nodes	10958	11075	11331	12292	14328	19954
	# elements	55296	55836	57042	62382	74354	102604
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.96
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.53
30	# nodes	10958	11309	12337	14043	17629	28738
	# elements	55292	57240	61930	72056	93656	146974
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.97
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.55
40	# nodes	10958	11001	11091	11408	12156	13952
	# elements	55296	55392	55872	57544	61794	71276
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.95
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.07	2.06	2.06	1.57
50	# nodes	10958	11087	11408	12291	14527	20542
	# elements	55296	55908	57432	62354	75464	105856
	$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.96
	$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.56

Table 2: *Computed values of* $\varepsilon_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \varepsilon_r$ *and* $\mu_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \mu_r$ *on the adaptively refined meshes. Here regularization parameters,* γ_1, γ_2*, are fixed constants. Computations of CIP are done with the noise* $p = 0.07$ *in* (170).

Test	Regularization parameter / Noise		coarse mesh	1 ref. mesh	2 ref. mesh	3 ref. mesh	4 ref. mesh	5 ref. mesh
1	Fixed / Noise 17%	# nodes	10958	11020	11170	11673	13330	16552
		# elements	55296	55542	56258	58994	68516	85226
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.95
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.56
2	Iterative rule 1 (89) / Noise 17%	# nodes	10958	11030	11165	11683	13244	16388
		# elements	55292	56224	59054	72056	68038	84338
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.97
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	2.06
3	Iterative rule 2 (91) / Noise 17%	# nodes	10958	11030	11165	11683	13245	16387
		# elements	55292	55566	56224	59054	68044	84332
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.99
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	2.06
4	Fixed / Noise 35%	# nodes	10958	11049	11270	11876	13943	18576
		# elements	55296	55680	56742	60096	72074	95586
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.96
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	1.56
5	Iterative rule 1 (89) / Noise 35%	# nodes	10958	11049	11271	11897	14042	18578
		# elements	55296	55680	56764	60210	72640	95614
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	12.99
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	2.06
6	Iterative rule 2 (91) / Noise 35%	# nodes	10958	11049	11273	11891	14050	18568
		# elements	55296	55680	56756	60176	72684	95586
		$\varepsilon_{\mathrm{r}}^{\mathrm{comp}}$	13	13	13	13	13	13
		$\mu_{\mathrm{r}}^{\mathrm{comp}}$	2.06	2.06	2.06	2.06	2.06	2.06

Table 3: *Computed values of* $\varepsilon_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \varepsilon_r$ *and* $\mu_{\mathrm{r}}^{\mathrm{comp}} := \max\limits_{\Omega_{FEM}} \mu_r$*. Computations of CIP are done with the two noise level* $p = 0.17$ *and* $p = 0.35$ *in* (170)*. Here regularization parameters,* γ_1, γ_2*, are given fixed constants (tests 1,4), iteratively computed via rule 1 given by* (89) *(tests 2,5) and iteratively computed via rule 2 given by* (91) *(tests 3,6).*

(a) $\max \varepsilon_r \approx 3.324$

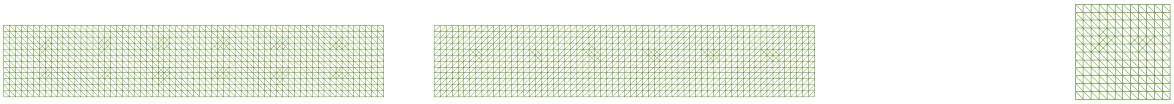
(b) $\max \mu_r \approx 1.04$

Figure 19: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after first iteration in conjugate gradient method on a coarse mesh with 3% additive noise (*$p = 0.03$ *in* (170)*), frequency* $\omega = 20$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*

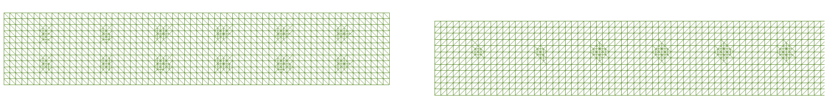
(a) $\max \varepsilon_r \approx 13$

(b) $\max \mu_r \approx 2.06$

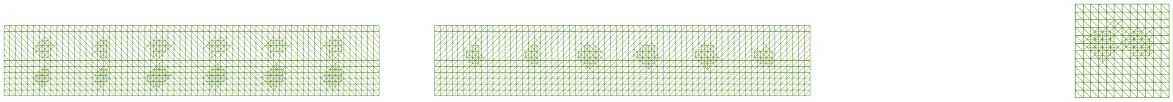
Figure 20: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 8-th iteration in conjugate gradient method on a coarse mesh with 3% additive noise (p=0.03 in* (170)*), frequency* $\omega = 20$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



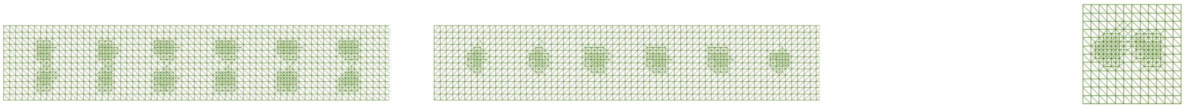
(a) 1 ref mesh, xy-view (b) xz-view (c) yz-view



(d) 2 ref mesh, xy-view (e) xz-view (f) yz-view

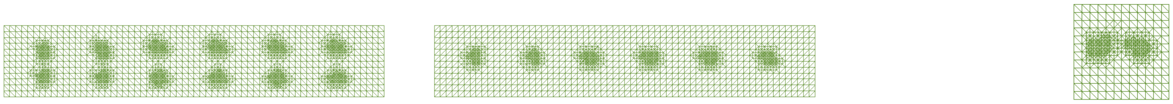


(h) 3 ref mesh, xy-view (i) xz-view (j) yz-view



(k) 4 ref mesh, xy-view (l) xz-view (m) yz-view

Figure 21: *Adaptively refined meshes of Table 1 for frequency* $\omega = 20$ *in* (168) *with additive noise 3% in data. Reconstructions on a coarse mesh are presented in Figure 20.*

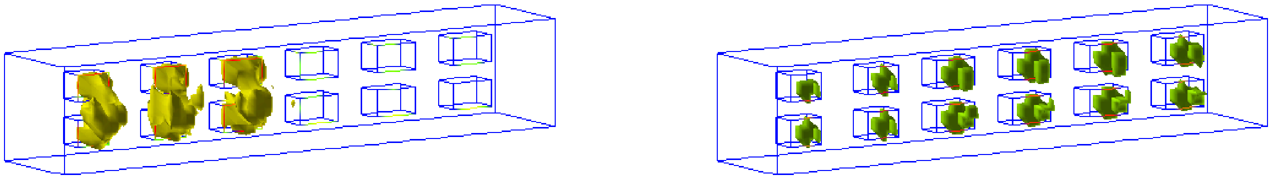


(a) 5 ref mesh, xy-view (b) xz-view (c) yz-view

Figure 22: *Different projections of five times adaptively refined mesh for computed images of Figure 23.*

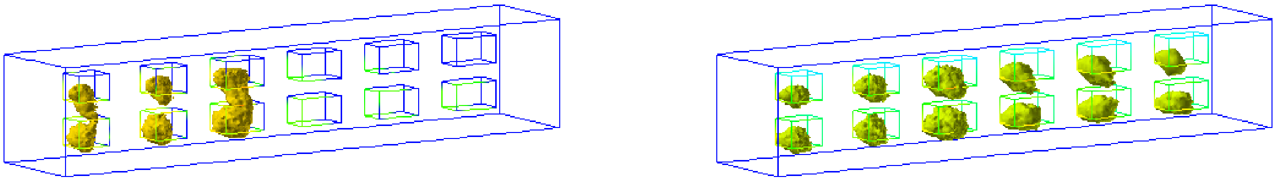
(a) $\max \varepsilon_r \approx 12.92$ (b) $\max \mu_r \approx 1.52$

Figure 23: *Computed images of reconstructed functions $\mu_r(\mathbf{x})$ and $\varepsilon_r(\mathbf{x})$ after second iteration in conjugate gradient method on a 5 times adaptively refined mesh presented in Figure 22 with 3% additive noise (p=0.03 in* (170)*), frequency $\omega = 20$ in* (168)*, regularization parameters, γ_1, γ_2, are fixed constants.*



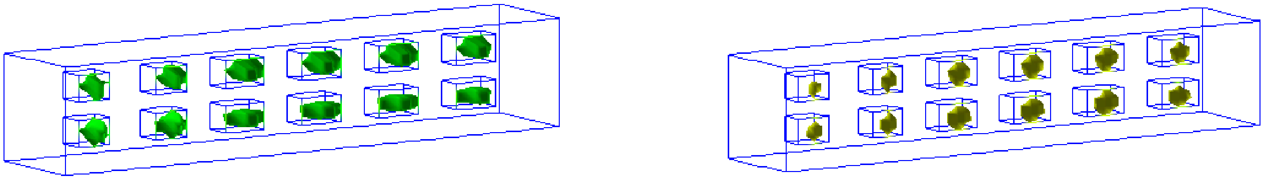
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 24: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 6-th iteration in conjugate gradient method on a coarse mesh with 3% additive noise (*$p = 0.03$ *in* (170)*), frequency* $\omega = 30$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



