Thesis for the Degree of Doctor of Philosophy

Efficient Adaptive Algorithms for an Electromagnetic Coefficient Inverse Problem

John Bondestam Malmberg

Division of Mathematics
Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
Göteborg, Sweden, 2017

Efficient Adaptive Algorithms for an Electromagnetic Coefficient Inverse Problem

John Bondestam Malmberg

ISBN 978-91-629-0203-2 (electronic version)
ISBN 978-91-629-0204-9 (printed version)

Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
SE–412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 10 00

Typeset with LaTeX.
Printed in Gothenburg, Sweden, 2017 by Ineko AB.

Abstract

This thesis comprises five scientific papers, all of which are focusing on the inverse problem of reconstructing a dielectric permittivity which may vary in space inside a given domain. The data for the reconstruction consist of time-domain observations of the electric field, resulting from a single incident wave, on a part of the boundary of the domain under consideration. The medium is assumed to be isotropic, non-magnetic, and non-conductive. We model the permittivity as a continuous function, and identify distinct objects by means of isosurfaces at threshold values of the permittivity.

Our reconstruction method is centred around the minimization of a Tikhonov functional, well known from the theory of ill-posed problems, where the minimization is performed in a Lagrangian framework inspired by optimal control theory for partial differential equations. Initial approximations for the regularization and minimization are obtained either by a so-called approximately globally convergent method, or by a (simpler but less rigorous) homogeneous background guess.

The functions involved in the minimization are approximated with finite elements, or with a domain decomposition method with finite elements and finite differences. The computational meshes are refined adaptively with regard to the accuracy of the reconstructed permittivity, by means of an a posteriori error estimate derived in detail in the fourth paper.

The method is tested with success on simulated as well as laboratory measured data.

Keywords: coefficient inverse problem, inverse scattering, Maxwell's equations, approximate global convergence, finite element method, adaptivity, a posteriori error analysis

List of included papers

The following papers are included in this thesis:

- **Paper I.** Larisa Beilina, Nguyen Trung Thành, Michael V. Klibanov and **John Bondestam Malmberg**. Reconstruction of shapes and refractive indices from backscattering experimental data using the adaptivity. *Inverse problems* 30:105007, 2014.
- **Paper II.** Larisa Beilina, Nguyen Trung Thành, Michael V. Klibanov and **John Bondestam Malmberg**. Globally convergent and adaptive finite element methods in imaging of buried objects from experimental backscattering radar measurements. *Journal of Computational and Applied Mathematics* 289:371–391, 2015.
- **Paper III. John Bondestam Malmberg**. *A posteriori error estimate in the Lagrangian setting for an inverse problem based on a new formulation of Maxwell's system*, volume 120 of *Springer Proceedings in Mathematics and Statistics*, pages 42–53, Springer, 2015.
- **Paper IV. John Bondestam Malmberg**, and Larisa Beilina. An Adaptive Finite Element Method in Quantitative Reconstruction of Small Inclusions from Limited Observations. Manuscript submitted to *Applied Mathematics & Information Sciences*.
- **Paper V. John Bondestam Malmberg**, and Larisa Beilina. Iterative Regularization and Adaptivity for an Electromagnetic Coefficient Inverse Problem. Manuscript to appear in the *Proceedings of the 14th International Conference of Numerical Analysis and Applied Mathematics*.

The following papers are, for reasons of extent and consistency, not included in this thesis:

- John Bondestam Malmberg and Larisa Beilina. *Approximate globally convergent algorithm with applications in electrical prospecting*, volume 52 of *Springer Proceedings in Mathematics and Statistics*, pages 29–41, Springer, 2013.
- John Bondestam Malmberg. *A posteriori error estimation in a finite element method for reconstruction of dielectric permittivity.* Preprint.
- Larisa Beilina, Nguyen Trung Thành, Michael V. Klibanov and John Bondestam Malmberg. *Methods of Quantitative Reconstruction of Shapes and Refractive Indices from Experimental data*, volume 120 of *Springer Proceedings in Mathematics and Statistics*, pages 13–41, Springer, 2015.
- John Bondestam Malmberg, and Larisa Beilina. Reconstruction of a Dielectric Permittivity Function Using an Adaptive Finite Element Method. *Proceedings of ICEAA 2016.*

Contributions. The author of this thesis has contributed in the following manner to the appended papers:

In **Papers I** and **II** the author contributed to the development of the adaptive algorithm presented, performed and analyzed results of the numerical studies of this algorithm, and contributed significantly to the writing.

