Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/120101

An Adaptive Finite Element Method in Quantitative Reconstruction of Small Inclusions from Limited Observations

John Bondestam Malmberg and Larisa Beilina*

Department of Mathematicals Sciences, Chalmers University of Technology and Gothenburg University, 412 96 Gothenburg, Sweden

Received: 3 Sep. 2017, Revised: 2 Dec. 2017, Accepted: 8 Dec. 2017
Published online: 1 Jan. 2018

Abstract: We consider a coefficient inverse problem to determine the dielectric permittivity in Maxwell's equations, with data consisting of boundary measurements. The true dielectric permittivity is assumed to belong to an ideal space of very fine finite elements. The problem is treated using a Lagrangian approach to the minimization of a Tikhonov functional, where an adaptive finite element method forms the basis of the computations. A new a posteriori error estimate for the norm of the error in the reconstructed permittivity is derived. The adaptive algorithm is formulated and tested successfully in numerical experiments for the reconstruction of two, three, and four small inclusions with low contrast, as well as the reconstruction of a superposition of two Gaussian functions.

Keywords: coefficient inverse problem, finite element approximation, adaptivity

1 Introduction

In this note we study an adaptive finite element method for the reconstruction of a dielectric permittivity function $\varepsilon = \varepsilon(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, x_3) \in \Omega$, where $\Omega \subset \mathbb{R}^3$ is a bounded convex domain with piecewise smooth boundary Γ. We assume that the true permittivity belongs to an ideal space of very fine finite elements. The data considered consists of a finite number of observations of (one or two) backscattered or transmitted waves. This is a coefficient inverse problem (CIP) for Maxwell's equations, where the dielectric permittivity function ε, acting as the coefficient in the equations, characterizes an inhomogeneous, isotropic, non-magnetic, non-conductive medium in Ω. Although we are in this note concerned with the reconstruction of a real-valued coefficient ε, our future applications are, eventually, in medical imaging, such as microwave imaging of breast cancer, and early diagnosis of stroke (see [1,2,3,4] for details).

The method studied is based on a Lagrangian approach to the minimization of a Tikhonov functional, where the functions involved are approximated by piecewise polynomials in a finite element method. Such an approach to coefficient inverse problems of the type we consider has previously been studied extensively in the context of a two-stage procedure in [5,6,7,8]. The main application considered in those publications was detection of explosives. With such applications, the low amount of data one can expect – usually only backscattered data for a single incident wave of one frequency – makes the problem of reconstruction challenging.

The main theoretical result of this paper is Theorem 1, which provides a new, direct a posteriori error estimate for the norm of the error in the reconstructed dielectric permittivity function. Qualitative computable, or a posteriori, error estimates are an essential tool for finite element based adaptive algorithms in the optimization approach to our inverse problem. Previously, in [5], such an estimate was given for our coefficient inverse problem. However, this estimate was an indirect one, estimating the size of an error in the computed Lagrangian, as apposed to a direct estimate of the error in the computed permittivity as presented here. A similar estimate, also in the computed Lagrangian, was shown for a modified Maxwell system in [9].

We illustrate our theoretical results with several numerical examples. With the above-mentioned aspects of contrast and size of inclusions in mind, we present reconstructions of two, three, and four different small inclusions, respectively, of low contrast. We also evaluate

* Corresponding author e-mail: john.bondestam.malmberg@chalmers.se

how variations in the type of data collected affect the reconstruction. More precisely, we consider the effect of working with only backscattered data, with only transmitted data, with both backscattered and transmitted data, as well as backscattered data from two anti-parallel incident waves. In addition to these reconstructions of inclusions, we also present a reconstruction of a more complicated function, to wit, a superposition of two gaussians.

Here is an outline of the remaining part of this note: In the Section 2 we present the mathematical formulations of the direct and inverse problems and state the basic results prior to discretization of the problems. In Section 3 we state the finite element formulations, perform the error analysis, and summarize the results in terms of a mesh refinement strategy. The adaptive algorithm for solving the inverse problem is described in Section 4, and numerical examples are given in Section 5. Section 6 concludes the paper.

2 The direct and inverse problems

Before proceeding with the mathematical statement of the problem, we introduce some notation. For the bounded convex polyhedral domain $\Omega \subset \mathbb{R}^3$ with boundary Γ, we write $\Omega_T := \Omega \times (0,T)$ and $\Gamma_T := \Gamma \times (0,T)$, where $T > 0$ is a (sufficiently large) fixed time. If $X \subset \mathbb{R}^n$, $n \in \mathbb{N}$, is a domain, we denote by $\langle \cdot, \cdot \rangle_X$ and $\|\cdot\|_X$ the L_2-inner product and norm, respectively, over the domain X.

Let V_0^{per} be the space of Lipschitz continuous functions defined by

$$V_0^{\text{per}} := \{v \in C(\bar{\Omega}) : \nabla v \in [L_\infty(\Omega)]^3\}.$$

In the spirit of [6], we define a finite dimensional subspace of piecewise linears

$$V^{\text{per}} := \{v \in V_0^{\text{per}} : v|_K \in P^1(K) \, \forall K \in \mathscr{T}_{\text{f}}\},$$

where $\mathscr{T}_{\text{f}} = \{K\}$ is a very fine, quasi-uniform (see [10]), face-to-face partition of Ω into tetrahedra, and $P^1(K)$ denotes the set of polynomials of degree no greater than 1 over K. We equip this space with the L_2-norm over Ω, and define the set of admissible dielectric permittivity functions

$$\begin{aligned} U^{\text{per}} := \{ & v \in V^{\text{per}} : 1 \le v(\mathbf{x}) \le \varepsilon_{\max} \; \forall \mathbf{x} \in \Omega, \\ & v|_\Gamma \equiv 1, \; \nabla v|_\Gamma \equiv 0\} \end{aligned} \tag{1}$$

for some known upper bound $\varepsilon_{\max}$. The space V^{per} can be thought of as a very fine finite element space (see Section 3), hence it is of (large but) finite dimension, which, by the equivalence of norms on finite dimensional spaces, justifies the use of the L_2-norm rather than the more natural H^1-norm. Although this space is finite dimensional, it is considered too large to handle numerically, and our goal is to approximate $\varepsilon \in U^{\text{per}}$ by an element of a smaller subspace.

The set U^{per} is defined to describe a heterogeneous medium in Ω, immersed in a constant background with permittivity 1 in $\mathbb{R}^3 \setminus \Omega$.

Under the assumption that $\varepsilon \in U^{\text{per}}$, we consider the stabilized Maxwell system of [11,12] in time-domain for an isotropic, non-magnetic, and non-conductive medium in Ω. That is,

$$\begin{aligned} & \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} - \Delta \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}) - s\nabla(\nabla \cdot (\varepsilon \mathbf{E})) = 0 && \text{in } \Omega_T, \\ & \nabla \cdot (\varepsilon \mathbf{E}) = 0 && \text{in } \Omega_T, \\ & \frac{\partial \mathbf{E}}{\partial \mathbf{n}} = \mathbf{P} && \text{on } \Gamma_T, \\ & \mathbf{E}(\cdot, 0) = \frac{\partial \mathbf{E}}{\partial t}(\cdot, 0) = 0 && \text{in } \Omega, \end{aligned} \tag{2}$$

where we have expanded the usual double curl term as $\nabla \times (\nabla \times \mathbf{E}) = -\Delta \mathbf{E} + \nabla(\nabla \cdot \mathbf{E})$ to simplify further manipulations of the equation. We use the notation $\frac{\partial}{\partial \mathbf{n}} = \mathbf{n} \cdot \nabla$, where $\mathbf{n}$ denotes the outward unit normal on Γ. The function $\mathbf{P} \in [L_2(\Gamma_T)]^3$ is given Neumann data (see Section 4 of [7] for details), and $s \ge 1$ is a stabilization parameter. For well-posedness of problems of this class, we refer to [13,14].

The mathematical statement of the coefficient inverse problem is:

Inverse Problem. *Given time-resolved boundary observations* $\mathbf{G} \in [L_2(\Gamma_T)]^3$ *of the electric field, determine* $\varepsilon \in U^{\text{per}}$ *such that* $\mathbf{E} = \mathbf{G}$ *on* Γ_T.

Here the observations $\mathbf{G}$ are defined on the whole boundary Γ_T. For the case of incomplete data, which we consider in our numerical examples, we prescribe $\mathbf{E} - \mathbf{G} = 0$ on all parts of Γ_T where we have no observations.

Inverse problems, such as the one above, are typically ill-posed in nature. Thus we cannot expect to be able to find an exact unique solution for any given data $\mathbf{G}$ (which in practice *will* contain noise). Instead we will follow the concept of regularization (see for instance [15,16]) and assume that the given data $\mathbf{G}$ is a perturbation of ideal data $\mathbf{G}^*$, for which there exists a unique solution ε^* to the Inverse Problem. The goal is then to systematically compute an approximation of ε^*, a so-called regularized solution which hereafter will be denoted simply by ε, which is as close to ε^* as can be achieved given the level of noise $\|\mathbf{G} - \mathbf{G}^*\|_{\Gamma_T}$. Such a regularized solution is obtained by minimizing a Tikhonov functional, to be defined below (see equation (4)).

Uniqueness of the solution of coefficient inverse problems with a finite number of observations is typically obtained via the method of Carleman estimates [17]. Examples where this method is applied to inverse problems for Maxwell's equations can be found in, for example, [18], additional examples are [19] where simultaneous reconstruction of two coefficients is

considered, and [20,21] for bi-isotropic and anisotropic media. However, this technique requires non-vanishing initial conditions for the underlying partial differential equation, which is not the case here. Thus, to date, uniqueness of the solution to the problem we study is not known. For the purpose of this work, we will assume that uniqueness holds. This assumption is justified by the numerical reconstruction results presented in the experimental works [7,8].

We introduce the space

$$V^{\mathrm{dir}} := \{\mathbf{v} \in [H^1(\Omega_T)]^3 : \mathbf{v}(\cdot, 0) = 0\}$$

for solutions to the direct problem, and

$$V^{\mathrm{adj}} := \{\mathbf{v} \in [H^1(\Omega_T)]^3 : \mathbf{v}(\cdot, T) = 0\}$$

for adjoint solutions. Both spaces are equipped with the usual norm and inner product on $[H^1(\Omega_T)]^3$. Then, multiplying the first equation in (2) by a test function $\boldsymbol{\phi} \in V^{\mathrm{adj}}$ and integrating over Ω_T yields, after integration by parts,

$$\begin{aligned}
0 = &-\left\langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}, \tfrac{\partial \boldsymbol{\phi}}{\partial t} \right\rangle_{\Omega_T} + \left\langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}(\cdot, T), \boldsymbol{\phi}(\cdot, T) \right\rangle_{\Omega} \\
&- \left\langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}(\cdot, 0), \boldsymbol{\phi}(\cdot, 0) \right\rangle_{\Omega} + \langle \nabla \mathbf{E}, \nabla \boldsymbol{\phi} \rangle_{\Omega_T} \\
&- \left\langle \tfrac{\partial \mathbf{E}}{\partial \mathbf{n}}, \boldsymbol{\phi} \right\rangle_{\Gamma_T} - \langle \nabla \cdot \mathbf{E}, \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} \\
&+ \langle \nabla \cdot \mathbf{E}, \mathbf{n} \cdot \boldsymbol{\phi} \rangle_{\Gamma_T} + s \langle \nabla \cdot (\varepsilon \mathbf{E}), \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} \\
&- s \langle \nabla \cdot (\varepsilon \mathbf{E}), \mathbf{n} \cdot \boldsymbol{\phi} \rangle_{\Gamma_T} \\
= &- \left\langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}, \tfrac{\partial \boldsymbol{\phi}}{\partial t} \right\rangle_{\Omega_T} + \langle \nabla \mathbf{E}, \nabla \boldsymbol{\phi} \rangle_{\Omega_T} \\
&- \langle \nabla \cdot \mathbf{E}, \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} + s \langle \nabla \cdot (\varepsilon \mathbf{E}), \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} \\
&- \langle \mathbf{P}, \boldsymbol{\phi} \rangle_{\Gamma_T} \\
:= &\, \mathscr{D}(\varepsilon, \mathbf{E}, \boldsymbol{\phi}),
\end{aligned} \tag{3}$$

where the second equality holds because $\boldsymbol{\phi}(\cdot, T) = 0$, $\frac{\partial \mathbf{E}}{\partial t}(\cdot, 0) = 0$, $\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = \mathbf{P}$ on Γ_T, and $\nabla \cdot \mathbf{E} = \nabla \cdot (\varepsilon \mathbf{E}) = 0$ on Γ_T. This last observation follows from the second equation of (2), and the fact that $\varepsilon \in U^{\mathrm{per}}$ implies that $\varepsilon \equiv 1$ in some neighborhood of Γ. We arrive at the following weak description of the electric field:

Continuous Direct Problem. *Given $\varepsilon \in U^{\mathrm{per}}$, determine $\mathbf{E} \in V^{\mathrm{dir}}$ such that $\mathscr{D}(\varepsilon, \mathbf{E}, \boldsymbol{\phi}) = 0$ for every $\boldsymbol{\phi} \in V^{\mathrm{adj}}$.*

Let $\mathbf{E}_\varepsilon \in V^{\mathrm{dir}}$ denote the solution to the Continuous Direct Problem for a given $\varepsilon \in U^{\mathrm{per}}$. We can then define the Tikhonov functional $F : U^{\mathrm{per}} \to \mathbb{R}_+$,

$$F(\varepsilon, \mathbf{E}_\varepsilon) := \frac{1}{2} \|(\mathbf{E}_\varepsilon - \mathbf{G}) z_\delta\|^2_{\Gamma_T} + \frac{\alpha}{2} \|\varepsilon - \varepsilon_0\|^2_{\Omega}, \tag{4}$$

where $\alpha > 0$ is a regularization parameter and $z_\delta = z_\delta(t) \in C^\infty([0, T])$ is a cut-off function for the data, dropping from a constant level of 1 to a constant level of 0 within the small interval $(T - \delta, T - \delta/2)$, $0 < \delta \ll T$, as schematically shown in Figure 1. The function z_δ is introduced to ensure data compatibility in the adjoint problem arising in the minimization of (4). To simplify the notation, we write $F(\varepsilon) := F(\varepsilon, \mathbf{E}_\varepsilon)$.

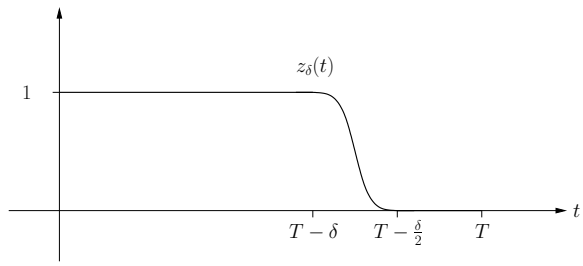


Fig. 1: Schematic illustration of the cut-off function z_δ appearing in the Tikhonov functional (4).