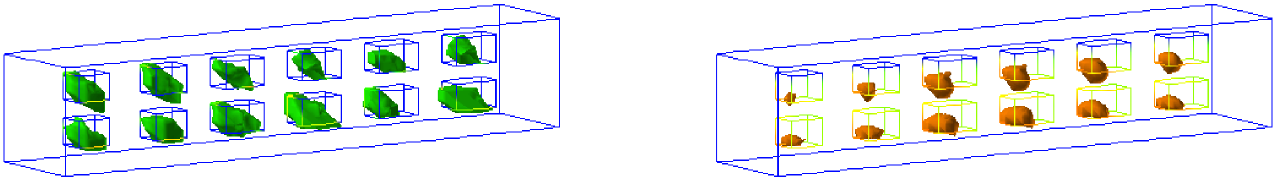
(a) $\max \varepsilon_r \approx 12.98$ (b) $\max \mu_r \approx 1.56$

Figure 25: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 3% additive noise (*$p = 0.03$ *in* (170)*), frequency* $\omega = 30$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



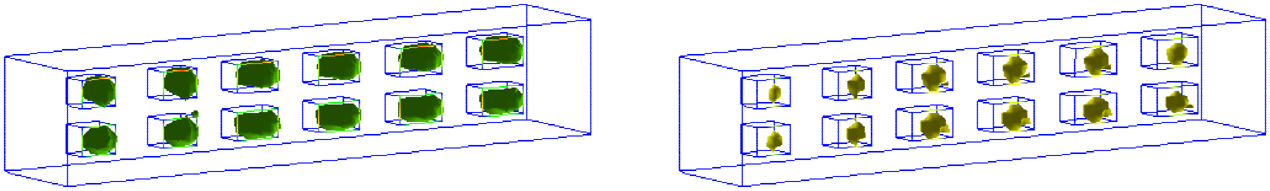
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 26: *Computed images of reconstructed functions $\mu_r(\mathbf{x})$ and $\varepsilon_r(\mathbf{x})$ after 10-th iteration in conjugate gradient method on a coarse mesh with 3% additive noise ($p = 0.03$ in (170)), frequency $\omega = 40$ in (168), regularization parameters, γ_1, γ_2, are fixed constants.*



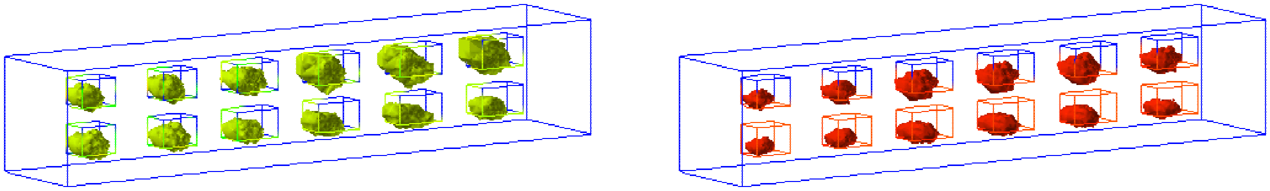
(a) $\max \varepsilon_r \approx 12.97$ (b) $\max \mu_r \approx 1.565$

Figure 27: *Computed images of reconstructed functions $\mu_r(\mathbf{x})$ and $\varepsilon_r(\mathbf{x})$ after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 3% additive noise ($p = 0.03$ in (170)), frequency $\omega = 40$ in (168), regularization parameters, γ_1, γ_2, are fixed constants.*



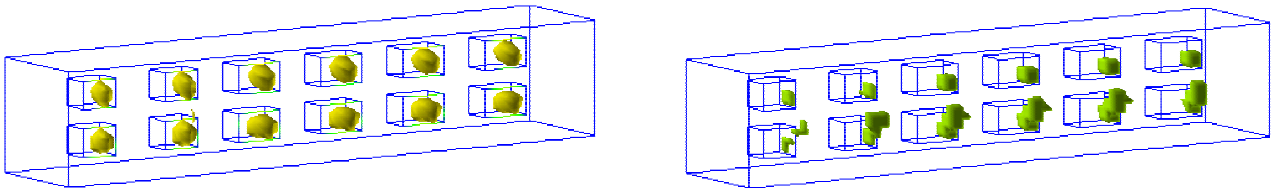
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 1.65$

Figure 28: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 10-th iteration in conjugate gradient method on a coarse mesh with 3% additive noise (*$p = 0.03$ *in* (170)*), frequency* $\omega = 50$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants..*



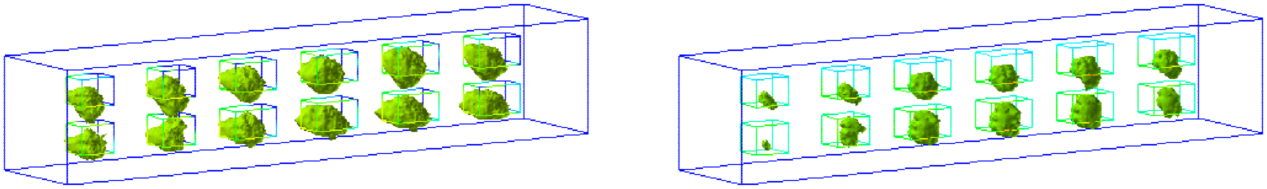
a) $\max \varepsilon_r \approx 12.96$ b) $\max \mu_r \approx 1.464$

Figure 29: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 3% additive noise (*$p = 0.03$ *in* (170)*), frequency* $\omega = 50$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



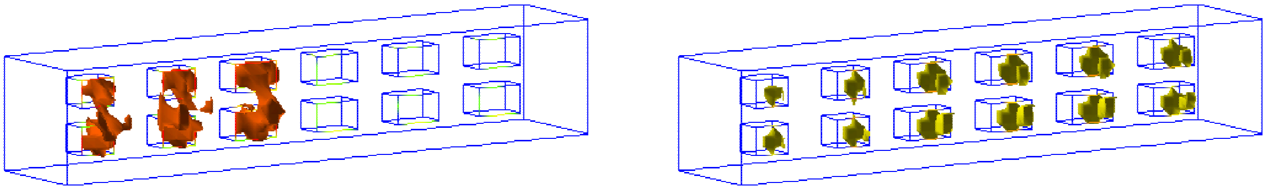
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 30: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 8-th iteration in conjugate gradient method on a coarse mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 20$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



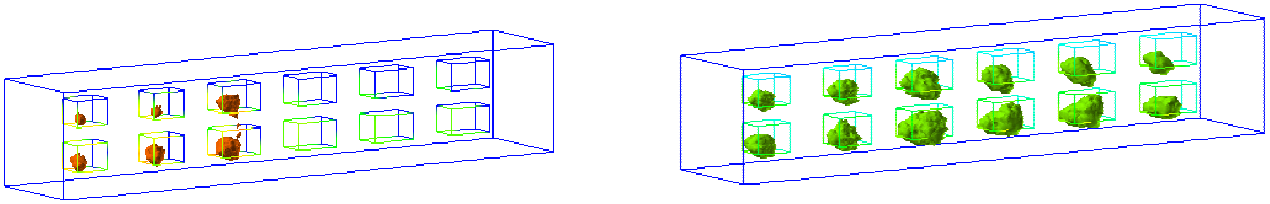
(a) $\max \varepsilon_r \approx 12.96$ (b) $\max \mu_r \approx 1.53$

Figure 31: *Interpolation of computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 10-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 20$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



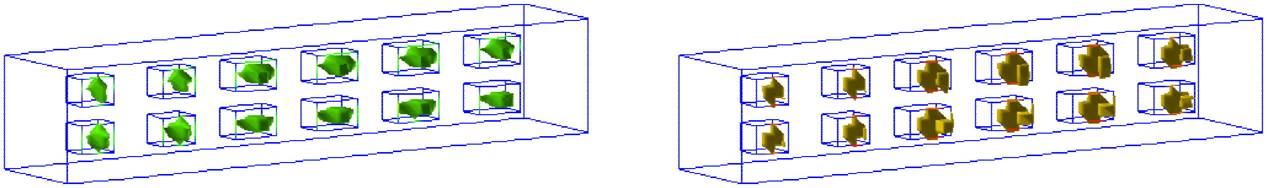
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 32: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 6-th iteration in conjugate gradient method on a coarse mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 30$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