Paper III is entirely the work of the author, who developed the theory presented and wrote the paper.

In **Paper IV**, the author developed the theory presented, and did the major part of the writing.

In **Paper V**, the author designed and performed the numerical study presented, and did the major part of the writing.

Acknowledgments

A hearty thank you to Larisa Beilina, for excellent supervision – I could not have wished for better! – during my work on this thesis.

Another big thank you to Mohammad Asadzadeh, for many long and interesting discussions, and much good advice.

A third big thank you to friends and colleagues, for five very, very good years.

John Bondestam Malmberg
Gothenburg, 2017

Contents

Introduction

The thesis you have before you comprises five scientific papers, each of which deals in one way or another with computational aspects of inverse electromagnetic scattering. The purpose of this introduction is to present a background of this subject, to introduce the main mathematical tools which are used in the appended papers, and to highlight the main results of each of those papers. We will also briefly compare the methods presented in this work to other methods which are commonly used for the solution of similar problems.

Since inverse electromagnetic scattering, or electromagnetic imaging, constitutes the common underlying theme of each of the appended papers, this will be our natural starting point for this exposition.

1. Inverse electromagnetic scattering

The term *inverse electromagnetic scattering* refers to the process of inferring properties of a medium, or object, through observation of its interaction with electromagnetic radiation. Which properties one seeks to image, and what type of radiation used, depends on the application. In X-ray tomography, one seeks to image the interior of a patients body by observing its interaction with X-rays. With a radar, one detects the presence of objects, for example aircraft, in air with the aid of radio waves. All told, the applications of inverse electromagnetic scattering are very numerous in such diverse fields as geology, medicine, process industry, and security. The central idea of the process remains the same in most applications, and as such we will describe it below.

As the first step of the imaging process, electromagnetic signals of predetermined properties (frequency, duration, point of origin, and so on) are generated by a transmitter. These signals propagate into the medium to be imaged, where they are scattered – transmitted, reflected, and absorbed in various combinations determined by the properties of the medium. Some of these scattered signals are measured by one or several detectors, and thus we have a relation where the generated signals are transformed to the measured, or observed, scattered signals, in a manner which is determined by the medium. By combining these relations with a suitable mathematical model (typically Maxwell's equations, or simplifications thereof) for the scattering process in terms of the properties of the medium, we can attempt to solve the inverse scattering problem (which we return to in Section 2)

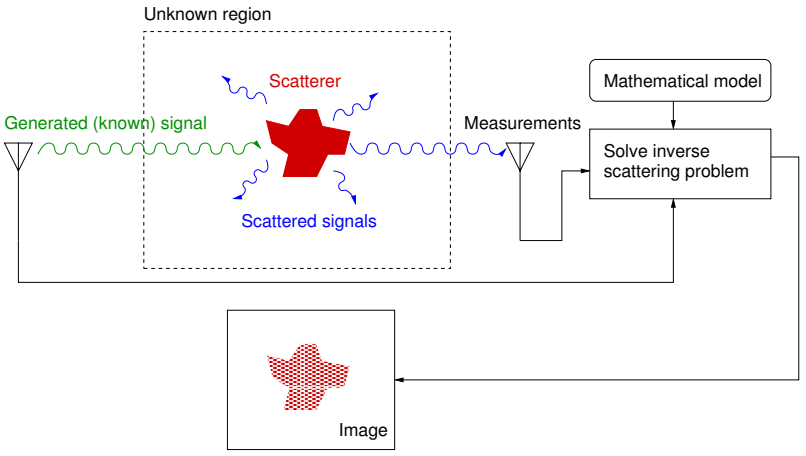

FIGURE 1. Schematic visualization of the inverse electromagnetic scattering process.

to get the image of the medium. The whole process is illustrated schematically in Figure 1.