How to choose the regularization parameter α with respect to the level of noise in the data is a widely studied topic. Several methods exist. Examples are the (generalized) discrepancy principle [16] and iterative methods [22]. In future studies, we are planning to investigate iterative methods, similarly to [23]. However, for the results presented here, we regard α as a fixed parameter.

We assume that the initial approximation ε_0 is sufficiently close to an ideal solution ε^*, corresponding to noiseless data $\mathbf{G}^*$ in the Inverse Problem, to ensure local strong convexity of the Tikhonov functional. By theorem 3.1 of [6], this can be achieved by applying the approximately globally convergent method discussed in that paper. Then F is strongly convex in a neighborhood $\mathscr{N} \subset V^{\mathrm{per}}$ of ε_0, containing ε^*. More precisely, we have the estimate

$$\frac{\alpha}{2} \|\varepsilon_1 - \varepsilon_2\|^2_{\Omega} \le F'(\varepsilon_1; \varepsilon_1 - \varepsilon_2) - F'(\varepsilon_2; \varepsilon_1 - \varepsilon_2) \tag{5}$$

for every $\varepsilon_1, \varepsilon_2 \in \mathscr{N} \cap U^{\mathrm{per}}$, where $F'(\varepsilon; \bar{\varepsilon})$ denotes the Fréchet derivative of F at ε, acting on $\bar{\varepsilon}$.

Throughout the remaining part of this text we will assume that the hypothesis of Theorem 3.1 of [6], and hence strong convexity, holds. Then we may seek a minimizer $\varepsilon \in U^{\mathrm{per}}$ of F by applying any gradient based method (such as steepest descent, quasi-Newton, or conjugate gradient), starting from ε_0.

Such an approach requires that we compute the Fréchet derivative of F, which is complicated, since it involves the implicit dependence of $\mathbf{E}_\varepsilon$ upon ε. To simplify the analysis, in the spirit of optimal control (see for example [24,25] for the general theory and some specific examples), we introduce the Lagrangian associated to the problem of minimizing $F(\varepsilon, \mathbf{E})$, $\varepsilon \in U^{\mathrm{per}}$, $\mathbf{E} \in V^{\mathrm{dir}}$, with $\mathscr{D}(\varepsilon, \mathbf{E}, \boldsymbol{\phi}) = 0$ for all $\boldsymbol{\phi} \in V^{\mathrm{adj}}$ acting as a constraint. This Lagrangian is

$$L(u) := F(\varepsilon, \mathbf{E}) + \mathscr{D}(\varepsilon, \mathbf{E}, \boldsymbol{\lambda}),$$

where $u = (\varepsilon, \mathbf{E}, \boldsymbol{\lambda}) \in U := U^{\text{per}} \times V^{\text{dir}} \times V^{\text{adj}} \subset V := V^{\text{per}} \times V^{\text{dir}} \times V^{\text{adj}}$, $F(\varepsilon, \mathbf{E})$ was defined in (4), and $\mathscr{D}(\varepsilon, \mathbf{E}, \boldsymbol{\lambda})$ was defined in (3).

We can now minimize F over U^{per} by finding a stationary point of L over U. With the strong convexity as above, this would imply that we solve

Continuous Saddle Point Problem. *Find $u \in U$ such that $L'(u; v) = 0$ for every $v \in V$.*

Again we use the notation $L'(u; v)$ for the Fréchet derivative of L at u, acting on v. It can be shown that

$$L'(u; v) = \frac{\partial L}{\partial \varepsilon}(u; \bar{\varepsilon}) + \frac{\partial L}{\partial \mathbf{E}}(u; \bar{\mathbf{E}}) + \frac{\partial L}{\partial \boldsymbol{\lambda}}(u; \bar{\boldsymbol{\lambda}}),$$

where $u = (\varepsilon, \mathbf{E}, \boldsymbol{\lambda}) \in U$, $v = (\bar{\varepsilon}, \bar{\mathbf{E}}, \bar{\boldsymbol{\lambda}}) \in V$, and

$$\begin{aligned}
\frac{\partial L}{\partial \varepsilon}(u; \bar{\varepsilon}) &= \alpha \langle \varepsilon - \varepsilon_0, \bar{\varepsilon} \rangle_{\Omega} - \left\langle \tfrac{\partial \mathbf{E}}{\partial t} \cdot \tfrac{\partial \boldsymbol{\lambda}}{\partial t}, \bar{\varepsilon} \right\rangle_{\Omega_T} \\
&\quad + s \langle \nabla \cdot \boldsymbol{\lambda}, \nabla \cdot (\bar{\varepsilon} \mathbf{E}) \rangle, \\
\frac{\partial L}{\partial \mathbf{E}}(u; \bar{\mathbf{E}}) &= \langle (\mathbf{E} - \mathbf{G}) z_\delta^2, \bar{\mathbf{E}} \rangle_{\Gamma_T} - \left\langle \varepsilon \tfrac{\partial \boldsymbol{\lambda}}{\partial t}, \tfrac{\partial \bar{\mathbf{E}}}{\partial t} \right\rangle_{\Omega_T} \\
&\quad + \langle \nabla \boldsymbol{\lambda}, \nabla \bar{\mathbf{E}} \rangle_{\Omega_T} - \langle \nabla \cdot \boldsymbol{\lambda}, \nabla \cdot \bar{\mathbf{E}} \rangle_{\Omega_T} \\
&\quad + s \langle \nabla \cdot \boldsymbol{\lambda}, \nabla \cdot (\varepsilon \bar{\mathbf{E}}) \rangle := \mathscr{A}(\varepsilon, \boldsymbol{\lambda}, \bar{\mathbf{E}}), \\
\frac{\partial L}{\partial \boldsymbol{\lambda}}(u; \bar{\boldsymbol{\lambda}}) &= \mathscr{D}(\varepsilon, \mathbf{E}, \bar{\boldsymbol{\lambda}}).
\end{aligned} \tag{6}$$

In particular, we note that the solution $u = (\varepsilon, \mathbf{E}, \boldsymbol{\lambda})$ to the Continuous Saddle Point Problem must satisfy $\mathscr{D}(\varepsilon, \mathbf{E}, \bar{\boldsymbol{\lambda}}) = 0$ for every $\bar{\boldsymbol{\lambda}} \in V^{\text{adj}}$ and $\mathscr{A}(\varepsilon, \boldsymbol{\lambda}, \bar{\mathbf{E}}) = 0$ for every $\bar{\mathbf{E}} \in V^{\text{dir}}$. The former means that $\mathbf{E}$ solves Continuous Direct Problem and the latter that $\boldsymbol{\lambda}$ solves the following adjoint problem:

Continuous Adjoint Problem. *Given $\varepsilon \in U^{\text{per}}$, determine $\boldsymbol{\lambda} \in V^{\text{adj}}$ such that $\mathscr{A}(\varepsilon, \boldsymbol{\lambda}, \boldsymbol{\phi}) = 0$ for every $\boldsymbol{\phi} \in V^{\text{dir}}$.*

The functional $\mathscr{A}$ in the Continuous Adjoint Problem was defined in (6). The problem can be seen as a weak analogue of the following system, adjoint to (2):

$$\begin{aligned}
&\varepsilon \frac{\partial^2 \boldsymbol{\lambda}}{\partial t^2} - \Delta \boldsymbol{\lambda} + \nabla(\nabla \cdot \boldsymbol{\lambda}) - s\varepsilon \nabla(\nabla \cdot \boldsymbol{\lambda}) = 0 && \text{in } \Omega_T, \\
&\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{n}} = -(\mathbf{E} - \mathbf{G}) z_\delta^2 && \text{on } \Gamma_T, \\
&\boldsymbol{\lambda}(\cdot, T) = \frac{\partial \boldsymbol{\lambda}}{\partial t}(\cdot, T) = 0 && \text{in } \Omega.
\end{aligned}$$

These observations will be used in the error analysis to be described below. But first we shall make some remarks concerning the relation between the Fréchet derivative of Tikhonov functional and that of the Lagrangian.

Let $u_\varepsilon = (\varepsilon, \mathbf{E}_\varepsilon, \boldsymbol{\lambda}_\varepsilon)$ be the element of U obtained by taking $\mathbf{E}_\varepsilon$ as the solution to the Continuous Direct Problem and $\boldsymbol{\lambda}_\varepsilon$ as the solution to the Continuous Adjoint Problem for the given $\varepsilon \in U^{\text{per}}$. Then, under assumption of sufficient stability of the weak solutions $\mathbf{E}_\varepsilon$ and $\boldsymbol{\lambda}_\varepsilon$ with respect to ε, the observation that

$$F(\varepsilon) = F(\varepsilon, \mathbf{E}_\varepsilon) = F(\varepsilon, \mathbf{E}_\varepsilon) + \mathscr{D}(\varepsilon, \mathbf{E}_\varepsilon, \boldsymbol{\lambda}_\varepsilon) = L(u_\varepsilon),$$

(as $\mathscr{D}(\varepsilon, \mathbf{E}_\varepsilon, \boldsymbol{\lambda}_\varepsilon) = 0$) leads to

$$F'(\varepsilon; \cdot) = \frac{\partial L}{\partial \varepsilon}(u_\varepsilon; \cdot). \tag{7}$$

Estimate (5) and identity (7) will play an important role in the error analysis for the Tikhonov functional and for the coefficient.

3 Finite element formulations and error analysis

In this section we will give finite element formulations for discretizing the Continuous Direct, Saddle Point, and Adjoint Problems. After that we will turn to the error analysis. We begin by defining finite-dimensional analogues of the spaces V^{per}, V^{dir}, V^{adj}, and V, as well as subsets corresponding to U^{per} and U.

Let $\mathscr{T}_h := \{K\}$ be a quasi-uniform face-to-face partition of Ω into tetrahedra, such that $\mathscr{T}_{\text{f}}$ can be obtained from $\mathscr{T}_h$ by subdivision of tetrahedra. Let $\mathscr{I}_\tau$ be a uniform partition of $(0, T)$ into subintervals $(t_k, t_{k+1}]$, $t_k = k\tau$, $k = 0, \ldots, N_\tau$, of length $\tau = T/N_\tau$. With $\mathscr{T}_h$ we associate a mesh-function $h = h(\mathbf{x})$ such that

$$h(\mathbf{x}) = \operatorname{diam}(K) \tag{8}$$

for $\mathbf{x} \in K \in \mathscr{T}_h$. On these meshes we define

$$\begin{aligned}
V_h^{\text{per}} &:= \{v \in V_0^{\text{per}} : v|_K \in P^q(K)\ \forall K \in \mathscr{T}_h\}, \\
U_h^{\text{per}} &:= V_h^{\text{per}} \cap U^{\text{per}}, \\
V_h^{\text{dir}} &:= \{v \in V^{\text{dir}} : v|_{K \times I} \in [P^1(K)]^3 \times P^1(I) \\
&\qquad \forall K \in \mathscr{T}_h\ \forall I \in \mathscr{I}_\tau\}, \\
V_h^{\text{adj}} &:= \{v \in V^{\text{adj}} : v|_{K \times I} \in [P^1(K)]^3 \times P^1(I) \\
&\qquad \forall K \in \mathscr{T}_h\ \forall I \in \mathscr{I}_\tau\}, \\
V_h &:= V_h^{\text{per}} \times V_h^{\text{dir}} \times V_h^{\text{adj}}, \\
U_h &:= U_h^{\text{per}} \times V_h^{\text{dir}} \times V_h^{\text{adj}},
\end{aligned}$$

where $P^n(X)$ denotes the space of polynomials of degree at most $n \in \mathbb{N}$ over X, and the degree q used in the finite-dimensional analogue V_h^{per} of V^{per} is at least 1. Observe that the dependence on the step size τ in time is not explicitly included in the notation for the finite-dimensional spaces. This is justified by the fact that τ should be selected with regard to h in accordance with the Courant-Friedrichs-Lewy condition.

Using these spaces we can state finite element versions of the Continuous Direct and Adjoint Problems as follows:

Discrete Direct Problem. *Given $\varepsilon \in U^{\text{per}}$, determine $\mathbf{E}_h \in V_h^{\text{dir}}$ such that $\mathscr{D}(\varepsilon, \mathbf{E}_h, \boldsymbol{\phi}_h) = 0$ for every $\boldsymbol{\phi}_h \in V_h^{\text{adj}}$.*

Discrete Adjoint Problem. *Given $\varepsilon \in U^{\text{per}}$, determine $\boldsymbol{\lambda}_h \in V_h^{\text{adj}}$ such that $\mathscr{A}(\varepsilon, \boldsymbol{\lambda}_h, \boldsymbol{\phi}_h) = 0$ for every $\boldsymbol{\phi} \in V_h^{\text{dir}}$.*

The finite-dimensional analogue for the Continuous Saddle Point Problem is:

Discrete Saddle Point Problem. *Find* $u_h = (\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \in U_h$ *such that* $L'(u_h, v) = 0$ *for every* $v \in V_h$.

The same remark that was made in conjunction with the Continuous Saddle Point Problem is also valid here: it holds that $\mathbf{E}_h$ solves the Discrete Direct Problem, and $\boldsymbol{\lambda}_h$ solves the Discrete Adjoint Problem, for $\varepsilon = \varepsilon_h$.

Having stated the Inverse Problem, the continuous problems, and their discrete counterparts, let us briefly summarize how these problems relate to each other, before turning to the error analysis.

Recall that our main goal is to find the dielectric permittivity which gave rise to the observed data $\mathbf{G}$, in other words to solve the Inverse Problem. Due to ill-posedness, this goal is practically unattainable, and we focus instead on finding a regularized solution, which can be done by finding a stationary point to the Lagrangian, that is, by solving the Continuous Saddle Point Problem. This, in turn, relies upon solving the Continuous Direct, and Adjoint Problems. However, the continuous problems cannot, in general, be solved exactly, but approximately through their finite-dimensional analogues, the Discrete Direct, Adjoint, and Saddle Point Problems.

The purpose of the error estimation below is to quantify the discrepancy between the solutions to the continuous problems, and the solutions to the discrete problems, in order to be able to adaptively refine the latter three problems so that the solution to the Discrete Saddle Point Problem, the approximation of the regularized solution, fits the solution to the Continuous Saddle Point Problem, the true regularized solution, as closely as desired. We will now focus on this estimation.