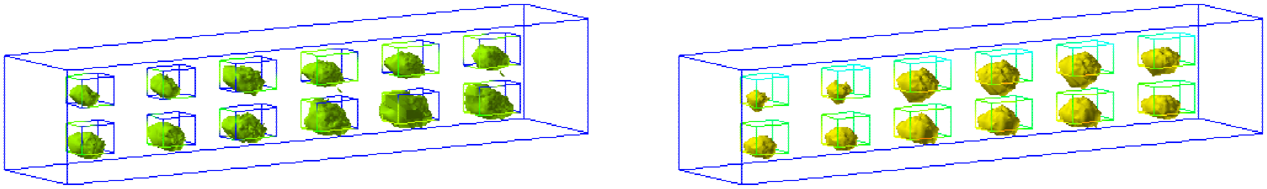
(a) $\max \varepsilon_r \approx 12.97$ (b) $\max \mu_r \approx 1.55$

Figure 33: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 30$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



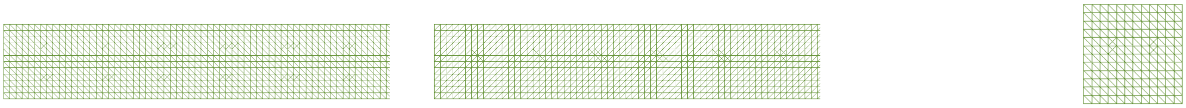
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 34: *Computed images of reconstructed functions* $\varepsilon_r(\mathbf{x})$*and* $\mu_r(\mathbf{x})$ *from backscattering data on a coarse mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*

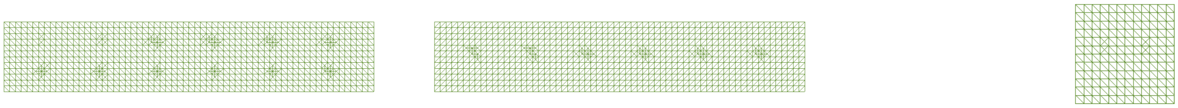


(a) $\max \varepsilon_r \approx 12.95$ (b) $\max \mu_r \approx 1.57$

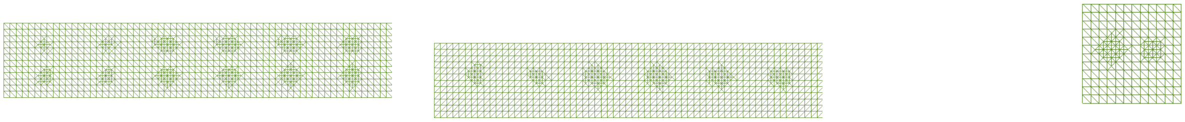
Figure 35: *Computed images of reconstructed functions* $\varepsilon_r(\mathbf{x})$ *and* $\mu_r(\mathbf{x})$ *from backscattering data on a 5 times adaptively refined mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



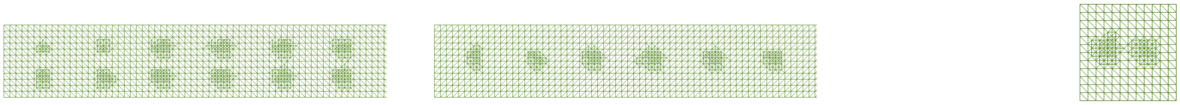
(a) 1 ref mesh, xy-view (b) xz-view (c) yz-view



(d) 2 ref mesh, xy-view (e) xz-view (f) yz-view

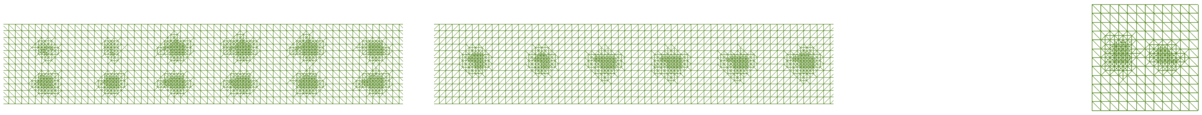


(h) 3 ref mesh, xy-view (i) xz-view (j) yz-view

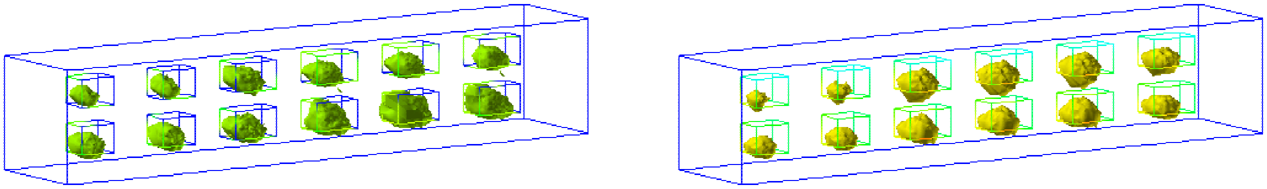


(k) 4 ref mesh, xy-view (l) xz-view (m) yz-view

Figure 36: *Adaptively refined meshes of Table 2 when* $\omega = 40$ *in* (170) *with 7% additive noise in data.*

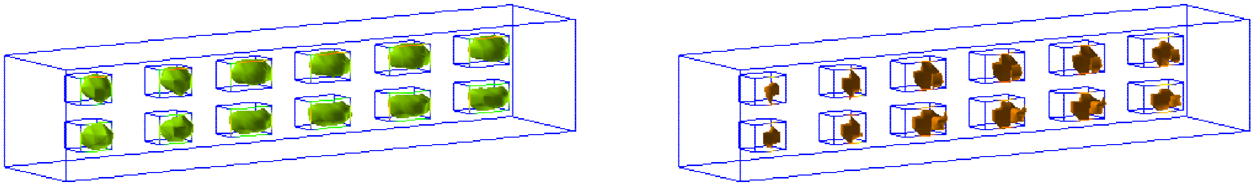


(a) 5 ref mesh, xy-view (b) xz-view (c) yz-view



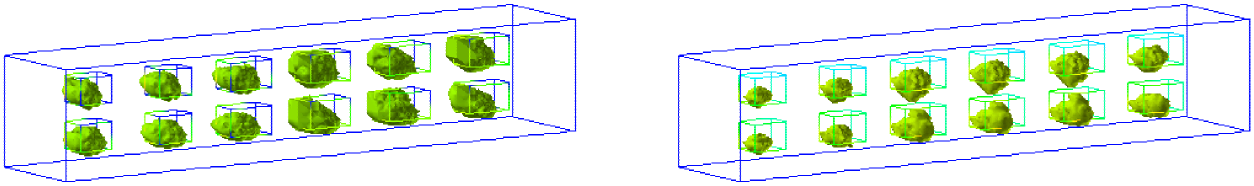
a) $\max \varepsilon_r \approx 12.95$ b) $\max \mu_r \approx 1.57$

Figure 37: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 7% additive noise (*$p = 0.07$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



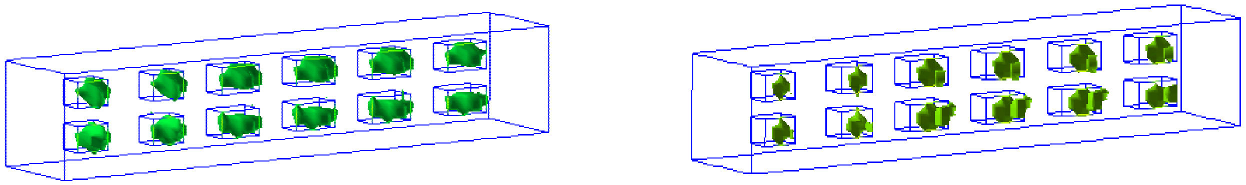
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 38: *Computed images of reconstructed functions $\mu_r(\mathbf{x})$ and $\varepsilon_r(\mathbf{x})$ after 7-th iteration in conjugate gradient method on a coarse mesh with 7% additive noise ($p = 0.07$ in (170)), frequency $\omega = 50$, regularization parameters, γ_1, γ_2, are fixed constants.*



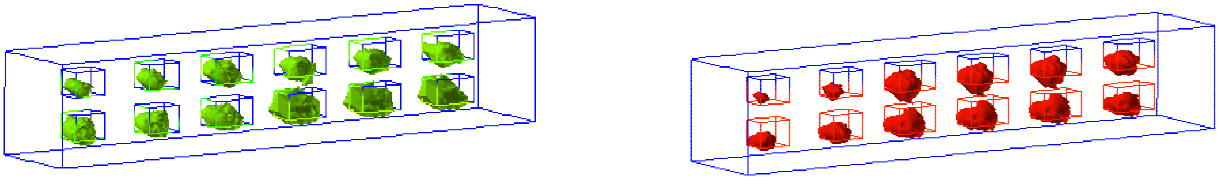
(a) $\max \varepsilon_r \approx 12.96$ (b) $\max \mu_r \approx 1.56$

Figure 39: *Computed images of reconstructed functions $\mu_r(\mathbf{x})$ and $\varepsilon_r(\mathbf{x})$ after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 7% additive noise ($p = 0.07$ in (170)), frequency $\omega = 50$ in (168), regularization parameters, γ_1, γ_2, are fixed constants.*