In this thesis, our focus is the last step of the procedure outlined above, that of solving the inverse scattering problem to obtain the properties of the medium from the relation between incident signals and corresponding scattered signals. The process of generating and measuring signals is beyond the scope of this work, and will be considered only as a limiting factor for the type and amount of measured signals which we assume to be available. In the next section we will give some additional details on this.

1.1. Imaging of explosives and tumours. There are two main applications which we have in mind in the appended papers: detection of hidden explosives (in the air or buried in the ground) using radio waves, primarily considered in Papers I and II, and detection of cancer tumours with microwaves, primarily considered in Paper IV. These applications impose certain restrictions on the type of data one can expect to have access to. Here we present a short overview, for more details, we refer to [**? ? ?**].

In detection of explosives, in particular explosives which are hidden in the ground, a potentially hazardous target with an unknown location, it is important to obtain an image quickly, without access to

measurement around the target. Thus we only work with backscattered data measured on the same side of the target as the original signal was emitted, and data from only one incident signal. Since backscattered signals are usually weaker than transmitted data, this makes the problem of imaging more challenging, due to the very limited amount of data.

For microwave imaging of tumours, the challenges are slightly different. To keep radiation doses low, it is still desirable to work with as few signals as possible. On the other hand, the approximate location of the target, that is, the organ to be imaged, is known in advance, allowing for measurements also of transmitted, stronger, signals. However, the contrast between healthy and unhealthy tissues are significantly lower than in the case of detection of explosives. Moreover, due to the conductivity of the tissues of the human body, the signals are damped, and thus weaker. Of these two additional complications, the first one, that of low contrasts, are addressed in Paper IV, while that of losses induced by conductivity remains an ongoing work.

2. Inverse problems and regularization

As we remarked above, the main focus in this thesis is the solution of an inverse scattering problem. To be more precise, we reconstruct the dielectric permittivity of the imaged medium, which we assume to be non-magnetic, non-conductive, and isotropic, from data consisting of time-resolved observations of the electric field generated by a single pulse. This is an inverse problem in the sense that we, from given observations (the electric field), seek to determine the physical conditions (the permittivity) which caused the particular set of observations. Inverse problems are often ill-posed, which makes them more challenging to solve than their direct counterparts – here to compute the electric field from a given dielectric permittivity.

Ill-posed problems are defined as problems which are not well-posed, where a well-posed problem, in turn, is defined in the sense of Hadamard as a problem with the following three properties:

(1) A solution exists for any admissible data,
(2) there is only one, unique, solution for each given set of data, and
(3) the solution varies continuously with the data.

Given that the electric field is observed during a sufficiently long time interval, we assume that the last two properties of uniqueness and continuous data dependence are satisfied for our problem. Although we are not aware of a proof for the particular case we consider, this assumption appears reasonable in view of related uniqueness results such as [**?** **?** **?** **?** **?**]. We cannot make any similar assumptions regarding the first property, that of existence of a solution for any given, and in practice noisy, set of data, and thus our problem is ill-posed.

The rigorous theory of ill-posed problems was founded by Tikhonov around 1960 [**?**], and has since then been developed and applied by many others, see for example [**?** **?** **?**]. We shall now outline the general idea, using our inverse problem as an example.