We begin by introducing some additional notation. For $v = (\varepsilon, \mathbf{E}, \boldsymbol{\lambda}) \in V$ we denote (with some slight abuse of notation) its interpolant in V_h by

$$\Pi_h v = (\Pi_h \varepsilon, \Pi_h \mathbf{E}, \Pi_h \boldsymbol{\lambda}),$$

and the interpolation error by

$$r_h v = v - \Pi_h v = (r_h \varepsilon, r_h \mathbf{E}, r_h \boldsymbol{\lambda}).$$

We will also need to consider jumps of discontinuous functions over $\mathscr{T}_h$ and $\mathscr{I}_\tau$. Let K_1, $K_2 \in \mathscr{T}_h$ share a common face f. For $\mathbf{x} \in f$ we define

$$\{\mathbf{v}\}_{\mathrm{s}}(\mathbf{x}) := \mathbf{n}_1 \cdot \left(\lim_{\mathbf{y}\to\mathbf{x}, \mathbf{y}\in K_1} \mathbf{v}(\mathbf{y}) - \lim_{\mathbf{y}\to\mathbf{x}, \mathbf{y}\in K_2} \mathbf{v}(\mathbf{y}) \right), \quad (9)$$

where $\mathbf{n}_1$ is the outward unit normal of K_1. This is well-defined since interchanging the roles of K_1 and K_2 changes the sign of both the outward unit normal and the quantity inside the parentheses. We extend $\{\cdot\}_{\mathrm{s}}$ to every face in $\mathscr{T}_h$ by defining $\{v\}_{\mathrm{s}}(\mathbf{x}) = 0$ for $\mathbf{x} \in K \cap \Gamma$, $K \in \mathscr{T}_h$. The corresponding maximal jump is defined by

$$[\mathbf{v}]_{\mathrm{s}}(\mathbf{x}) := \max_{\mathbf{y}\in\partial K} |\{\mathbf{v}\}_{\mathrm{s}}(\mathbf{y})|, \quad \mathbf{x} \in K \in \mathscr{T}_h, \quad (10)$$

where ∂K denotes the boundary of K.

For jumps in time, we define

$$\{v\}_{\mathrm{t}}(t_k) := \lim_{s\to 0+} \big(v(t_k + s) - v(t_k - s)\big), \quad (11)$$

with $\{v\}_{\mathrm{t}}(0) = \{v\}_{\mathrm{t}}(T) = 0$, and

$$[v]_{\mathrm{t}}(t) := \max\{|\{v\}_{\mathrm{t}}(t_k)|, |\{v\}_{\mathrm{t}}(t_{k+1})|\} \quad (12)$$

for $t \in (t_k, t_{k+1})$.

In the theorems and proofs to be presented, we will frequently use the symbols $\approx$ and $\lesssim$ to denote approximate equality and inequality, respectively, where higher order terms (with respect to mesh-size or errors) are neglected. We let C denote various constants of moderate size which are independent of the mesh-sizes and the unknown functions.

We now proceed to an error estimation for the coefficient. An error estimate for the Tikhonov functional will follow as a corollary.

Theorem 1. *(A posteriori error estimate for the coefficient.) Suppose that the initial approximation* ε_0 *and the regularization parameter* α *are such that the strong convexity estimate* (5) *holds. Let* $u = (\varepsilon, \mathbf{E}, \boldsymbol{\lambda}) \in U$ *be the solution to the Continuous Saddle Point Problem, and let* $u_h = (\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \in U_h$ *be the solution to the Discrete Saddle Point Problem, computed on meshes* $\mathscr{T}_h$ *and* $\mathscr{I}_\tau$. *Then there exists a constant* C, *which does not depend on* u, u_h, h, *or* τ, *such that*

$$\|\varepsilon - \varepsilon_h\|_\Omega \lesssim \frac{2C}{\alpha}(\eta + \|R_\varepsilon\|_\Omega), \quad (13)$$

where $\eta = \eta(u_h)$ *is defined by*

$$\begin{aligned} \eta := & \left\langle \tfrac{1}{\tau}\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}} + s|\nabla\cdot\boldsymbol{\lambda}_h|, h[\nabla \mathbf{E}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}} \right\rangle_{\Omega_T} \\ & + \left\langle \tfrac{1}{\tau}\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}, h[\nabla \boldsymbol{\lambda}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}} \right\rangle_{\Omega_T} \\ & + s\left\langle |\nabla\cdot\boldsymbol{\lambda}_h|, [\nabla \mathbf{E}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \nabla\cdot\mathbf{E}_h}{\partial t}\right]_{\mathrm{t}} \right\rangle_{\Omega_T} \\ & + s\left\langle |\nabla\cdot\mathbf{E}_h| + |\mathbf{E}_h|, [\nabla \boldsymbol{\lambda}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \nabla\cdot\boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}} \right\rangle_{\Omega_T}, \end{aligned} \quad (14)$$

and

$$\begin{aligned} R_\varepsilon := & \ \alpha(\varepsilon_h - \varepsilon_0) - \int_0^T \frac{\partial \mathbf{E}_h}{\partial t}\cdot\frac{\partial \boldsymbol{\lambda}_h}{\partial t}\,\mathrm{d}t \\ & + \frac{s}{2h}\int_0^T [(\nabla\cdot\boldsymbol{\lambda}_h)\mathbf{E}_h]_{\mathrm{s}}\,\mathrm{d}t. \end{aligned} \quad (15)$$

Proof. Using strong convexity (5), we obtain

$$\|\varepsilon - \varepsilon_h\|_\Omega^2 \le \frac{2}{\alpha}\left(F'(\varepsilon; \varepsilon - \varepsilon_h) - F'(\varepsilon_h; \varepsilon - \varepsilon_h)\right).$$

Since ε minimizes $F(\varepsilon)$ we have $F'(\varepsilon; \varepsilon - \varepsilon_h) = 0$ and thus

$$\begin{aligned} \|\varepsilon - \varepsilon_h\|_\Omega^2 &\le \frac{2}{\alpha}\left|F'(\varepsilon_h; \varepsilon - \varepsilon_h)\right| \\ &= \frac{2}{\alpha}\left|\frac{\partial L}{\partial \varepsilon}(\tilde{u}; \varepsilon - \varepsilon_h)\right|, \end{aligned} \quad (16)$$

where we have denoted by $\tilde{\mathbf{E}}$ and $\tilde{\boldsymbol{\lambda}}$ the solutions to the Continuous Direct Problem and the Continuous Adjoint Problem, respectively, with permittivity ε_h, and set $\tilde{u} = (\varepsilon_h, \tilde{\mathbf{E}}, \tilde{\boldsymbol{\lambda}}) \in U$. The last equality follows from (7). We remark that (16) implies $\|\varepsilon - \varepsilon_h\|_\Omega \leq \frac{2}{\alpha}\|F'(\varepsilon_h)\|$, which would be an effective estimate. Unfortunately, we cannot compute $F'(\varepsilon_h)$ exactly, since it depends on the exact solutions $\tilde{\mathbf{E}}$ and $\tilde{\boldsymbol{\lambda}}$ as indicated in the text; hence further estimation is needed.

We expand, using the triangle inequality,

$$\begin{aligned}\left|\frac{\partial L}{\partial \varepsilon}(\tilde{u}; \varepsilon - \varepsilon_h)\right| \leq \\ \left|\frac{\partial L}{\partial \varepsilon}(\tilde{u}; \varepsilon - \varepsilon_h) - \frac{\partial L}{\partial \varepsilon}(u_h; \varepsilon - \varepsilon_h)\right| \\ + \left|\frac{\partial L}{\partial \varepsilon}(u_h; \varepsilon - \varepsilon_h)\right| \\ := |\Theta_1| + |\Theta_2|,\end{aligned} \tag{17}$$

and estimate the two terms $|\Theta_1|$ and $|\Theta_2|$ separately.

For Θ_1 we assume that the second partial derivatives of L exist, and use the linearization

$$\begin{aligned}\Theta_1 &= \frac{\partial L}{\partial \varepsilon}(\tilde{u}; \varepsilon - \varepsilon_h) - \frac{\partial L}{\partial \varepsilon}(u_h; \varepsilon - \varepsilon_h) \\ &= \frac{\partial^2 L}{\partial \varepsilon^2}(u_h; \varepsilon_h - \varepsilon_h; \varepsilon - \varepsilon_h) + o(\|\varepsilon_h - \varepsilon_h\|_\Omega) \\ &\quad + \frac{\partial^2 L}{\partial \mathbf{E}\partial \varepsilon}(u_h; \tilde{\mathbf{E}} - \mathbf{E}_h; \varepsilon - \varepsilon_h) + o(\|\tilde{\mathbf{E}} - \mathbf{E}_h\|_{H^1(\Omega_T)}) \\ &\quad + \frac{\partial^2 L}{\partial \boldsymbol{\lambda}\partial \varepsilon}(u_h; \tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h; \varepsilon - \varepsilon_h) + o(\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h\|_{H^1(\Omega_T)}),\end{aligned}$$

where $\frac{\partial^2 L}{\partial \mathbf{E}\partial \varepsilon}$ and $\frac{\partial^2 L}{\partial \boldsymbol{\lambda}\partial \varepsilon}$ denote mixed second partial Fréchet derivatives of L. The first two terms vanish, since the first components of $\tilde{u}$ and u_h are both ε_h, and again the remainder terms are neglected as they are of higher order with respect to the error. Thus, after exchanging the order of differentiation, we are left with

$$\begin{aligned}\Theta_1 &\approx \frac{\partial^2 L}{\partial \mathbf{E}\partial \varepsilon}(u_h; \tilde{\mathbf{E}} - \mathbf{E}_h; \varepsilon - \varepsilon_h) \\ &\quad + \frac{\partial^2 L}{\partial \boldsymbol{\lambda}\partial \varepsilon}(u_h; \tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h; \varepsilon - \varepsilon_h) \\ &= D_1|_{\varepsilon-\varepsilon_h}\frac{\partial L}{\partial \mathbf{E}}(u_h; \tilde{\mathbf{E}} - \mathbf{E}_h) \\ &\quad + D_1|_{\varepsilon-\varepsilon_h}\frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; \tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h),\end{aligned} \tag{18}$$

where $D_1|_{\varepsilon-\varepsilon_h}$ denotes differentiation with respect to the first component in u_h and action on $\varepsilon - \varepsilon_h$.

We split $\tilde{\mathbf{E}} - \mathbf{E}_h = (\tilde{\mathbf{E}} - \Pi_h\tilde{\mathbf{E}}) + (\Pi_h\tilde{\mathbf{E}} - \mathbf{E}_h) = r_h\tilde{\mathbf{E}} + (\Pi_h\tilde{\mathbf{E}} - \mathbf{E}_h)$ and use the fact that $\boldsymbol{\lambda}_h$ solves the Discrete Adjoint Problem with coefficient ε_h, so that $\frac{\partial L}{\partial \mathbf{E}}(u_h; \Pi_h\tilde{\mathbf{E}} - \mathbf{E}_h) = 0$ as $\Pi_h\tilde{\mathbf{E}} - \mathbf{E}_h \in V_h^{\mathrm{dir}}$. This gives

$$\begin{aligned}&\frac{\partial L}{\partial \mathbf{E}}(u_h; \tilde{\mathbf{E}} - \mathbf{E}_h) \\ &\quad = \frac{\partial L}{\partial \mathbf{E}}(u_h; r_h\tilde{\mathbf{E}}) + \frac{\partial L}{\partial \mathbf{E}}(u_h; \Pi_h\tilde{\mathbf{E}} - \mathbf{E}_h) \\ &\quad = \frac{\partial L}{\partial \mathbf{E}}(u_h; r_h\tilde{\mathbf{E}}).\end{aligned} \tag{19}$$

Similarly, we have

$$\begin{aligned}&\frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; \tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h) \\ &\quad = \frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; r_h\tilde{\boldsymbol{\lambda}}) + \frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; \Pi_h\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_h) \\ &\quad = \frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; r_h\tilde{\boldsymbol{\lambda}})\end{aligned} \tag{20}$$

as $\mathbf{E}_h$ solves the Discrete Direct Problem with coefficient ε_h.

Combining (18), (19), and (20), and recalling (6) gives

$$\begin{aligned}\Theta_1 &\approx D_1|_{\varepsilon-\varepsilon_h}\left(\frac{\partial L}{\partial \mathbf{E}}(u_h; r_h\tilde{\mathbf{E}}) + \frac{\partial L}{\partial \boldsymbol{\lambda}}(u_h; r_h\tilde{\boldsymbol{\lambda}})\right) \\ &= -\left\langle (\varepsilon - \varepsilon_h)\tfrac{\partial r_h\tilde{\mathbf{E}}}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_{\Omega_T} \\ &\quad + s\left\langle \nabla\cdot((\varepsilon - \varepsilon_h)r_h\tilde{\mathbf{E}}), \nabla\cdot\boldsymbol{\lambda}_h\right\rangle_{\Omega_T} \\ &\quad - \left\langle (\varepsilon - \varepsilon_h)\tfrac{\partial \mathbf{E}_h}{\partial t}, \tfrac{\partial r_h\tilde{\boldsymbol{\lambda}}}{\partial t}\right\rangle_{\Omega_T} \\ &\quad + s\left\langle \nabla\cdot((\varepsilon - \varepsilon_h)\mathbf{E}_h, \nabla\cdot r_h\tilde{\boldsymbol{\lambda}}\right\rangle_{\Omega_T}.\end{aligned}$$

We now aim to lift time derivatives from the interpolation residuals $r_h\tilde{\mathbf{E}}$ and $r_h\tilde{\boldsymbol{\lambda}}$ by splitting the integral over $[0, T]$ into the sum of integrals over the subintervals in $\mathscr{I}_\tau$, and integrating by parts in each subinterval. Thus we obtain

$$\begin{aligned}&\left\langle (\varepsilon - \varepsilon_h)\tfrac{\partial r_h\tilde{\mathbf{E}}}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_{\Omega_T} \\ &\quad = \sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\left\langle (\varepsilon - \varepsilon_h)\tfrac{\partial r_h\tilde{\mathbf{E}}}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_\Omega \mathrm{d}t \\ &\quad = \sum_{k=1}^{N_\tau}\Bigg(-\int_{t_{k-1}}^{t_k}\left\langle (\varepsilon - \varepsilon_h)r_h\tilde{\mathbf{E}}, \tfrac{\partial^2 \boldsymbol{\lambda}_h}{\partial t^2}\right\rangle_\Omega \mathrm{d}t \\ &\qquad + \left\langle (\varepsilon - \varepsilon_h)r_h\tilde{\mathbf{E}}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_\Omega\Bigg|_{t=t_k} \\ &\qquad - \left\langle (\varepsilon - \varepsilon_h)r_h\tilde{\mathbf{E}}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_\Omega\Bigg|_{t=t_{k-1}}\Bigg).\end{aligned}$$

We note that $\frac{\partial^2 \boldsymbol{\lambda}_h}{\partial t^2} \equiv 0$ on each subinterval, since $\boldsymbol{\lambda}_h$ is piecewise linear, and identify the jumps form (11), obtaining

$$\begin{aligned}&\left\langle (\varepsilon - \varepsilon_h)\tfrac{\partial r_h\tilde{\mathbf{E}}}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_{\Omega_T} \\ &\quad = \sum_{k=1}^{N_\tau}\left\langle (\varepsilon - \varepsilon_h)r_h\tilde{\mathbf{E}}, \left\{\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\}_{\mathrm{t}}\right\rangle_\Omega\Bigg|_{t=t_k}.\end{aligned}$$