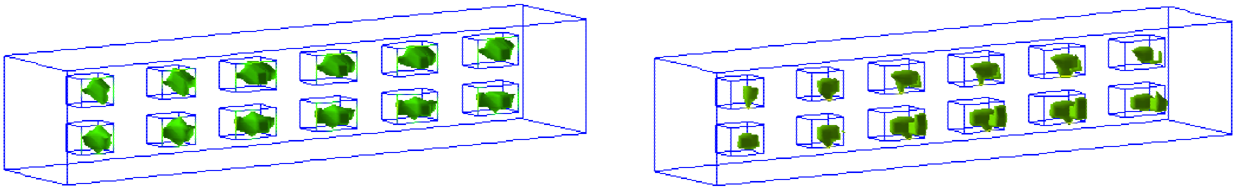
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 1.58$

Figure 40: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after first iteration in conjugate gradient method on first adaptively refined mesh with 12% additive noise (*$p = 0.12$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



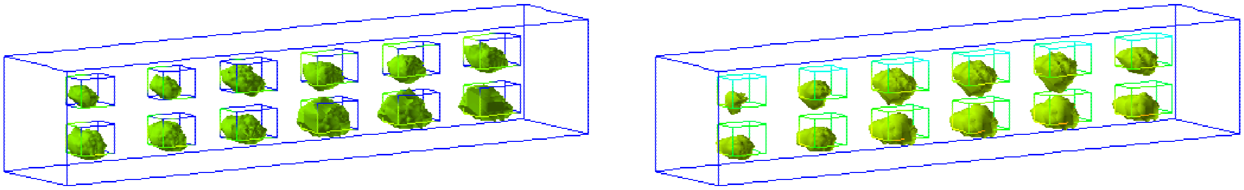
(a) $\max \varepsilon_r \approx 12.9933$ (b) $\max \mu_r \approx 1.374$

Figure 41: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 10-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 12% additive noise (*$p = 0.12$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



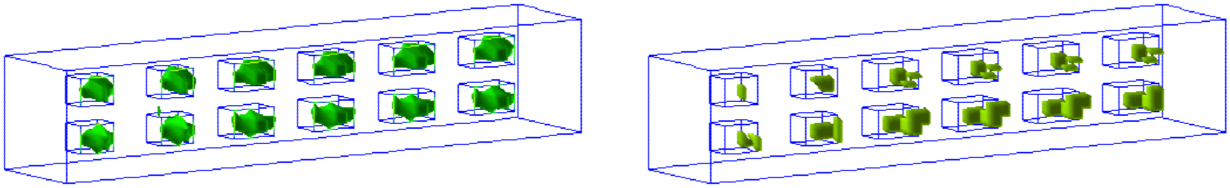
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 42: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 6-th iteration in conjugate gradient method on a coarse mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



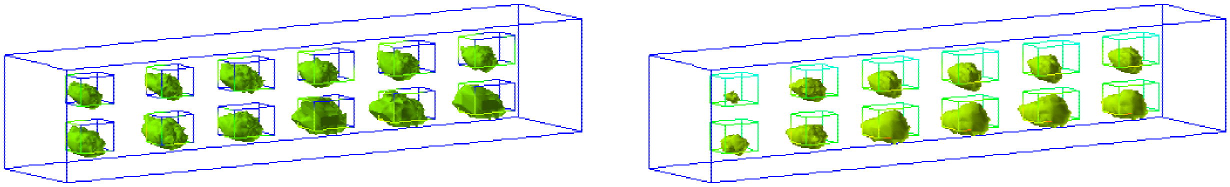
(a) $\max \varepsilon_r \approx 12.95$ (b) $\max \mu_r \approx 1.56$

Figure 43: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *from backscattering data on a 5 times adaptively refined mesh, with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



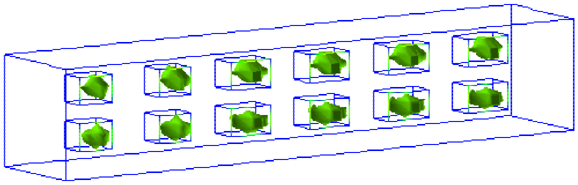
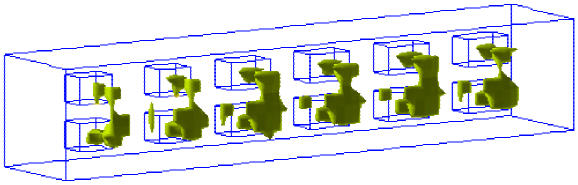
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 44: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 7-th iteration in conjugate gradient method on a coarse mesh with 25% additive noise (*$p = 0.25$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,*γ_1, γ_2*, are fixed constants.*



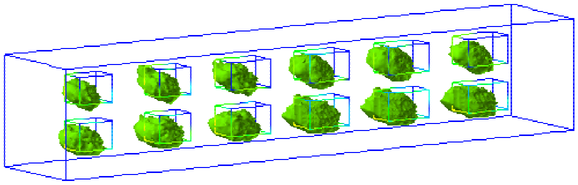
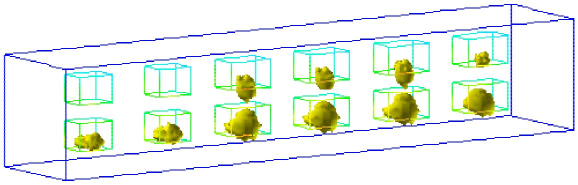
a) $\max \varepsilon_r \approx 12.97$ b) $\max \mu_r \approx 1.571$

Figure 45: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 25% additive noise (*$p = 0.25$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



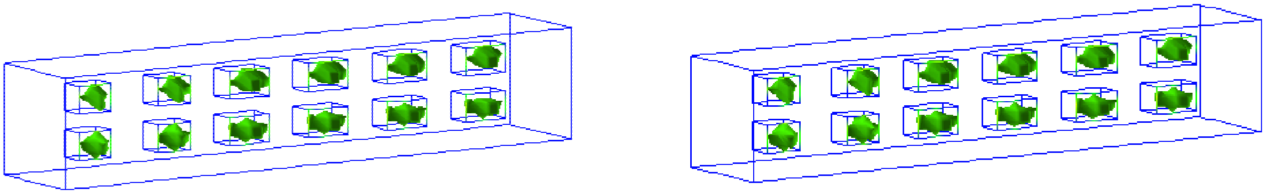
(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

Figure 46: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 7-th iteration in conjugate gradient method on a coarse mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



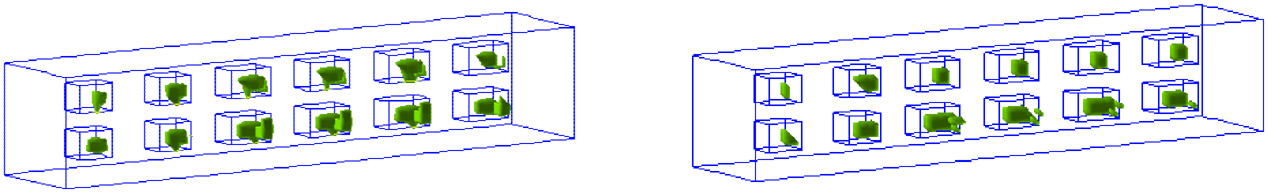
a) $\max \varepsilon_r \approx 12.9687$ b) $\max \mu_r \approx 1.562$

Figure 47: *Computed images of reconstructed functions* $\mu_r(\mathbf{x})$ *and* $\varepsilon_r(\mathbf{x})$ *after 3-th iteration in conjugate gradient method on a 5 times adaptively refined mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters,* γ_1, γ_2*, are fixed constants.*



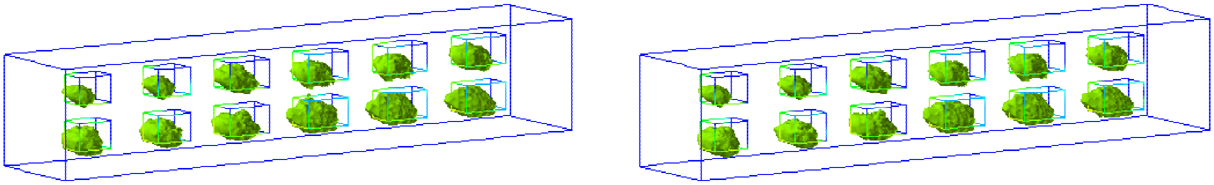
(a) $\max \varepsilon_r \approx 13$ (b) $\max \varepsilon_r \approx 13$

Figure 48: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *from backscattering data on a coarse mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89)*.*



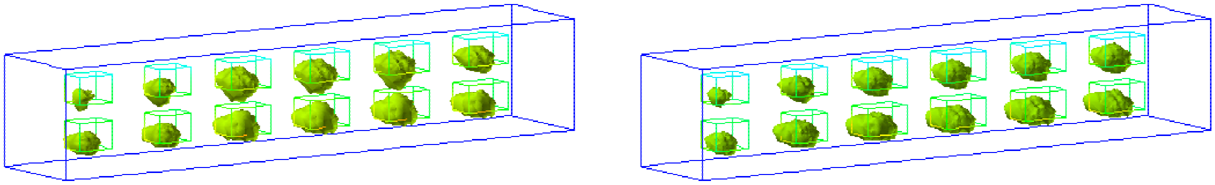
(a) $\max \mu_r \approx 2.06$ (b) $\max \mu_r \approx 2.06$

Figure 49: *Comparison between computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a coarse mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89)*.*



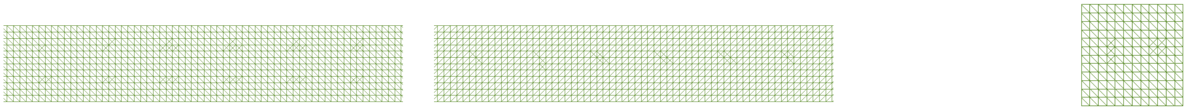
(a) $\max \varepsilon_r \approx 12.95$ (b) $\max \varepsilon_r \approx 12.97$

Figure 50: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *from backscattering data on a 5 times adaptively refined mesh, with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168) *with different regularization parameters* γ_1, γ_2*: **(a)** fixed* γ_1, γ_2*, **(b)** iteratively computed* γ_1, γ_2 *via rule 1 given by* (89).