Next, we use the approximation

$$f(t_k) \approx \frac{1}{\tau}\int_{t_{k-1}}^{t_k} f(t)\,\mathrm{d}t$$

to obtain terms defined in the whole time-interval. That is

$$\begin{aligned}
&\left\langle (\varepsilon-\varepsilon_h)\tfrac{\partial r_h\tilde{\mathbf{E}}}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_{\Omega_T}\\
&\quad\approx \sum_{k=1}^{N_\tau}\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\left\langle (\varepsilon-\varepsilon_h) r_h\tilde{\mathbf{E}}, \left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}\right\rangle_{\Omega}\mathrm{d}t\\
&\quad= \left\langle (\varepsilon-\varepsilon_h) r_h\tilde{\mathbf{E}}, \tfrac{1}{\tau}\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}\right\rangle_{\Omega_T}.
\end{aligned}$$

Similarly, we have

$$\left\langle (\varepsilon-\varepsilon_h)\tfrac{\partial \mathbf{E}_h}{\partial t}, \tfrac{\partial r_h\tilde{\boldsymbol{\lambda}}}{\partial t}\right\rangle_{\Omega_T} = \left\langle (\varepsilon-\varepsilon_h)\tfrac{1}{\tau}\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}, r_h\tilde{\boldsymbol{\lambda}}\right\rangle_{\Omega_T}.$$

Thus

$$\begin{aligned}
|\Theta_1| \lesssim{}& \left\langle |\varepsilon-\varepsilon_h|\,\tfrac{1}{\tau}\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}, \left|r_h\tilde{\mathbf{E}}\right|\right\rangle_{\Omega_T}\\
&+ s\left\langle \left|\nabla\cdot((\varepsilon-\varepsilon_h)r_h\tilde{\mathbf{E}})\right|, |\nabla\cdot\boldsymbol{\lambda}_h|\right\rangle_{\Omega_T}\\
&+ \left\langle |\varepsilon-\varepsilon_h|\,\tfrac{1}{\tau}\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}, \left|r_h\tilde{\boldsymbol{\lambda}}\right|\right\rangle_{\Omega_T}\\
&+ s\left\langle |\nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h)|, \left|\nabla\cdot r_h\tilde{\boldsymbol{\lambda}}\right|\right\rangle_{\Omega_T}\\
\le{}& \|\varepsilon-\varepsilon_h\|_{L_\infty(\Omega)}\left\langle \tfrac{1}{\tau}\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}, \left|r_h\tilde{\mathbf{E}}\right|\right\rangle_{\Omega_T}\\
&+ \|\varepsilon-\varepsilon_h\|_{L_\infty(\Omega)}\left\langle \tfrac{1}{\tau}\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}, \left|r_h\tilde{\boldsymbol{\lambda}}\right|\right\rangle_{\Omega_T}\\
&+ s\,\|\varepsilon-\varepsilon_h\|_{L_\infty(\Omega)}\left\langle \left|\nabla\cdot r_h\tilde{\mathbf{E}}\right|, |\nabla\cdot\boldsymbol{\lambda}_h|\right\rangle_{\Omega_T}\\
&+ s\,\|\varepsilon-\varepsilon_h\|_{L_\infty(\Omega)}\left\langle |\nabla\cdot\mathbf{E}_h|, \left|\nabla\cdot r_h\tilde{\boldsymbol{\lambda}}\right|\right\rangle_{\Omega_T}\\
&+ s\,\|\nabla(\varepsilon-\varepsilon_h)\|_{L_\infty(\Omega)}\left\langle \left|r_h\tilde{\mathbf{E}}\right|, |\nabla\cdot\boldsymbol{\lambda}_h|\right\rangle_{\Omega_T}\\
&+ s\,\|\nabla(\varepsilon-\varepsilon_h)\|_{L_\infty(\Omega)}\left\langle |\mathbf{E}_h|, \left|\nabla\cdot r_h\tilde{\boldsymbol{\lambda}}\right|\right\rangle_{\Omega_T}.
\end{aligned} \tag{21}$$

Note that, by Theorem 4.5.11 of [10], there is a positive constant $C \propto h_{\mathrm{f}}^{-5/2}$ (in the notation of the cited theorem, the exponent is $m-l+\min\{0, n/p-n/q\}$, where in our case $m=0$, $l=1$, $n=3$, $p=\infty$, and $q=2$), where h_{f} is the mesh size in V^{per}, such that

$$\|\varepsilon-\varepsilon_h\|_{L_\infty(\Omega)} + \|\nabla(\varepsilon-\varepsilon_h)\|_{L_\infty(\Omega)} \le C\,\|\varepsilon-\varepsilon_h\|_{\Omega}.$$

Finally, we use standard interpolation estimates (see for instance [26]) for $r_h\tilde{\mathbf{E}}$

$$\begin{aligned}
\left|r_h\tilde{\mathbf{E}}\right| &\le C\left(h^2\left|D^2\tilde{\mathbf{E}}\right| + \tau^2\left|\frac{\partial^2\mathbf{E}}{\partial t^2}\right|\right)\\
&\approx C\left(h^2\left|h^{-1}[\nabla\mathbf{E}_h]_{\mathrm{s}}\right| + \tau^2\left|\tau^{-1}\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}\right|\right)\\
&\lesssim C\left(h[\nabla\mathbf{E}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}\right),
\end{aligned}$$

as well as for $r_h\tilde{\boldsymbol{\lambda}}$, $\nabla\cdot r_h\tilde{\mathbf{E}}$, and $\nabla\cdot r_h\tilde{\boldsymbol{\lambda}}$

$$\begin{aligned}
\left|r_h\tilde{\boldsymbol{\lambda}}\right| &\lesssim C\left(h[\nabla\boldsymbol{\lambda}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}\right),\\
\left|\nabla\cdot r_h\tilde{\mathbf{E}}\right| &\lesssim C\left([\nabla\mathbf{E}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \nabla\cdot\mathbf{E}_h}{\partial t}\right]_{\mathrm{t}}\right),\\
\left|\nabla\cdot r_h\tilde{\boldsymbol{\lambda}}\right| &\lesssim C\left([\nabla\boldsymbol{\lambda}_h]_{\mathrm{s}} + \tau\left[\tfrac{\partial \nabla\cdot\boldsymbol{\lambda}_h}{\partial t}\right]_{\mathrm{t}}\right),
\end{aligned}$$

where the jumps $[\cdot]_{\mathrm{s}}$ and $[\cdot]_{\mathrm{t}}$ were defined in equations (9), (10), (11), and (12). Applying these estimates in (21), we get

$$\Theta_1 \lesssim C\eta\,\|\varepsilon-\varepsilon_h\|_{\Omega}, \tag{22}$$

with η as in (14).

Turning to Θ_2 of (17), we recall from (6) that

$$\begin{aligned}
\Theta_2 &= \frac{\partial L}{\partial \varepsilon}(u_h; \varepsilon-\varepsilon_h)\\
&= \alpha\,\langle \varepsilon_h-\varepsilon_0, \varepsilon-\varepsilon_h\rangle_{\Omega}\\
&\quad - \left\langle (\varepsilon-\varepsilon_h)\tfrac{\partial \mathbf{E}_h}{\partial t}, \tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\right\rangle_{\Omega_T}\\
&\quad + s\,\langle \nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h), \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\Omega_T}.
\end{aligned} \tag{23}$$

Starting with the last term in the above expression, we split the integral over Ω into the sum of integrals over $K\in\mathscr{T}_h$, and integrate by parts to get

$$\begin{aligned}
&\langle \nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h), \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\Omega_T}\\
&\quad= \sum_{K\in\mathscr{T}_h}\langle \nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h), \nabla\cdot\boldsymbol{\lambda}_h\rangle_{K_T}\\
&\quad= \sum_{K\in\mathscr{T}_h}\Big(-\langle (\varepsilon-\varepsilon_h)\mathbf{E}_h, \nabla(\nabla\cdot\boldsymbol{\lambda}_h)\rangle_{K_T}\\
&\qquad\qquad + \langle (\varepsilon-\varepsilon_h)\mathbf{n}\cdot\mathbf{E}_h, \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\partial K_T}\Big),
\end{aligned}$$

where $K_T := K\times(0,T)$, and $\partial K_T := (\partial K)\times(0,T)$. we observe that $\nabla\cdot(\nabla\boldsymbol{\lambda}_h)\equiv 0$ on each $K\in\mathscr{T}_h$ for the piecewise linear $\boldsymbol{\lambda}_h$, and identify the jumps from (9), obtaining

$$\begin{aligned}
&\langle \nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h), \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\Omega_T}\\
&\quad= \sum_{K\in\mathscr{T}_h}\langle (\varepsilon-\varepsilon_h)\mathbf{n}\cdot\mathbf{E}_h, \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\partial K_T}\\
&\quad= \frac{1}{2}\sum_{K\in\mathscr{T}_h}\langle \{(\nabla\cdot\boldsymbol{\lambda}_h)\mathbf{E}_h\}_{\mathrm{s}}, \varepsilon-\varepsilon_h\rangle_{\partial K_T},
\end{aligned}$$

where the factor $\frac{1}{2}$ appears since every non-zero jump is encountered exactly twice in the sum over $K\in\mathscr{T}_h$.

Seeking to obtain expressions defined over the whole of Ω, as opposed to those defined only on boundaries of elements $K\in\mathscr{T}_h$, we use the following approximation, similar to the one used above for the jumps in time:

$$\int_{\partial K} f\,\mathrm{d}S \approx \int_K \frac{f}{h_K}\,\mathrm{d}\mathbf{x}.$$

This yields

$$\begin{aligned}&\langle \nabla\cdot((\varepsilon-\varepsilon_h)\mathbf{E}_h), \nabla\cdot\boldsymbol{\lambda}_h\rangle_{\Omega_T}\\ &\approx \frac{1}{2}\sum_{K\in\mathscr{T}_h}\left\langle \tfrac{1}{h_K}[(\nabla\cdot\boldsymbol{\lambda}_h)\mathbf{E}_h]_s, \varepsilon-\varepsilon_h\right\rangle_{K_T}\\ &= \left\langle \tfrac{1}{2h}[(\nabla\cdot\boldsymbol{\lambda}_h)\mathbf{E}_h]_s, \varepsilon-\varepsilon_h\right\rangle_{\Omega_T}.\end{aligned}$$

With the above estimate in (23), we can now conclude that

$$\begin{aligned}|\Theta_2| \lesssim \Big| &\langle \alpha(\varepsilon_h-\varepsilon_0), \varepsilon-\varepsilon_h\rangle_\Omega\\ &-\left\langle \int_0^T \tfrac{\partial \mathbf{E}_h}{\partial t}\cdot\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t}\,\mathrm{d}t, \varepsilon-\varepsilon_h\right\rangle_\Omega\\ &+\left\langle \tfrac{s}{2h}\int_0^T [(\nabla\cdot\boldsymbol{\lambda}_h)\mathbf{E}_h]_s\,\mathrm{d}t, \varepsilon-\varepsilon_h\right\rangle_\Omega\Big|\\ \le\ & \|R_\varepsilon\|_\Omega\,\|\varepsilon-\varepsilon_h\|_\Omega,\end{aligned} \tag{24}$$

with R_ε as defined (15).

Combining estimates (22) and (24) with (16) and (17), we conclude that

$$\|\varepsilon-\varepsilon_h\|_\Omega^2 \lesssim \frac{2C}{\alpha}\left(\eta\,\|\varepsilon-\varepsilon_h\|_\Omega + \|R_\varepsilon\|_\Omega\,\|\varepsilon-\varepsilon_h\|_\Omega\right),$$

and the result (13) follows.

We see that if the numerical errors for solving the direct and adjoint problems are relatively small, that is, when $\tilde{u}\approx u_h$ with relatively high accuracy so that the jump and divergence terms in (14) are small, then $\|R_\varepsilon\|_\Omega \approx \|F'(\varepsilon_h)\|$ dominates the error estimate.

Corollary 1. *(A posteriori error estimate for the Tikhonov functional.) Under the hypothesis of Theorem 1, we have*

$$|F(\varepsilon)-F(\varepsilon_h)| \lesssim \frac{4C^2}{\alpha^2}\left(\eta^2+\|R_\varepsilon\|_\Omega^2\right),$$

with η as defined in (14), and R_ε as in (15).

Proof. Using the definition of the Fréchet derivative and (7), we get

$$\begin{aligned}F(\varepsilon)-F(\varepsilon_h) &= F'(\varepsilon_h;\varepsilon-\varepsilon_h)+o(\|\varepsilon-\varepsilon_h\|_\Omega^2)\\ &= \frac{\partial L}{\partial \varepsilon}(\tilde{u};\varepsilon-\varepsilon_h)+o(\|\varepsilon-\varepsilon_h\|_\Omega^2).\end{aligned}$$

Neglecting the remainder term as it is of higher order with respect to the error, and estimating $\frac{\partial L}{\partial \varepsilon}(\tilde{u};\varepsilon-\varepsilon_h)=\Theta_1+\Theta_2$ as in the proof of Theorem 1, we obtain

$$|F(\varepsilon)-F(\varepsilon_h)| \lesssim \frac{2C}{\alpha}(\eta+\|R_\varepsilon\|_\Omega)\,\|\varepsilon-\varepsilon_h\|_\Omega.$$

Applying Theorem 1 to estimate $\|\varepsilon-\varepsilon_h\|_\Omega$, we arrive at

$$|F(\varepsilon)-F(\varepsilon_h)| \lesssim \frac{4C^2}{\alpha^2}(\eta+\|R_\varepsilon\|_\Omega)^2 \le C\left(\eta^2+\|R_\varepsilon\|_\Omega^2\right).$$

We will conclude this section by describing how these theorems can be translated into concrete recommendations for refining the computational mesh in the adaptive algorithm, outlined in Section 4.

3.1 Mesh refinement recommendations

From Theorem 1 and Corollary 1, it is clear that the error in the reconstruction can be estimated in terms of two quantities, η, and R_ε. The former essentially represents how well the Continuous Direct and Adjoint Problems are approximated by their finite element counterparts, the Discrete Direct and Adjoint Problems, respectively. The latter, R_ε, describes the error incurred by the approximation of the coefficient itself.

In our computational experience from [7,8], on given meshes $\mathscr{T}_h$ and $\mathscr{I}_\tau$, the solutions to direct and adjoint problems are in general approximated better than the coefficient itself. As remarked above, this implies that $\eta \ll \|R_\varepsilon\|_\Omega$, and thus contributions to the error are the greatest in regions where $|R_\varepsilon| \approx |F'(\varepsilon_h)|$ is close to its maximum value. Thus we propose the following mesh refinement recommendation:

Mesh Refinement Recommendation 1 *Using Theorem 1 we conclude that we should refine the mesh in neighborhoods of those points in Ω where the function $|R_\varepsilon|$ attains its maximal values. More precisely, let $\beta\in(0,1)$ be a tolerance number which should be chosen in computational experiments. Then, refine the mesh $\mathscr{T}_h$ in such subdomains of Ω where*

$$|R_\varepsilon| \ge \beta \max_\Omega |R_\varepsilon|.$$

Since a relatively large value of the reconstructed coefficient ε_h indicates a region where the permittivity is different from the background value of 1, we can also propose the following heuristic:

Mesh Refinement Recommendation 2 *We should refine the mesh in neighborhoods of those points in Ω where the function $|\varepsilon_h|$ attains its maximal values. More precisely, we refine the mesh in such subdomains of Ω where*

$$|\varepsilon_h| \ge \widetilde{\beta}\max_\Omega |\varepsilon_h|,$$

where $\widetilde{\beta}\in(0,1)$ is a number which should be chosen computationally.

4 Adaptive algorithms for the inverse problem

In this section we will present different algorithms which can be used for the solution of the inverse problem we consider: usual conjugate gradient algorithm and two different adaptive finite element algorithms. Conjugate gradient algorithm is applied on every finite element mesh $\mathscr{T}_h$ which we use in computations. We note that in our adaptive algorithms the time mesh $\mathscr{I}_\tau$ is refined globally accordingly to the Courant-Friedrichs-Lewy condition of [27].

Taking into account the remark of Section 3.1 we denote by

$$R_\varepsilon^n(\mathbf{x}) := \alpha(\varepsilon_h - \varepsilon_0) - \int_0^T \frac{\partial \mathbf{E}_h^n}{\partial t}(\mathbf{x},t)\cdot\frac{\partial \boldsymbol{\lambda}_h^n}{\partial t}(\mathbf{x},t)\,\mathrm{d}t + s\int_0^T \nabla\cdot\mathbf{E}_h^n(\mathbf{x},t)\nabla\cdot\boldsymbol{\lambda}_h^n(\mathbf{x},t)\,\mathrm{d}t, \tag{25}$$

where functions $\boldsymbol{\lambda}_h^n$, and $\mathbf{E}_h^n$ are finite element solutions of direct and adjoint problems computed with $\varepsilon := \varepsilon_h^n$, respectively, and n is the iteration number in the conjugate gradient algorithm.

4.1 Conjugate Gradient Algorithm

Here we outline the conjgate gradient algorithm, which will be used in the two adaptive algorithms presented below.

Algorithm 1 *(Conjugate Gradient Algorithm)*
Step 0. Discretize the computational space-time domain $\Omega\times[0,T]$ using partitions $\mathscr{T}_h$ and $\mathscr{I}_\tau$, respectively, see Section 3. Start with the initial approximations $\varepsilon_h^0 = \varepsilon_0$ and compute the sequence of ε_h^n, $n = 1,2,\ldots$, as:
Step 1. Compute solutions $\mathbf{E}_h^n$ and $\boldsymbol{\lambda}_h^n$ of the Discrete Direct and Adjoint Problems, respectively, using the coefficient ε_h^n.

Step 2. Update the coefficient on $\mathscr{T}_h$ and $\mathscr{I}_\tau$ via the conjugate gradient method

$$\varepsilon_h^{n+1}(\mathbf{x}) = \varepsilon_h^n(\mathbf{x}) + \gamma_\varepsilon^n d_\varepsilon^n(\mathbf{x}),$$

where

$$d_\varepsilon^n(\mathbf{x}) = -R_\varepsilon^n(\mathbf{x}) + \beta_\varepsilon^n d_\varepsilon^{n-1}(\mathbf{x}),$$

with

$$\beta_\varepsilon^n = \frac{\|R_\varepsilon^n\|_\Omega^2}{\|R_\varepsilon^{n-1}\|_\Omega^2},$$

$d_\varepsilon^0(\mathbf{x}) = -R_\varepsilon^0(\mathbf{x})$, and γ_ε^n are step-sizes in the gradient update which can be computed as in [28]

$$\gamma_\varepsilon^n = -\frac{\langle R_\varepsilon^n, d_\varepsilon^n\rangle_\Omega}{\alpha\,\|d_\varepsilon^n\|_\Omega^2}. \tag{26}$$

Step 3. Stop computing updates ε_h^n at the iteration $M := n$ and obtain the function $\varepsilon_h := \varepsilon_h^M$ if either $\|R_\varepsilon^n\|_\Omega \le \theta$, where θ is the tolerance in n updates of the gradient method, or norms $\|\varepsilon_h^n\|_\Omega$ are stabilized. Otherwise update n to $n+1$ and return to Step 1.

4.2 Adaptive algorithms

In this section we present two different adaptive algorithms for the solution of our coefficient inverse problem (more precisely, the Discrete Saddle Point Problem, approximating the solution to the Continuous Saddle Point Problem), where in the first adaptive algorithm we apply Mesh Refinement Recommendation 1 of Section 3.1, while in the second adaptive algorithm we use Mesh Refinement Recommendation 2 of Section 3.1.

We define the minimizer of the Tikhonov functional (4) and its approximated finite element solution on k times adaptively refined mesh $\mathscr{T}_{h_k}$ by ε and ε_{h_k}, correspondingly. The latter is obtained at the final step of the conjugate gradient iteration of Section 4.1 on the mesh $\mathscr{T}_{h_k}$.

Algorithm 2 *(The First Adaptive Algorithm)*
Step 0. Choose an initial space-time mesh $\mathscr{T}_{h_0}\times\mathscr{I}_{\tau_0}$ in $\Omega\times[0,T]$. Compute ε_{h_k}, $k>0$, via following steps:

Step 1. Obtain numerical solution ε_{h_k} on $\mathscr{T}_{h_k}$ using the Conjugate Gradient Method of Section 4.1. Denote the number of conjugate gradient steps by M_k.

Step 2. In accordance with the first mesh refinement recommendation, refine such elements in the mesh $\mathscr{T}_{h_k}$ where the expression

$$\left|R_{\varepsilon,k}^M\right| \ge \beta_k \max_\Omega \left|R_{\varepsilon,k}^M\right|, \tag{27}$$

where $R_{\varepsilon,k}^M$ is the gradient on the last iteration of the conjugate gradient method on the k times adaptively refined meshis satisfied, and the tolerance numbers $\beta_k\in(0,1)$ are chosen by the user.

Step 3. Define a new refined mesh as $\mathscr{T}_{h_{k+1}}$ and construct a new time partition $\mathscr{I}_{\tau_{k+1}}$ such that the Courant-Friedrichs-Lewy condition of [27] is satisfied. Interpolate ε_{h_k} on the new mesh $\mathscr{T}_{h_{k+1}}$, perform Steps 1–3 on the space-time mesh $\mathscr{T}_{h_{k+1}}\times\mathscr{I}_{\tau_{k+1}}$, and end up with $\varepsilon_{h_{k+1}}$.

Step 4. Stop mesh refinements when either $\left\|\varepsilon_{h_k}-\varepsilon_{h_{k-1}}\right\|_\Omega < \theta_1$, or for some n, $\|R_\varepsilon^n\|_\Omega < \theta_2$, where θ_i, $i=1,2$ are tolerances chosen by the user, and R_ε^n is the gradient on the n:th iteration of the conjugate gradient method on the new mesh. We then set the final number of refinements $k_{\mathrm{rec}} := k$, and the reconstructed coefficient $\varepsilon_{\mathrm{rec}} := \varepsilon_{h_k}$.

Algorithm 3 *(The Second Adaptive Algorithm)*
This algorithm follows the same procedure as the First Adaptive Algorithm, except that the refinement criterion (27) is replaced, in accordance with the second mesh refinement recommendation, by

$$\left|\varepsilon_{h_k}\right| \ge \widetilde{\beta}_k \max_\Omega \left|\varepsilon_{h_k}\right| \tag{28}$$

for some tolerance numbers $\widetilde{\beta}_k \in (0,1)$, possibly different from β_k, chosen by the user.

Before continuing with the details of our numerical examples, some remarks are in order:

–Firstly, we comment on how to choose the tolerance numbers β_k, and $\widetilde{\beta}_k$ in (27), (28). Their values depend on the concrete values of $\max_\Omega \left|R^M_{\varepsilon,k}\right|$ and $\max_\Omega \left|\varepsilon_{h_k}\right|$, correspondingly. If we take values of β_k and $\widetilde{\beta}_k$ very close to 1 then we will refine the mesh in very a narrow region of the domain Ω, and if we will choose β_k, and $\widetilde{\beta}_k \approx 0$ then almost all elements in the finite element mesh will be refined. Thus, we will get global and not local mesh refinement. Our numerical tests of Section 5 show that the choice of β_k, and $\widetilde{\beta}_k = 0.7$ is a quasi-optimal one, since with these values of the parameters β_k, and $\widetilde{\beta}_k$ the finite element mesh $\mathscr{T}_{h_k}$ is refined exactly at the places, where we have computed non-trivial parts of the functions ε_{h_k}.

–Secondly, to compute norms $\left\|\varepsilon_{h_k} - \varepsilon_{h_{k-1}}\right\|_\Omega$ in Step 3 of the adaptive algorithms, the solution $\varepsilon_{h_{k-1}}$ is interpolated from the mesh $\mathscr{T}_{h_{k-1}}$ to the mesh $\mathscr{T}_{h_k}$.

5 Numerical examples

In this section we present numerical studies of the solution of our inverse problem using the adaptive algorithms of Section 4.2. The algorithms are efficiently implemented in the software package WavES ([29]), using the domain decomposition technique of [30].

To do that we here consider a rectangular domain Ω, which we enlarge to another rectangular domain $\widetilde{\Omega} \supset \Omega$. We then decompose $\widetilde{\Omega}$ into two subregions Ω_{FEM} and Ω_{FDM} such that $\Omega_{\text{FEM}} = \Omega$, and $\Omega_{\text{FEM}} \cap \Omega_{\text{FDM}} = \emptyset$. In Ω_{FEM} we will use the finite element method (FEM) and, in Ω_{FDM}, the finite difference method (FDM). The boundary $\partial\widetilde{\Omega}$ of the domain $\widetilde{\Omega}$ is such that $\partial\widetilde{\Omega} = \partial_1\widetilde{\Omega} \cup \partial_2\widetilde{\Omega} \cup \partial_3\widetilde{\Omega}$, where $\partial_1\widetilde{\Omega}$ and $\partial_2\widetilde{\Omega}$ are, respectively, front and back sides of $\widetilde{\Omega}$, and $\partial_3\widetilde{\Omega}$ is the union of left, right, top and bottom sides of this domain. We will collect time-dependent observations over $S_T := \partial_1\widetilde{\Omega} \times (0,T)$ at the backscattering side $\partial_1\widetilde{\Omega}$ of $\widetilde{\Omega}$. We also define $S_{1,1} := \partial_1\widetilde{\Omega} \times (0,t_1]$, $S_{1,2} := \partial_1\widetilde{\Omega} \times (t_1,T)$, $S_2 := \partial_2\widetilde{\Omega} \times (0,T)$ and $S_3 := \partial_3\widetilde{\Omega} \times (0,T)$.

As in [7,8] we initialize only the component E_2 of the electric field $\mathbf{E} = (E_1, E_2, E_3)$ on S_T as a plane wave $f(t)$ such that

$$f(t) = \begin{cases} \sin(\omega t) & \text{if } 0 < t < 2\pi/\omega, \\ 0 & \text{if } t > 2\pi/\omega. \end{cases} \tag{29}$$

We assume that the function $\varepsilon \equiv 1$ inside Ω_{FDM}. The numerical tests of our previous studies [31] show that the best reconstruction results are obtained for $\omega = 40$ in (29). Thus, we perform our tests with $\omega = 40$ in (29).

In our computations, we consider computational domains $\widetilde{\Omega} := (-0.8, 0.8)^3$, and $\Omega_{\text{FEM}} := (-0.7, 0.7)^3$, where the length scales are in decimeters. We choose the mesh size $h_0 = 0.05$ in the overlapping layers between Ω_{FEM} and Ω_{FDM} as well as in the coarse mesh. We note that we have generated our transmitted data using a four times locally refined initial mesh Ω_{FEM}, and in a such way we avoid variational crimes. To generate transmitted data we solve the model problem in time $T = [0, 3.0]$ with the time step $\tau = 0.006$ which satisfies to the Courant-Friedrichs-Lewy condition [27]. We then pollute our data additive noise to levels $\sigma = 3\%$, and 10%, respectively.

Similarly to [5,7,8,30,31] in all our computations we choose a constant regularization parameter $\alpha = 0.01$ because it gives the smallest relative error in the reconstruction of the function ε. This parameter was chosen via trial and error because of our computational experience: such a choice of the regularization parameter gave the smallest relative error $e_\varepsilon = \left\|\varepsilon - \varepsilon_{h_k}\right\|_\Omega / \left\|\varepsilon_{h_k}\right\|_\Omega$. An iteratively regularized adaptive finite element method when both dielectric permittivity and magnetic permeability are reconstructed, has recently been presented in [23]. Here, iterative regularization is performed via the algorithms of [22]. We also refer to [15, 16] for different techniques for the choice of regularization parameters.

We perform four different tests:

Test 1: The goal of this numerical test is to reconstruct a smooth function ε only inside Ω_{FEM}. We define this function for $\mathbf{x} \in \Omega_{\text{FEM}}$ as

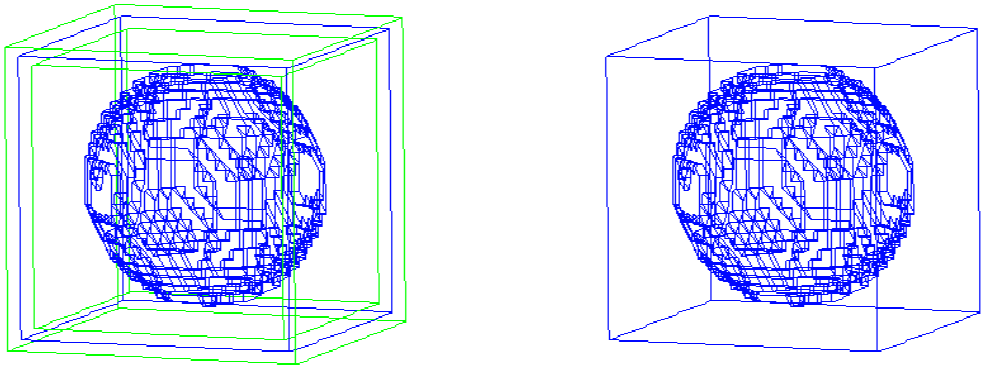
$$\begin{aligned} \varepsilon(\mathbf{x}) &:= 1.0 + 1.0\mathrm{e}^{-|\mathbf{x}-\mathbf{x}_1|^2/0.2} + 1.0\mathrm{e}^{-|\mathbf{x}-\mathbf{x}_2|^2/0.2}, \\ \mathbf{x}_1 &:= (0.3, 0.0, 0.0) \in \Omega_{\text{FEM}}, \\ \mathbf{x}_2 &:= (-0.4, 0.2, 0.0) \in \Omega_{\text{FEM}}. \end{aligned} \tag{30}$$

Test 2: In this test we reconstruct three small inclusions of diameter $d = 2$ mm with the centers of the inclusions at $(-0.3, 0.0, -0.25)$, $(0.3, 0.2, -0.25)$ and $(0.3, -0.2, -0.25)$, respectively, and $\varepsilon = 2.0$ inside the inclusions.