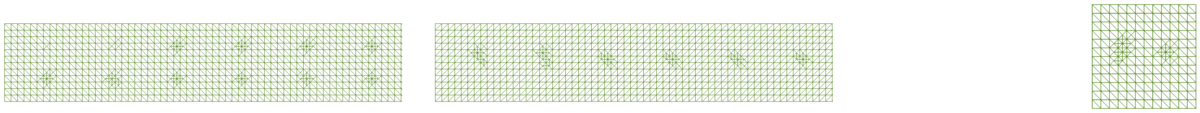


(a) $\max \mu_r \approx 1.56$ (b) $\max \mu_r \approx 2.06$

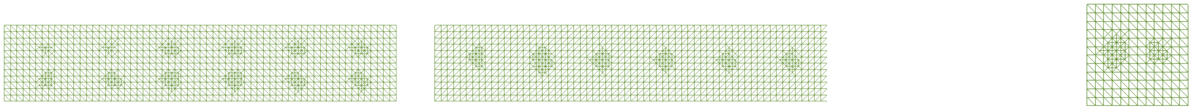
Figure 51: *Comparison between computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, regularization parameters* γ_1, γ_2*: **(a)** fixed* γ_1, γ_2*, **(b)** iteratively computed* γ_1, γ_2 *via rule 1 given by* (89).



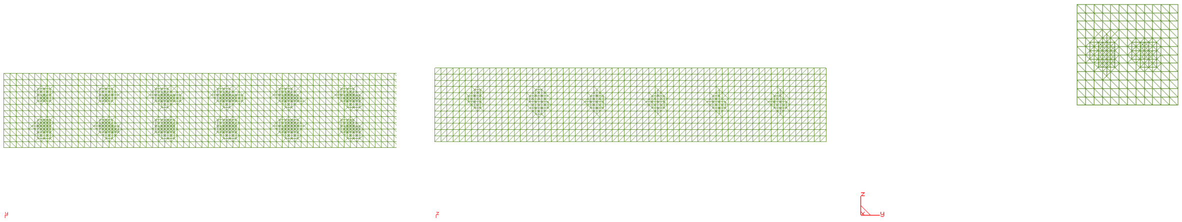
(a) 1 ref mesh, xy-view (b) xz-view (c) yz-view



(d) 2 ref mesh, xy-view (e) xz-view (f) yz-view

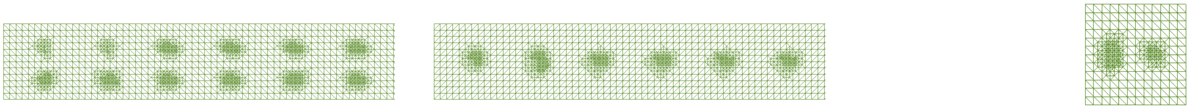


(h) 3 ref mesh, xy-view (i) xz-view (j) yz-view

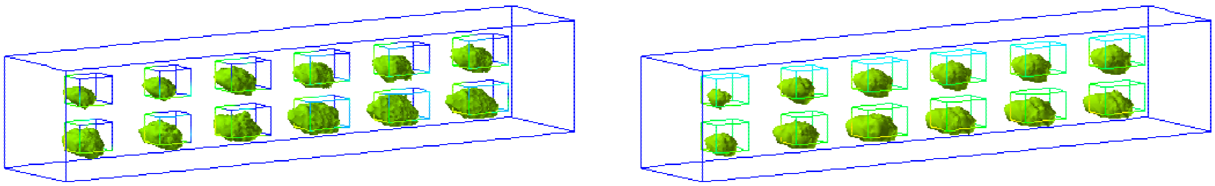


(k) 4 ref mesh, xy-view (l) xz-view (m) yz-view

Figure 52: *Adaptively refined mesh of test 2 in Table 3.*

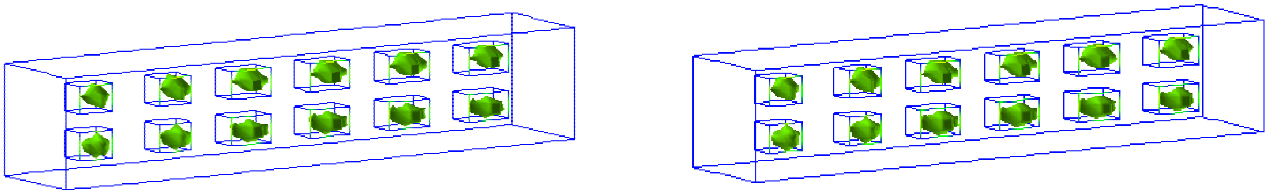


(a) 5 ref mesh, xy-view (b) xz-view (c) yz-view



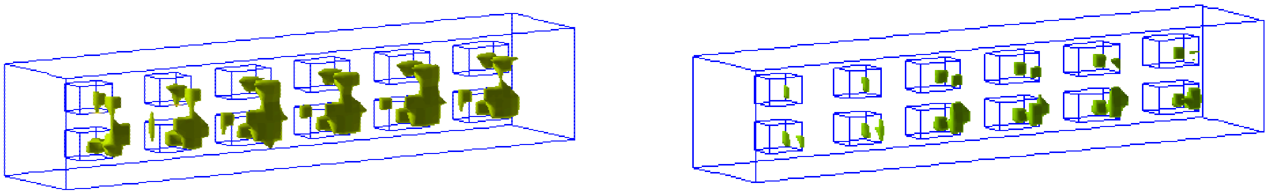
(a) $\max \varepsilon_r \approx 12.97$ (b) $\max \mu_r \approx 2.06$

Figure 53: *Computed images of reconstructed functions $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$ on a 5 times adaptively refined mesh with 17% additive noise ($p = 0.35$ in (170)), frequency $\omega = 40$ in (168), regularization parameters, γ_1, γ_2, iteratively computed via rule 1 given by (89).*



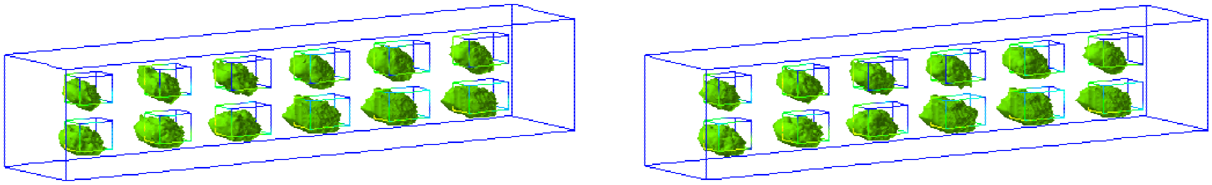
(a) $\max \varepsilon_r \approx 13$ (b) $\max \varepsilon_r \approx 13$

Figure 54: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a coarse mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters,* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89)*.*



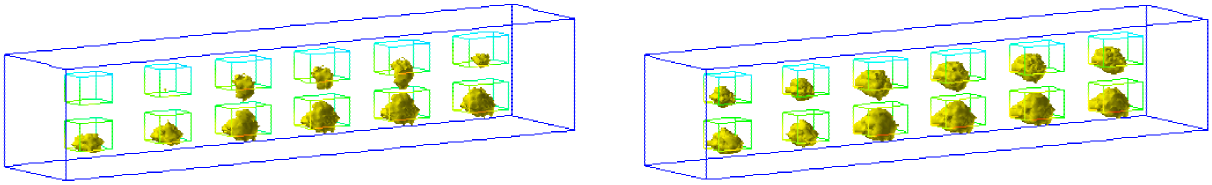
(a) $\max \mu_r \approx 2.06$ (b) $\max \mu_r \approx 2.06$

Figure 55: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a coarse mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89)*.*



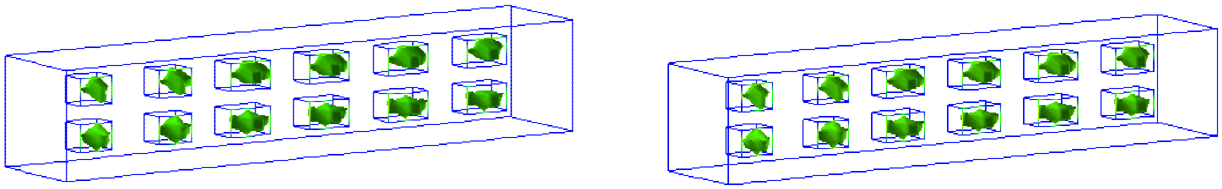
(a) $\max \varepsilon_r \approx 12.96$ (b) $\max \varepsilon_r \approx 12.99$

Figure 56: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89).