Test 3: In this test we reconstruct four small inclusions of diameter $d = 2$ mm with the centers of the inclusions at $(-0.3, 0.0, 0.25)$, $(0.0, 0.2, 0.25)$, $(0.0, -0.2, -0.25)$, and $(0.3, -0.2, -0.25)$, respectively, and $\varepsilon = 2.0$ inside the inclusions.

Test 4: The inclusions of this test are the same as in Test 3, but here the data consists of measurements of two backscattered wave fields: one backscattered field initiated at the front boundary $\partial_1\widetilde{\Omega}$, and another one at the back boundary $\partial_2\widetilde{\Omega}$.

We start to run the adaptive algorithm with a homogeneous initial approximation $\varepsilon_0 \equiv 1.0$ in Ω_{FEM}. In our recent work [31], it was shown that such choice of the



a) Test1: $\widetilde{\Omega} = \Omega_{\text{FEM}} \cup \Omega_{\text{FDM}}$ b) Test 1: Ω_{FEM}

Fig. 2: *Domain decomposition in numerical tests of Section 5. a) The decomposed domain* $\widetilde{\Omega} = \Omega_{\text{FEM}} \cup \Omega_{\text{FDM}}$. *b) The finite element domain* Ω_{FEM}.

initial approximation gives a good reconstruction. Such homogeneous initial approximations were also used in [32]. See also [5,7,8,30] for a similar choice of initial approximations. We also assume the upper bound $\varepsilon_{\max} = 5$ in (1). This is a reasonable value, given that our target applications are in medical imaging, and typically involves low contrasts.

To get final images of our reconstructed function ε_{h_k} we use a post-processing procedure which has been described before in [5,7,30,31]. The post-processing allows us to identify distinct shapes from the continuous reconstruction ε_{rec}.

5.1 Reconstructions

5.1.1 Test 1

In this section we present numerical results of the reconstruction of the function ε given by (30). Tables 1–2 present computed results of the reconstructions on adaptively refined meshes after applying the First Adaptive Algorithm. Figures 3–6 display results of the reconstruction of the function given by (30) with additive noise of the level $\sigma = 10\%$. Quite similar results are obtained for $\sigma = 3\%$, see Tables 1, and 2, and thus they are not presented here. In Figures 3–6 we observe that the location of the maximal value of the function (30) is imaged correctly. It follows from Figure 5 and Table 1 that on the coarse mesh we obtain good contrast, with $\max_{\Omega_{\text{FEM}}} \varepsilon_{h_0} = 1.94$. However, Figure 5 reveals that it is desirable to improve the location of the maxima of the reconstructed function in x_3 direction as well as remove some artifacts which appeared in the reconstruction on the coarse mesh.

The reconstruction ε_{rec} of ε on a final, five times adaptively refined mesh are presented in Figure 6. We observe significant improvement of the reconstruction of the function ε obtained on the final adaptively refined mesh: the artifacts are removed and the reconstructed function is moved more closer to the exact function in x_3 direction, compared with results of Figure 5.

5.1.2 Test 2

In the test of this section we consider the problem arisen in microwave imaging: reconstruction of high contrast in malign tumors and detection of small tumor sizes (less than 1 cm) in breast cancer screening. Most of existing numerical methods for solution of these problems uses minimization of conventional least-squares functionals and Gauss-Newton methods. See, for example, [1,33,34,35,36,37,38]. We propose to use an adaptive finite element method which will allow achievement of high contrast in the malign tumor and efficiently detect very small sizes of inclusions during adaptive mesh refinement. Although only reconstruction of a real dielectric permittivity function is considered in our test, the obtained results can be extended for the reconstruction of the complex permittivity function which is one of the goals of our current research.

We tested the Second Adaptive Algorithm on the reconstruction of three small inclusions with centers at $(-0.3, 0.0, -0.25)$, $(0.3, 0.2, -0.25)$, and $(0.3, -0.2, -0.25)$ located inside spherical geometry of Figure 2, see also Figure 9 where they are visualized. These inclusions model three small malign tumors of a size 2 mm. We performed simulations with two additive noise levels in the data: $\sigma = 3\%$ and $\sigma = 10\%$, see Tables 1–2 for the results.

The reconstruction of the three inclusions on the initial coarse mesh with $\sigma = 10\%$ is presented on the left figures of Figure 9. In this figure and Table 1 we observe

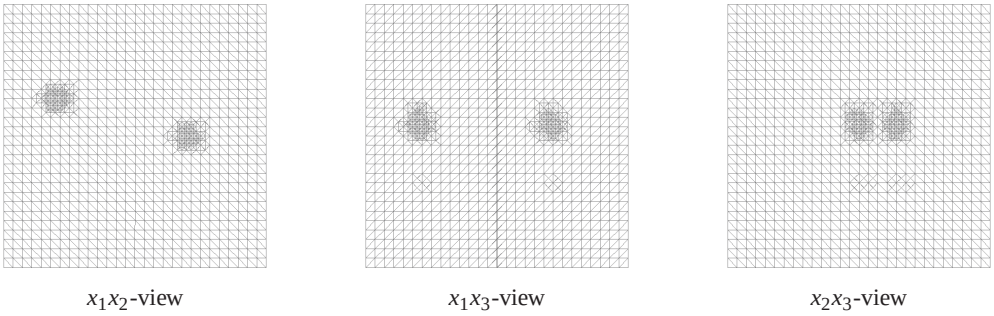


Fig. 3: *Test 1. Five times adaptively refined mesh when the level of the noise in the data was* $\sigma = 10\%$.

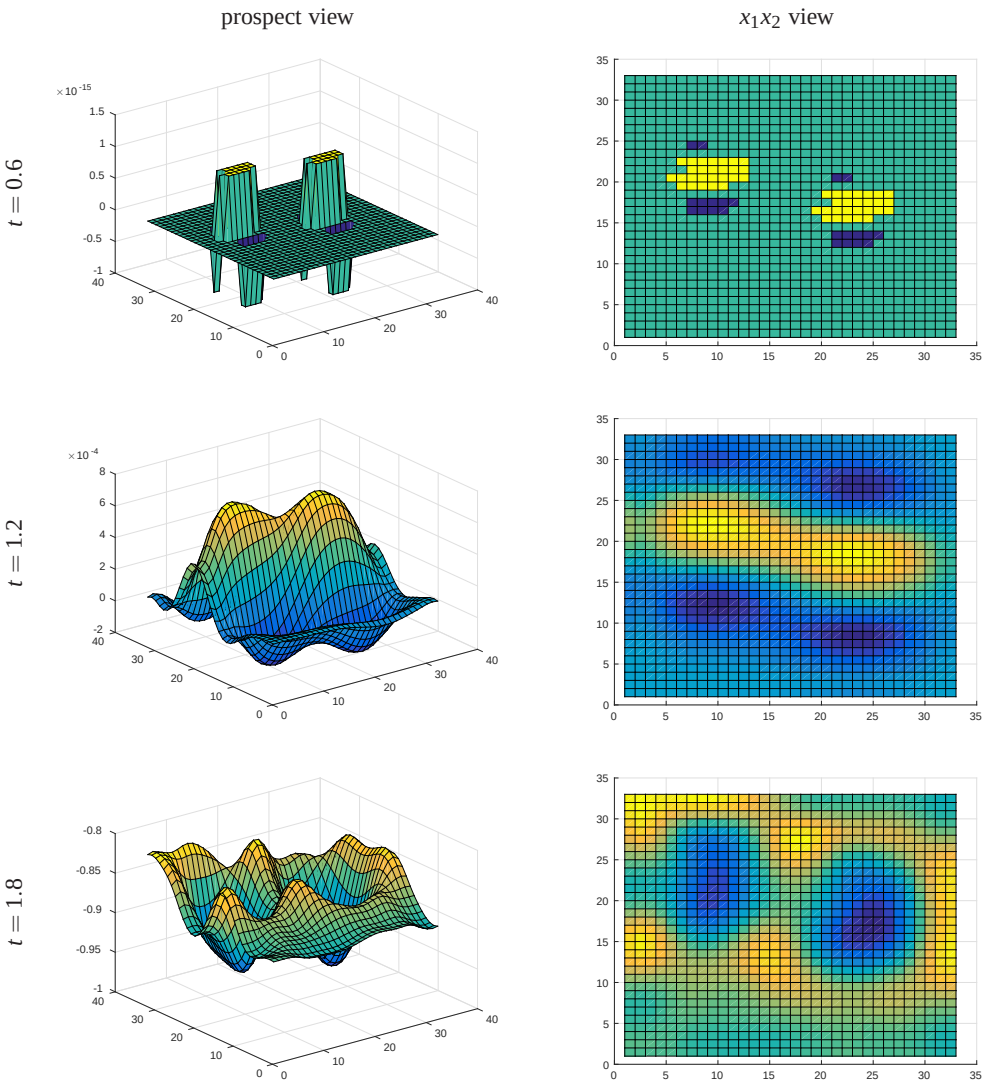


Fig. 4: *Test 1. Transmitted data of component* E_2 *at different times. The noise in the data is* $\sigma = 10\%$.

Table 1. *Results obtained on the coarse mesh. We present reconstructions of the maximal contrast* $\tilde{\varepsilon} = \max_{\Omega_{\mathrm{FEM}}} \varepsilon_{h_0}^{M_0}$ *together with computational errors in percents. Here,* M_0 *is the final number of iteration in the conjugate gradient method on the coarse mesh.*

$\sigma = 3\%$				$\sigma = 10\%$			
	$\tilde{\varepsilon}$	error, %	M_0		$\tilde{\varepsilon}$	error, %	M_0
Test 1	1.93	3.5	2	Test 1	1.94	3	2
Test 2	2.94	47	2	Test 2	2.81	40.5	2
Test 3	1.77	11.5	2	Test 3	2.04	2	2
Test 4	1.9	5	2	Test 4	2.03	1.5	2

Table 2. *Results obtained on* k_{rec} *times adaptively refined mesh. We present reconstructions of the maximal contrast* $\tilde{\varepsilon} = \max_{\Omega_{\mathrm{FEM}}} \varepsilon_{\mathrm{rec}}$ *together with computational errors in percents.*

$\sigma = 3\%$					$\sigma = 10\%$				
Case	$\tilde{\varepsilon}$	error, %	$M_{k_{\mathrm{rec}}}$	k_{rec}	Case	$\tilde{\varepsilon}$	error, %	$M_{k_{\mathrm{rec}}}$	k_{rec}
Test 1	2.04	2	5	4	Test 1	1.97	1.5	1	5
Test 2	1.99	0.5	1	5	Test 2	1.92	4	1	5
Test 3	1.55	22.5	1	2	Test 3	1.88	6	1	2
Test 4	1.9	5	1	1	Test 4	2.15	7.5	1	1

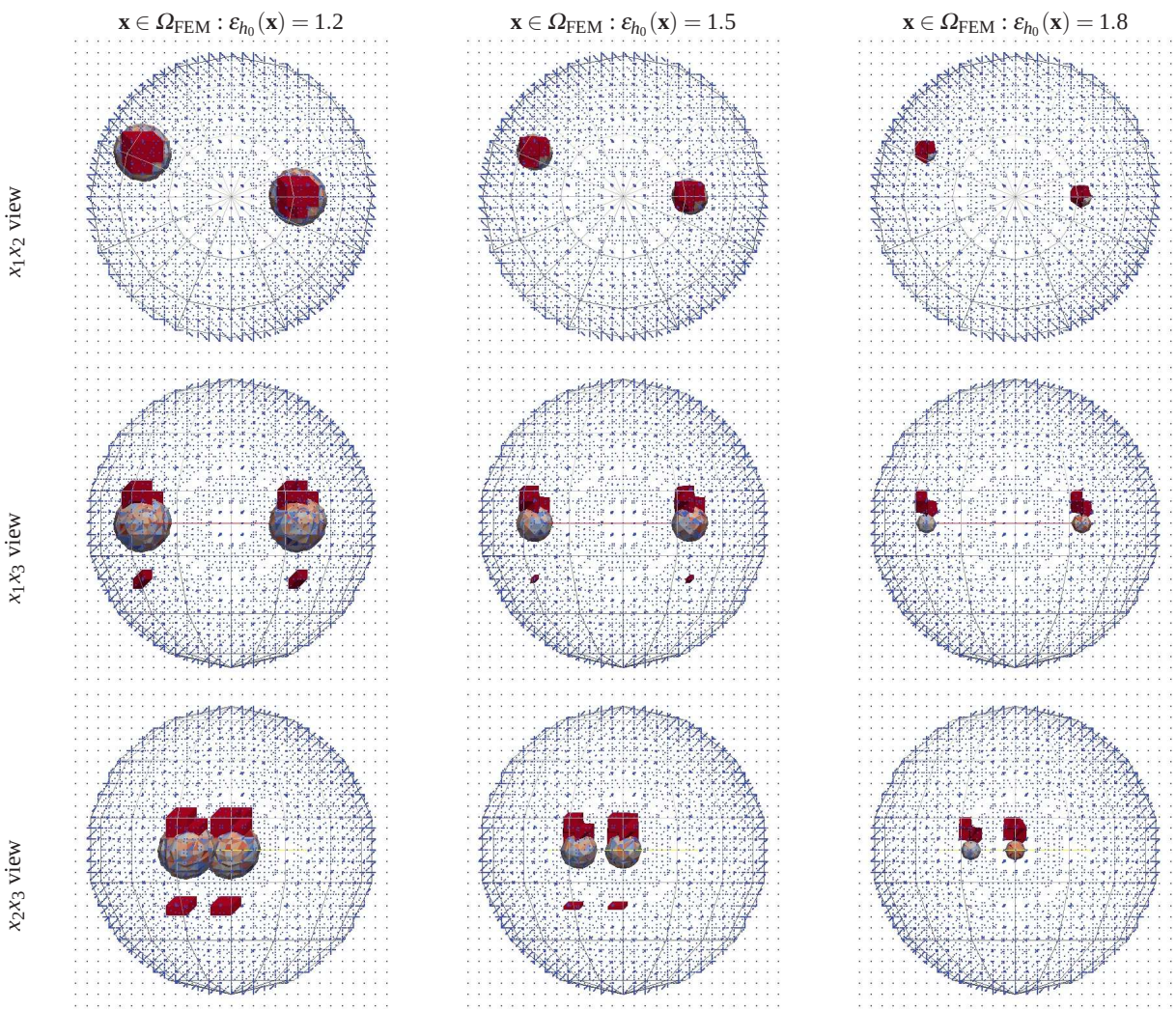


Fig. 5: *Test 1. In red: isosurfaces on a coarse mesh. Here,* $\max_{\Omega_{\mathrm{FEM}}} \varepsilon_{h_0} = 1.94$, *and the noise level in the data is* $\sigma = 10\%$. *For comparison we also present as wireframes the corresponding isosurface of the function* (30) *in every figure.*

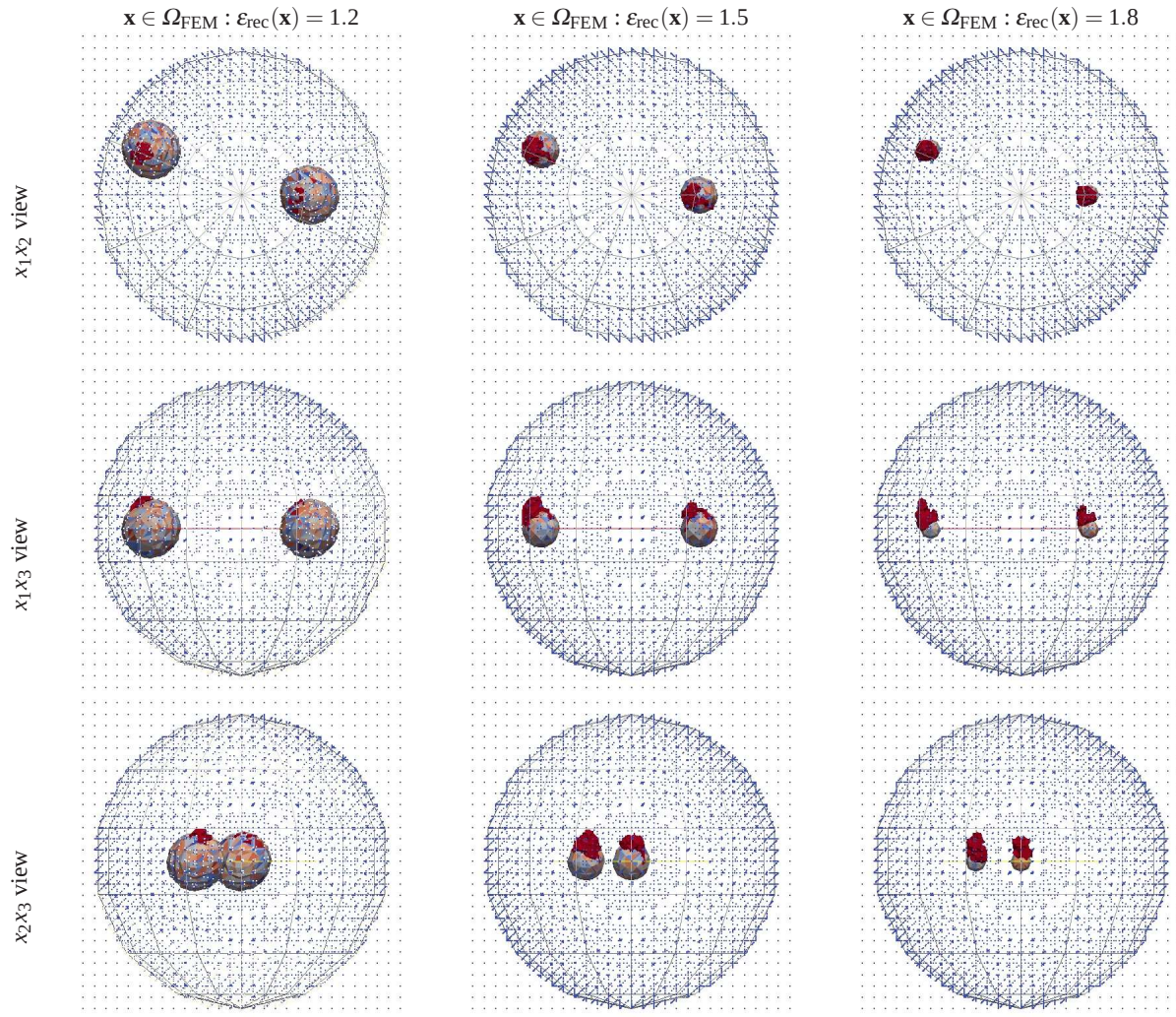


Fig. 6: *Test 1. In red: isosurfaces on the final mesh. Here,* ε_{rec} *was obtained on a five times refined mesh,* $\max_{\Omega_{\text{FEM}}} \varepsilon_{\text{rec}} = 1.97$*, and the noise level in the data is* $\sigma = 10\%$*. For comparison we also present as wireframes the corresponding isosurface of the function* (30) *in every figure.*

that, on the coarse mesh, we obtain quite correct locations of all inclusions and achieve maximal contrast of $\max_{\Omega_{\text{FEM}}} \varepsilon_{h_0} = 2.8$ in the inclusions. However, Figures 9-e), and g) show us that the locations of all inclusions in x_3-direction can still be improved. The figures on the right of Figure 9 present the reconstruction ε_{rec} of ε on the five times locally adaptively refined mesh. In Figure 9-f), and h) we observe an improvement of the reconstructions of the three inclusions in the x_3-direction on the final adaptively refined mesh, compared with reconstructions obtained on the coarse mesh. Comparing Tables 1 and 2 we also can conclude that adaptive mesh refinement allowed us to obtain more correct contrast for all three inclusions.

5.1.3 Test 3

This test is similar to Test 2, only here the goal was to reconstruct four small inclusions located at different parts of Ω_{FEM}: two inclusions were located closer to the backscattering boundary, and two other inclusions were placed closer to the transmission boundary of Ω_{FEM}, see Figure 10 where they are presented. These inclusions model four small malign tumors of a size 2 mm. We performed simulations with two additive noise levels in the data: $\sigma = 3\%$ and $\sigma = 10\%$, see Tables 1–2 for the results.

The reconstruction on the initial coarse mesh with a noise level $\sigma = 10\%$ in the data is presented on the left figures of Figure 10. From this figure and Table 1 we observe that we get quite correct locations of all

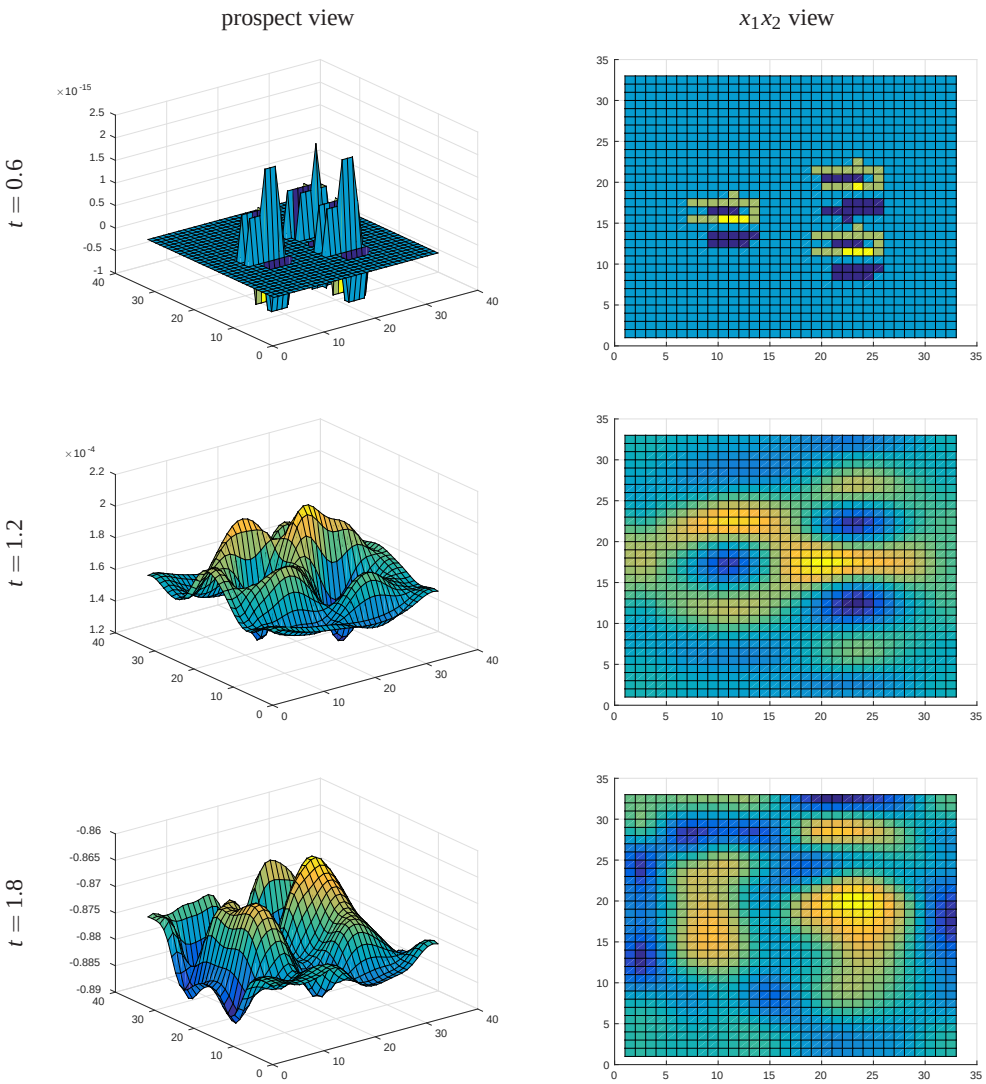


Fig. 7: *Test 2. Transmitted data of the component E_2 at different times. The noise level in the data is $\sigma = 10\%$.*

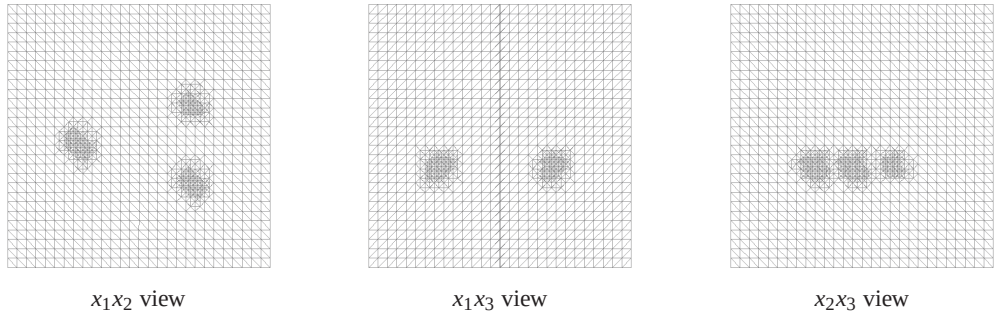


Fig. 8: *Test 2. Five times adaptively refined mesh when the noise level in the data was $\sigma = 10\%$.*

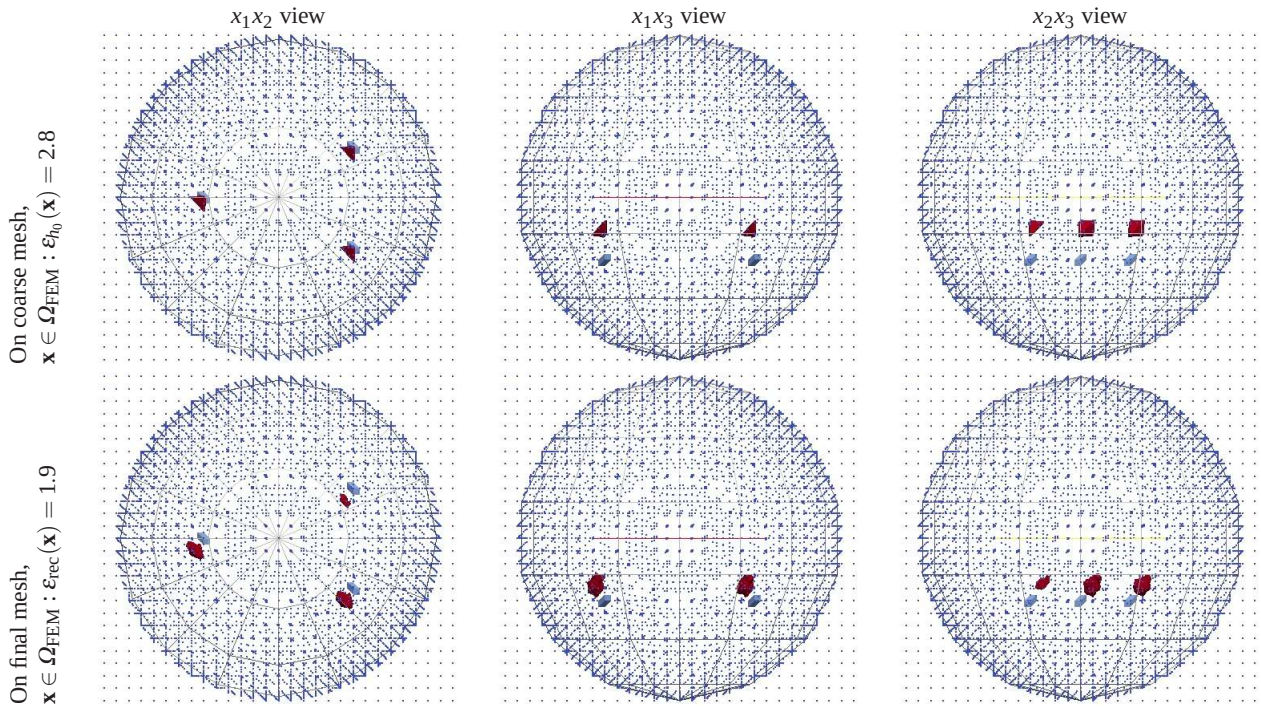


Fig. 9: *Test 2. Reconstructions (in red color) ε_{h_0} of ε obtained on the coarse mesh (upper figures), and $\varepsilon_{\mathrm{rec}}$ on the five times adaptively refined mesh (lower figures). The noise level in the data is $\sigma = 10\%$. For comparison we also present exact isosurfaces of the three small inclusions to be reconstructed (in light blue color).*