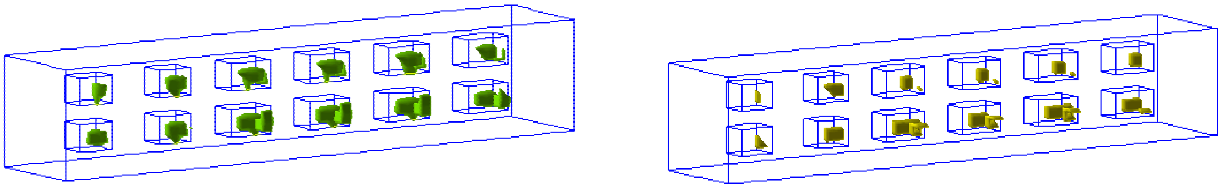
(a) $\max \mu_r \approx 1.56$ (b) $\max \mu_r \approx 2.06$

Figure 57: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89).



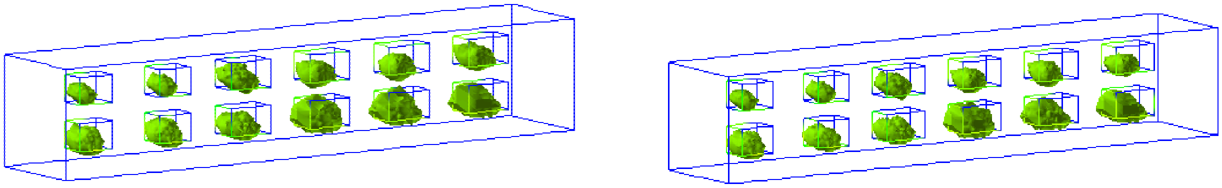
(a) $\max \varepsilon_r \approx 13$ (b) $\max \varepsilon_r \approx 13$

Figure 58: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a coarse mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



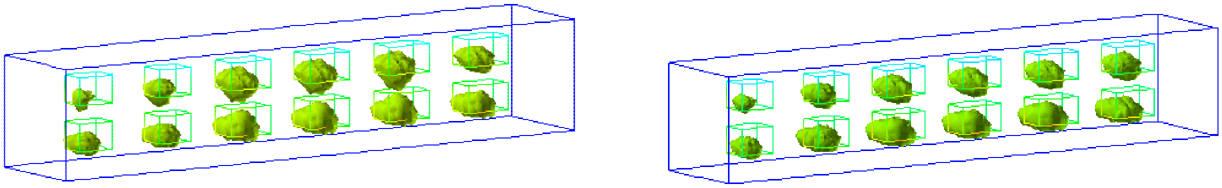
(a) $\max \mu_r \approx 2.06$ (b) $\max \mu_r \approx 2.06$

Figure 59: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a coarse mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



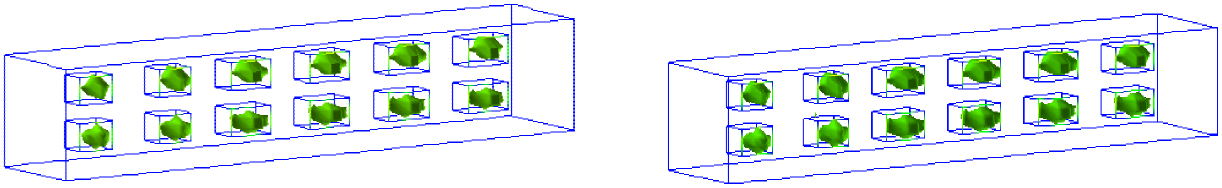
(a) $\max \varepsilon_r \approx 12.95$ (b) $\max \varepsilon_r \approx 12.99$

Figure 60: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



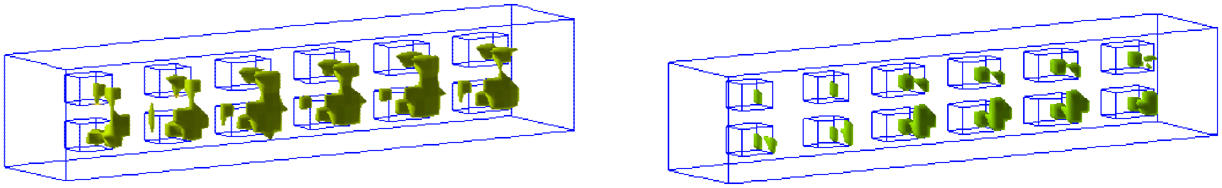
(a) $\max \mu_r \approx 1.56$ (b) $\max \mu_r \approx 2.06$

Figure 61: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 17% additive noise (*$p = 0.17$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



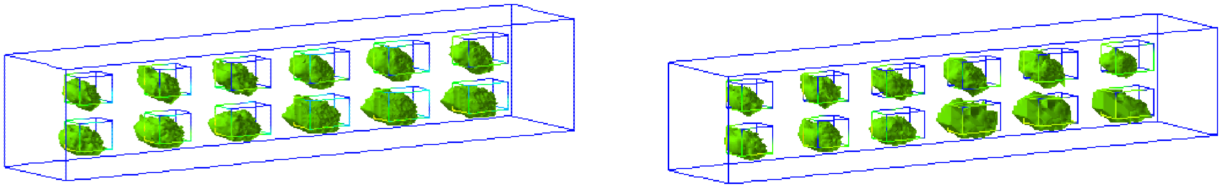
(a) $\max \varepsilon_r \approx 13$ (b) $\max \varepsilon_r \approx 13$

Figure 62: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a coarse mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



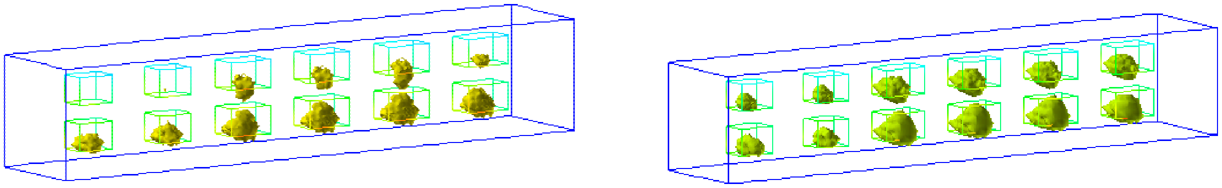
(a) $\max \mu_r \approx 2.06$ (b) $\max \mu_r \approx 2.06$

Figure 63: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a coarse mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).



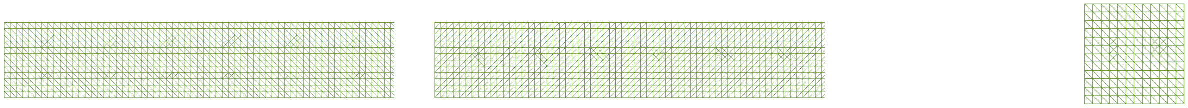
(a) $\max \varepsilon_r \approx 12.96$ (b) $\max \varepsilon_r \approx 13$

Figure 64: *Comparison of computed images of reconstructed function* $\varepsilon_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 35% additive noise* ($p = 0.35$ *in* (170))*, frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91)*.*

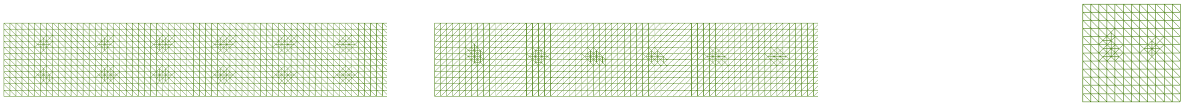


(a) $\max \mu_r \approx 1.56$ (b) $\max \mu_r \approx 2.06$

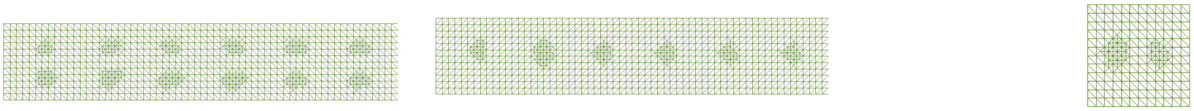
Figure 65: *Comparison of computed images of reconstructed function* $\mu_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 35% additive noise* ($p = 0.35$ *in* (170))*, frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)*** *fixed* γ_1, γ_2*,* ***(b)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91)*.*



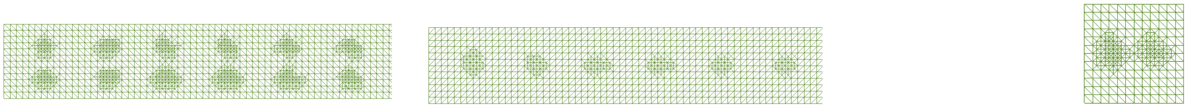
(a) 1 ref mesh, xy-view (b) xz-view (c) yz-view



(d) 2 ref mesh, xy-view (e) xz-view (f) yz-view

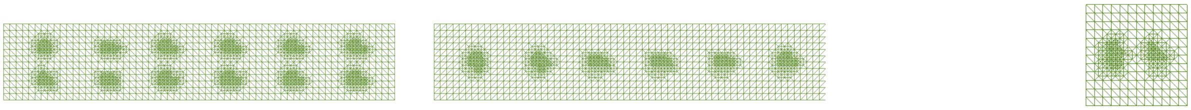


(h) 3 ref mesh, xy-view (i) xz-view (j) yz-view



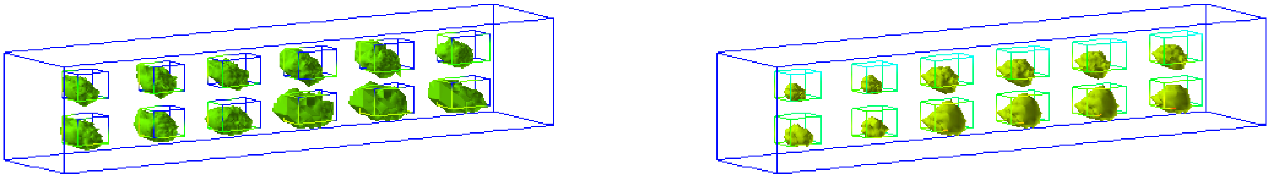
(k) 4 ref mesh, xy-view (l) xz-view (m) yz-view

Figure 66: *Adaptively refined meshes of test 6 in Table 3.*



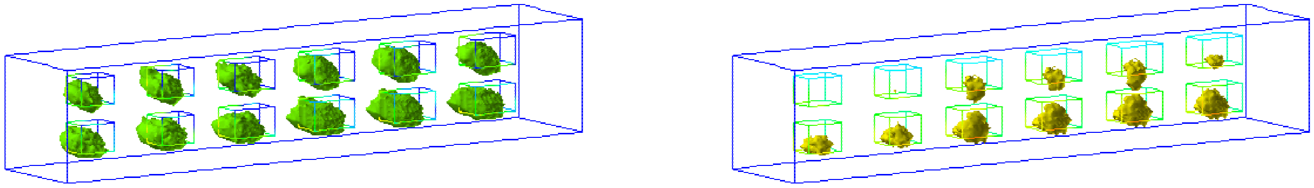
(a) 5 ref mesh, xy-view (b) xz-view (c) yz-view

Figure 67: *Different projections of five times adaptively refined mesh for computed images of Figure 68.*

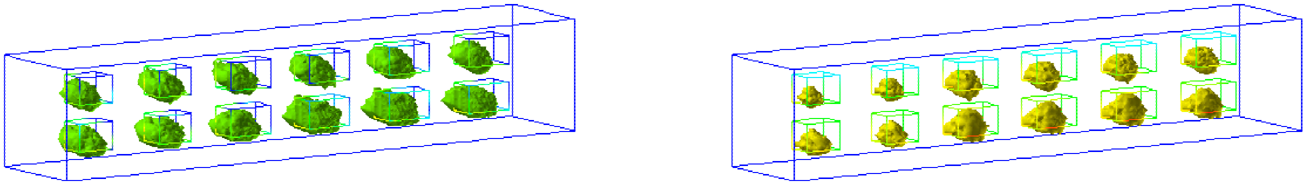


(a) $\max \varepsilon_r \approx 13$ (b) $\max \mu_r \approx 2.06$

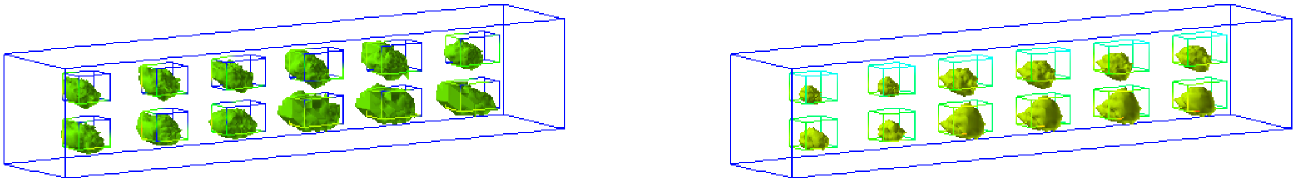
Figure 68: *Comparison of computed images of reconstructed functions $\varepsilon_r(\mathbf{x})$ and $\mu_r(\mathbf{x})$ on a 5 times adaptively refined mesh presented in Figure 67 with 35% additive noise ($p = 0.35$ in (170)), frequency $\omega = 40$ in (168), regularization parameters, γ_1, γ_2, iteratively computed via rule 2 given by (91).*



(a) $\max \varepsilon_r \approx 12.96$ (b) $\max \mu_r \approx 1.56$



(c) $\max \varepsilon_r \approx 12.99$ (d) $\max \mu_r \approx 2.06$



(e) $\max \varepsilon_r \approx 13$ (f) $\max \mu_r \approx 2.06$

Figure 69: *Comparison of computed images of reconstructed functions* $\varepsilon_r(\mathbf{x})$ *and* $\mu_r(\mathbf{x})$ *on a 5 times adaptively refined mesh with 35% additive noise (*$p = 0.35$ *in* (170)*), frequency* $\omega = 40$ *in* (168)*, with different regularization parameters* γ_1, γ_2*:* ***(a)-(b)*** *fixed* γ_1, γ_2*,* ***(c)-(d)*** *iteratively computed* γ_1, γ_2 *via rule 1 given by* (89)*,* ***(e)-(f)*** *iteratively computed* γ_1, γ_2 *via rule 2 given by* (91).

References

[1] F. Assous, P. Degond, E. Heintze and P. Raviart, On a finite-element method for solving the three-dimensional Maxwell equations, *J. Comput. Physics,* 109, 222–237, 1993.

[2] A. B. Bakushinskii, Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion, *Comput. Maths. Math. Phys.* , 24(4), 181-182, 1984.

[3] F. Bauer, M. Reiß, Regularization independent of the noise level: an analysis of quasi-optimality, *Inverse Problems*, 24 055009, 2008.

[4] L. Beilina, S. Hosseinzadegan, An Adaptive finite element method in reconstruction of coefficients in Maxwell's equations, *Paper in preparation*.

[5] L. Beilina, M. Kilbanov, *Approximate global convergence and adaptivity for coefficient inverse problems*, Springer, New-York, 2012.

[6] L. Beilina, M. Kilbanov, M. Kokurin, Adaptivity with relaxation for ill-posed problems and global convergence for a coefficient inverse problem, *Journal of Mathematical Sciences* 167, pp. 279-325, 2010.

[7] A. Bakushinsky, M. Y. Kokurin, A. Smirnova, *Iterative methods for ill-posed problems*, Inverse and Ill-Posed Problems Series 54, De Gruyter, 2011.

[8] A. Bakushinsky, A. Goncharsky, *Ill-Posed Problems: Theory and Applications*, Kluwer, Drodrecht, 1994.

[9] L. Beilina, M. Grote, Adaptive Hybrid Finite Element/Difference Method for Maxwell's equations, *TWMS J. of Pure and Applied Mathematics*, V.1(2), pp.176-197, 2010.

[10] L. Beilina, N.T. Thành, M. V.Kilbanov,J. B.Malmberg, Globally convergent and adaptive finite element methods in imaging of buried object from experimental back scattering radar measurements, *Journal of Computational and Applied Mathematics, Elsevier*, 2015.

[11] L. Beilina, N. T. Thành, M. V. Klibanov, J. Bondestam-Malmberg, Methods of quantitative reconstruction of shapes and refractive indices from experimental data, *Inverse Problems and Applications, Springer Proceedings in Mathematics & Statistics*, Vol. 120, 2015.

[12] L. Beilina, N. T. Thành, M. V. Klibanov, M. A. Fiddy, *Reconstruction from blind experimental data for an inverse problem for a hyperbolic equation*, *Inverse Problems*, 30, 025002, doi:10.1088/0266-5611/30/2/025002, 2014.