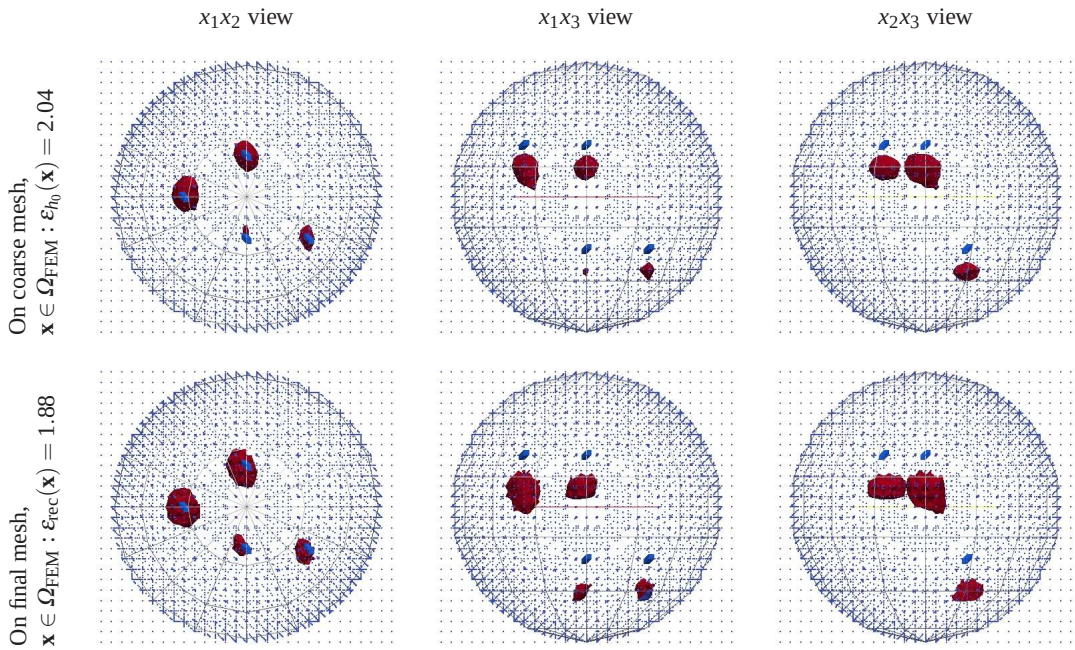


Fig. 10: *Test 3. Reconstruction of four inclusions (in red color) obtained on the coarse mesh (upper figures) and on the two times adaptively refined mesh (lower figures). The noise level in the data is $\sigma = 10\%$. For comparison we also present exact isosurfaces of the four small inclusions to be reconstructed (in light blue color).*

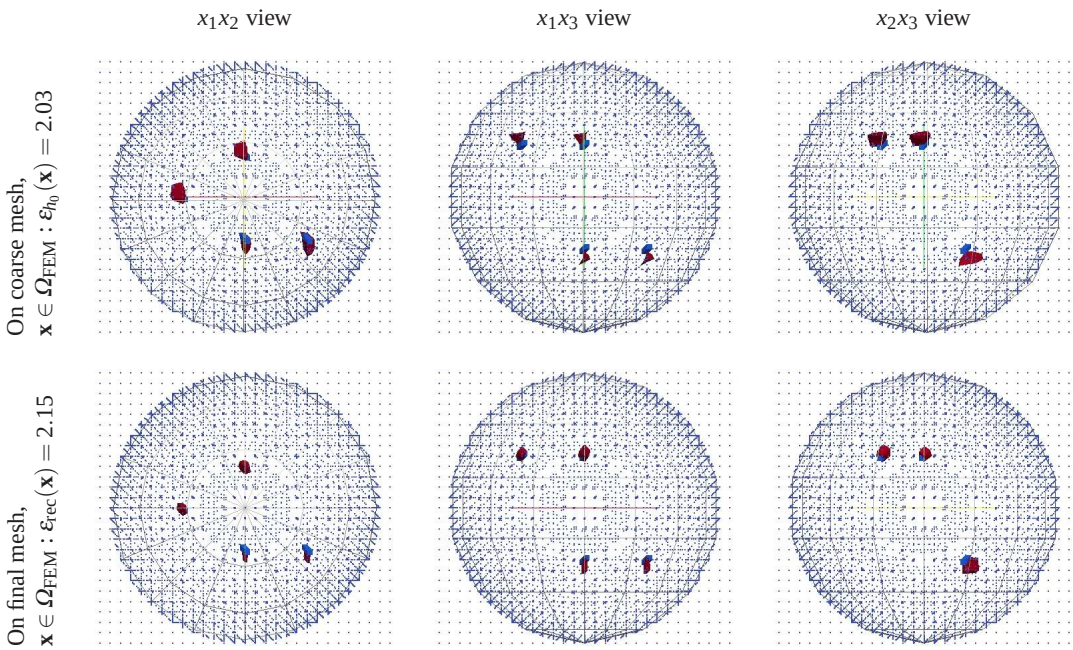


Fig. 11: *Test 4. Reconstruction obtained with two plane waves. We present reconstruction of the four inclusions (in red color) obtained on the coarse mesh (upper figures) and on the two times adaptively refined mesh (lower figures). The level of noise in the data is* $\sigma = 10\%$. *For comparison we also present exact isosurfaces of the four small inclusions to be reconstructed (in light blue color).*

inclusions and achieve a maximal contrast of $\max_{\Omega_{\mathrm{FEM}}} \varepsilon_{h_0} = 2.04$ on the coarse mesh. However, Figure 9-e), and g) show that the locations of all inclusions in the x_3-direction can still be improved. These figures also show that one of the four small inclusions is almost not present in the initial reconstruction.

The figures on the right of Figure 10 present the reconstruction $\varepsilon_{\mathrm{rec}}$ on the two times locally adaptively refined mesh. In Figure 10-f), and h) we observe that the two lower inclusions are well reconstructed. However, since in this quite challenging test we have used only transmitted data resulted from a single measurement of a plane wave, we do not observe significant improvement of the reconstruction of the four inclusions in the x_3-direction.

5.1.4 Test 4

In this test we decided improve the results of Test 3 and we tried to reconstruct the four inclusions using measurements of two backscattered wave fields. First, we initialized a plane wave at the front boundary $\partial_1\widetilde{\Omega}$ in time $[0, T]$ and collected backscattered data here. Next, we initialized a plane wave at the back boundary $\partial_2\widetilde{\Omega}$ in time $[T, 2T]$ and collected backscattered data for this wave field. We choose $T = 3$ and $\tau = 0.006$ as in all previous tests. Figure 11 and Tables 1–2 show the results of the reconstruction. Now we see that all inclusions are of the same size and they are reconstructed in correct positions and with correct contrasts on the coarse mesh as well as on the refined mesh. From Tables 1–2 we see that adaptive algorithm converged already after first mesh refinement. The drawback of this test is that computations of one optimization iteration took twice as much time as the corresponding iterations of the previous tests, because of two measurements of the backscattered wave fields.

6 Conclusion

In this work we have derived two a posteriori error estimates, for the direct error in the approximated permittivity as well as in the Tikhonov functional, in the finite element approximation of the Lagrangian approach to our coefficient inverse problem. Both estimates consist of two parts which can be interpreted as representing the error incurred by the approximation of the solution to the direct and adjoint problems, and the error incurred by the approximation of the coefficient itself, respectively. These estimates are important in the adaptive algorithms we have studied. Moreover, they further justify the use of similar error indicators in our previous works [7,8].

Numerically we have tested the adaptive algorithms with two different additive noise levels, $\sigma = 3\%$, and 10%, in the data. Our numerical tests show that with mesh refinements, as was expected, the quality of the reconstruction is improved a lot. Compare, for example,

the results of Figure 5 with those of Figure 6. Using these figures and tables 1 and 2 we observe that, with mesh refinements, the artifacts obtained on a coarse mesh are removed and the reconstructed function ε_{h_k} has more correct location in x_3 direction.

We can conclude that we have supported the tests of our previous works [6,7,8,30,31,39,40,41], and have shown that the adaptive finite element method is a powerful tool for the reconstruction of coefficients in Maxwell's equations from limited observations.

Our adaptive algorithms can also be applied for the case when edge elements are used for the numerical simulation of the solutions of forward and adjoint problems, see [42,43] for finite element analysis in this case. This, as well as development of different techniques for iterative choice of regularization the parameter in the Tikhonov functional, can be considered as a challenge for future research.

For the applications to medical imaging, an additional important challenge, both from a computational and theoretical point of view, is the extension of the adaptive Lagrangian methods to the case of a complex coefficient. This corresponds to considering a conductive medium, which is more realistic for organic tissues, and is a topic of our current research.

Acknowledgment

This research is supported by the Swedish Research Council (VR). The computations were performed on resources at Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

References

[1] A. Fhager, P. Hashemzadeh, and M. Persson, IEEE Trans. Biomed. Eng. **53**, 1594–1604 (2006).
[2] P.M. Meaney, Q. Fang, T. Rubaek, E. Demidenko, and K.D. Paulsen, Med. Phys. **34** 2014–2023 (2007).
[3] P.M. Meaney, K.D. Paulsen, and T.P. Ryan, IEEE Trans. Antennas Propag. **43**, 239–247 (1995).
[4] T.M. Grzegorczyk, P.M. Meaney, P.A. Kaufman, R.M diFlorio Alexander, and K.D. Paulsen, IEEE Trans Med Imaging., **31** 1584–1592, (2012).
[5] L. Beilina, Appl. Anal. **90**, 1461–1479 (2011).
[6] L. Beilina, M.V. Klibanov, and M.Yu. Kokurin, J. Math. Sci. **167**, 279–325 (2010).
[7] L. Beilina, Nguyen T.T., M.V. Klibanov, and J.B. Malmberg, Inverse Problems, **30**, 105007 (2014).
[8] L. Beilina, Nguyen T.T., M.V. Klibanov, and J.B. Malmberg, J. Comput. Appl. Math. **289**, 371–391 (2015).
[9] J.B. Malmberg, Springer Proceedings in Mathematics and Statistics **120**, 42–53 (2015).
[10] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics, Volume 15, Springer, New York, 2008.
[11] P. Monk, Finite Element Methods for Maxwell's Equations, Clarendon, Oxford, 2003.
[12] K.D. Poulsen and D.R. Lynch, IEEE Trans. Microwave Theory Techniques **39**, 395–404 (1991).
[13] O.A. Ladyzhenskaya, The boundary value problems of mathematical physics, Applied Mathematical Sciences, Volume 49, Springer-Verlag, New York, 1985.
[14] L.C. Evans, Partial differential equations, Graduate Studies in Mathematics, Volume 19, American Mathematical Society, Providence, RI, 2010.
[15] H.W. Engl, M. Hanke, and A. Neubauer, Regularization of Inverse Problems, Mathematics and its Applications, Volume 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
[16] A.N. Tikhonov, A.V. Goncharsky, V.V. Stepanov, and A.G. Yagola, Numerical Methods for the Solution of Ill-Posed Problems, Kluwer Academic Publishers, Dordrecht, 1995.
[17] A.L. Bukhgeĭm and M.V. Klibanov, Dokl. Akad. Nauk SSSR **260**, 269–272 (1981).
[18] M.V. Klibanov, Zh. Vychisl. Mat. i Mat. Fiz. **26**, 1063–1071 (1986).
[19] M. Bellassoued, M. Cristofol, and E. Soccorsi, Inverse Problems **28**, 095009 (2012).
[20] S. Li, SIAM J. Math. Anal. **37**, 1027–1043 (2005).
[21] S. Li and M. Yamamoto, Chin. Ann. Math. Ser. B **28**, 35–54 (2007).
[22] A.B. Bakushinsky, M.Yu. Kokurin, and A. Smirnova, Iterative Methods for Ill-Posed Problems: An Introduction, De Gruyter, Berlin, 2011.
[23] S. Hosseinzadegan, Iteratively regularized adaptive finite element method for reconstruction of coefficients in Maxwell's system, Master's thesis in applied mathematics, Department of Mathematical Sciences, Chalmers University of Technology and Gothenburg University, 2015.
[24] R. Becker, H. Kapp, and R. Rannacher, SIAM J. Control Optim. **39**, 113–132 (2000).
[25] K. Kraft and S. Larsson, BIT, **50**, 587–607 2010.
[26] C. Johnson and A. Szepessy, Comm. Pure Appl. Math. **48**, 199–234 (1995).
[27] F. Courant, K. Friedrichs, and H. Lewy, IBM J. Res. Develop. **11**, 215–234 (1967).
[28] O. Pironneau, Optimal shape design for elliptic systems, Springer-Verlag, Berlin, 1984.
[29] WavES: the software package. http://www.waves24.com.
[30] L. Beilina, Cent. Eur. J. Math. **11**, 702–733 (2013).
[31] L. Beilina and S. Hosseinzadegan, Applications of Mathematics, **61**, 253–286 (2016).
[32] C. Eyraud, A. Litman, A. Hérique, and W. Kofman, Inverse Problems **25**, 024005 (2009).
[33] W.T. Joines, Y. Zhang, C. Li, and R.L. Jirtle, Med. Phys. **21**, 547–550 (1994).
[34] A.E. Bulyshev, Souvorov A.E., S.Y. Semenov, V.G. Posukh, and Y.E. Sizov, Inverse Problems **20**, 1239–1259 (2004).
[35] W.C. Chew and Y.M. Wang, IEEE Trans. Med. Imaging **9**, 218–225 (1990).
[36] N. Joachimowicz, C. Pichot, and J.P. Hugonin, IEEE Trans. Antennas Propag. **39** 1742–1753, (1991).
[37] C. Eyraud, J.-M. Geffrin, and A. Litman, IEEE Antennas and Wireless Propagation Lett. **14**, 309–312 (2015).
[38] C. Eyraud, J.-M. Geffrin, and A. Litman, IEEE Microwave and Wireless Components Lett. **25**, 472–474 (2015).

[39] L. Beilina, Inverse problems and information technologies **1**, 73–116 (2002).
[40] L. Beilina and C. Johnson, Math. Models Methods Appl. Sci. **1**, 23–35 (2005).
[41] L. Beilina and C. Clason, SIAM J. Sci. Comp **28**, 382–402 (2006).
[42] P. Ciarlet Jr. and J. Zou, Numerische Mathematik **82**, 193–219 (1999).
[43] P. Ciarlet Jr., H. Wu, and J. Zou, SIAM J. Numer. Anal. **52**, 779–807 (2014).

John Bondestam Malmberg completed his PhD-degree in 2017, advised by Larisa Beilina, in the Department of Mathematical Sciences of Chalmers University of Technology Gothenburg University, Sweden. He graduated from Chalmers University of Technology, Sweden, in engineering mathematics in 2012. During his years as a doctoral student, J. B. Malmberg has authored or co-authored five scientific papers in peer-reviewed journals or conference proceedings, participated in five international scientific conferences or workshops, attended several national seminars or workshops as participant or invited speaker, and taught yearly courses in applied mathematics for undergraduate students. The topic of J. B. Malmberg's PhD thesis is numerical finite element and optimization methods for solving coefficient inverse problems for Maxwell's equations, with emphasis on a posteriori error analysis and simulation.

Larisa Beilina is working currently as an Associate Professor of Applied Mathematics at the Department of Mathematical Sciences at Chalmers University of Technology (CTH) and Gothenburg University (GU), Sweden. She obtained PhD degree at this university in 2003 and worked then in 2003-2005 as a PostDoc at University of Basel, Switzerland, and in 2007-2008 as a PostDoc at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway. She returned back in 2009 to the Department of Mathematical Sciences at CTH and GU and continued to work there as a researcher and teacher. Her main research achievements are in the solution of coefficient inverse problems (CIPs) using an adaptive finite element method, numerical methods for the wave propagation and electromagnetics, development of an approximatelly globally convergent method, numerical methods in optimization, domain decomposition methods, high-performing scientific computing using C++ and PETSc. L. Beilina has authored and co-authored four books published in Springer and more than 50 articles in peer-reviewed journals.