[13] L. Beilina, Energy estimates and numerical verification of the stabilized domain decomposition finite element/finite difference approach for the Maxwell's system in time domain, *CEJM*, 11(4), pp.702-733, DOI: 10.2478/s11533-013-0202-3, 2013.

[14] L. Beilina, Adaptive Finite Element Method for a coefficient inverse problem for the Maxwell's system, *Applicable Analysis*, V.90(10), pp.1461-1479, 2011.

[15] L. Beilina, M. Christofol and K. Niinimäki, Optimization approach for the simultaneous reconstruction of the dielectric permittivity and magnetic permeability functions from limited observations, *Inverse Problems and Imaging*, 9 (1), pp. 1-25, 2015.

[16] M.I. Belishev and V.M. Isakov, On the Uniqueness of the Recovery of Parameters of the Maxwell System from Dynamical Boundary Data, *Journal of Mathematical Sciences* 122, 3459-3469, 2004.

[17] L. Beilina and C. Johnson, A posteriori error estimation in computational inverse scattering, *Mathematical Models and Methods in Applied Sciences*, V.15, N.1, pp.23-37, 2005.

[18] M. Bellassoued, M. Cristofol, and E. Soccorsi, Inverse boundary value problem for the dynamical heterogeneous Maxwell's system, *Inverse Problems*, 28, 095009, 2012.

[19] S. C. Brenner and L. R. Scott, *Mathematical theory of finite element methods*, *Springer-Verlag*, Berlin, 1994.

[20] J. Bondestam Malmberg, A posteriori error estimate in the Lagrangian setting for an inverse problem based on a new formulation of Maxwell's system, *Inverse Problems and Applications, Springer Proceedings in Mathematics & Statistics*, Vol. 120, 2015, pp. 42-53, 2015.

[21] J. Bondestam Malmberg, A posteriori error estimation in a finite element method for reconstruction of dielectric permittivity, *ARXIV*, arXiv:1502.07658.

[22] P. Caro, Stable determination of the electromagnetic coefficients by boundary measurements, *Inverse Problems*, **26** (2010), 105014.

[23] P. Caro, P. Ola and M. Salo, Inverse boundary value problem for Maxwell equations with local data, *Comm. PDE*, **34** (2009), 1425-1464.

[24] J. M. Catalá-Civera, A.J. Canós-Marin, L. Sempere and E. de los Reyes, Microwave resonator for the non-invasive evaluation of degradation processes in liquid composites, *IEEE Electron. Lett.*, **37** (2001), 99-100.

[25] H. S. Cho, S. Kim, J. Kim and J. Jung, Determination of allowable defect size in multi-mode polymer waveguides fabricated by printed circuit board compatible processes, *J. Micromech. Microeng.*, **20** (2010), doi:10.1088/0960-1317/20/3/035035.

[26] G. C. Cohen, *Higher order numerical methods for transient wave equations*, Springer-Verlag, 2002.

[27] R. Courant, K. Friedrichs and H. Lewy, On the partial differential equations of mathematical physics, *IBM Journal of Reseach and Development*, **11** (1967), 215-234.

[28] K. Eriksson, D. Estep, C. Johnson, *Calculus in Several Dimensions*, Springer-Verlag, Berlin, 2004.

[29] M. Eller, V. Isakov, G. Nakamura and D.Tataru, Uniqueness and stability in the Cauchy problem for Maxwell and elasticity systems, *in "Non-linear Partial Differential Equations and their Applications", Collège de France Seminar*, **14** (2002), 329-349.

[30] A. Elmkies and P. Joly, Finite elements and mass lumping for Maxwell's equations: the 2D case, *Numerical Analysis, C. R. Acad. Sci. Paris*, **324** (1997), 1287-1293.

[31] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Boston: Kluwer Academic Publishers, 2000.

[32] Engquist B and Majda A, Absorbing boundary conditions for the numerical simulation of waves, *Math. Comp.*, **31** (1977), 629-651.

[33] M. J. Grote, A. Schneebeli, D. Schötzau, Interior penalty discontinuous Galerkin method for Maxwell's equations: Energy norm error estimates, *Journal of Computational and Applied Mathematics, Elsevier Science Publishers*, **204** (2007), 375-386.

[34] *Website of the Hindawi Publishing Corporation*, *http://www.hindawi.com/*.

[35] T.J.R., Hughes, *The Finite Element Method*, Prentice Hall, Englewood Cliffs, 1987.

[36] P, Joly, Variational Methods for Time-Dependent Propagation Problems, *Topics in Computational Wave Propagation, Lect. Notes Comput. Sci. Eng.*, 31, Springer, Berlin, 2003, 201–264.

[37] M.V. Klibanov, *Uniqueness of the solution of two inverse problems for a Maxwellian system*, *Computational Math. and Math.Phys.*, 26 (1986), 67-73.

[38] Y. Kurylev, M. Lassas and E. Somersalo, Maxwell's equations with a polarization independent wave velocity: Direct and inverse problems, *J. Math. Pures Appl.*, **86** (2006), 237-270.

[39] A.V. Kuzhuget, L. Beilina, M.V. Klibanov, A. Sullivan, L. Nguyen and M.A. Fiddy, Blind experimental data collected in the field and an approximately globally convergent inverse algorithm, *Inverse Problems*, **28** (2012), 095007.

[40] R. L. Lee and N. K. Madsen, A mixed finite element formulation for Maxwell's equations in the time domain, *Journal of Computational Physics*, **88** (1990), 284-304.

[41] S. Li and M. Yamomoto, Carleman estimate for Maxwell's Equations in anisotropic media and the observability inequality, *Journal of Physics: Conference Series*, **12** (2005) 110-115.

[42] S. Li and M. Yamamoto, An inverse source problem for Maxwell's equations in anisotropic media, *Applicable Analysis*, **84** (2005).

[43] P. B. Monk, *Finite Element methods for Maxwell's equations*, Oxford University Press, 2003.

[44] J. Monzó, A. Díaz, J.V. Balbastre, D. Sánchez-Hernández and E. de los Reyes Selective heating and moisture levelling in microwave-assisted trying of laminal materials:explicit model, *8-th Int. Conf. on Microwave and High Frequancy heating AMPERE, Bayreuth*, Germany, 2001.

[45] C. D. Munz, P. Omnes, R. Schneider, E. Sonnendrucker and U. Voss, Divergence correction techniques for Maxwell Solvers based on a hyperbolic model, *Journal of Computational Physics*, **161** (2000), 484–511.

[46] J.C. Nédélec, A new family of mixed finite elements in $\mathbb{R}^3$, *NUMMA*, **50** (1986), 57-81.

[47] P. Ola, L. Päivarinta and E. Somersalo, An inverse boundary value problem in electrodynamics, *Duke Math. J.*, **70** (1993), 617-653.

[48] K. D. Paulsen and D. R. Lynch, Elimination of vector parasities in Finite Element Maxwell solutions, *IEEE Trans.Microwave Theory Tech.*, **39** (1991), 395-404.

[49] O. Pironneau, *shape design for elliptic systems*, Springer-Verlag, Berlin, 1984.

[50] PETSc, Portable, Extensible Toolkit for Scientific Computation software, *http://www.mcs.anl.gov/petsc/*.

[51] J. Pitarch, M. Contelles-Cervera, F. L Peñaranda-Foix, J. M. Catalá-Civera, Determination of the permittivity and permeability for waveguides partially loaded with isotropic samples, *Measurement Science and Technology*, **17** (2006), 145-152.

[52] M. Salo, C. Kenig and G. Uhlmann, Inverse problems for the anisotropic Maxwell equations, *Duke Math. J.*, **157** (2011), 369-419.

[53] Y. G. Smirnov, Y. G Shestopalov, E. D. Derevyanchyk, Reconstruction of permittivity and permeability tensors of anisotropic materials in a rectangular waveguides from the reflection and transmission coefficients from different frequencies, *Proceedings of Progress in Electromagnetic Research Symposium*, Stockholm, Sweden, (2013), 290-295.

[54] D. R. Smith, S. Schultz, P. Markos and C. M. Soukoulis, Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients, *Physical Review B*, **65** (2002), DOI:10.1103/PhysRevB.65.195104.

[55] L. R. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math.Comp.*, **54** (1990), 483–493.

[56] A. N. Tikhonov, A. S. Leonov, A. G. Yagola, *Nonlinear ill-posed problems*, Chapman & Hall, 1998.

[57] N. T. Thành, L. Beilina, M. V. Klibanov, M. A. Fiddy, Reconstruction of the refractive index from experimental backscattering data using a globally convergent inverse method, *SIAM J. Scientific Computing*,36 (3) (2014), 273-293.

[58] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer, London, 1995.

[59] V. V. Vasin, L. B. Ryashko, *Ill-posed problems with a priori information*, *VNU*, Utrecht, 1995.

[60] WavES, the software package, http://www.waves24.com.

[61] Website of the Wikipedia, http://en.wikipedia.org/wiki/.

[62] Website of the wordpress, https://cmontalto.wordpress.com.