Let us denote the observed electric field by $\mathbf{E}_{\text{obs}} = \mathbf{E}_{\text{obs}}(t, \mathbf{x})$, where t is the time variable and $\mathbf{x}$ is the spatial variable. Further, we denote the dielectric permittivity by $\varepsilon = \varepsilon(\mathbf{x})$, noting that it may vary in space, but not in time. For any given permittivity ε, we have a corresponding electric field which we denote by $\mathbf{E}_\varepsilon = \mathbf{E}_\varepsilon(t, \mathbf{x})$. Our inverse problem can now be stated as to determine ε such that $\mathbf{E}_\varepsilon = \mathbf{E}_{\text{obs}}$ for $0 \leq t \leq T$ and $\mathbf{x} \in \Gamma_{\text{obs}}$, where $T > 0$ is a final time for our measurements, and Γ_{obs} is the region in space (typically a part of a plane), where our observations are made.

By the ill-posed nature of our inverse problem, we cannot guarantee that such a permittivity exists for any given noisy observation $\mathbf{E}_{\text{obs}}$. The ideas of Tikhonov gives us a way to circumvent this problem, and obtain what is known as a quasi-solution, which may be thought of as the best approximation of a solution as can be expected, given the amount of noise in the data.

The argument is the following: Suppose that there is an ideal set of data $\mathbf{E}^*_{\text{obs}}$, to which corresponds a unique ideal permittivity ε^* with $\mathbf{E}_{\varepsilon^*} = \mathbf{E}^*_{\text{obs}}$ as desired. We regard the actual observations $\mathbf{E}_{\text{obs}}$ as a perturbation of $\mathbf{E}^*_{\text{obs}}$, with a noise level $\delta > 0$ such that $\|\mathbf{E}_{\text{obs}} - \mathbf{E}^*_{\text{obs}}\| \leq \delta$. (Here and below we will use the notation $\|\cdot\|$ for a generic norm, the nature of which may vary depending on the quantity inside. More precise descriptions can be found in the appended papers.) Assume now that we have a method of finding a unique ε_δ, for every $\delta > 0$ and every $\mathbf{E}_{\text{obs}}$ with $\|\mathbf{E}_{\text{obs}} - \mathbf{E}^*_{\text{obs}}\| \leq \delta$, such that $\varepsilon_\delta \to \varepsilon^*$ as $\delta \to 0$. We then say that the ill-posed problem is regularizable, and that ε_δ for our observation $\mathbf{E}_{\text{obs}}$ is the regularized solution, or quasi-solution,

to the problem. It should be noted that we do not require precise knowledge of ε^* and $\mathbf{E}^*_{\text{obs}}$. The existence of such a pair of ideal data and solution is an abstract assumption. Only the noise level δ and the actual observations $\mathbf{E}_{\text{obs}}$ are needed to obtain ε_δ.

In addition, Tikhonov proposed a method of constructing ε_δ for regularizable ill-posed problems, via the minimization of the Tikhonov functional, here

$$F_\alpha(\varepsilon) \coloneqq \frac{1}{2} \|\mathbf{E}_\varepsilon - \mathbf{E}_{\text{obs}}\|^2 + \frac{\alpha}{2} \|\varepsilon - \varepsilon_0\|^2 , \tag{2.1}$$

where $\alpha = \alpha(\delta) > 0$ is a regularization parameter which should be selected appropriately, in particular we must have $\alpha(\delta) \to 0$ as $\delta \to 0$, and ε_0 is an initial approximation of ε^*. We will discuss the choice of α and ε_0 further below. Given appropriate choices of α and ε_0, a unique minimizer ε_α of F_α exists, and we can define $\varepsilon_\delta \coloneqq \varepsilon_{\alpha(\delta)}$.

2.1. Choice of regularization parameter. For the theory of regularization by Tikhonov's method, it is important to define the correct behaviour of the regularization parameter $\alpha = \alpha(\delta)$ as the noise level δ tends to zero. In a practical setting however, we cannot control the noise level, but instead we are given fixed data, having some fixed noise level. Thus, how to choose a good value of the regularization parameter for a given observation becomes another important question.

Numerous methods for doing so exist, ranging from pure heuristics to rigorous algorithms. Some examples are the discrepancy principle [**?**], the L-curve method [**?**], iterative regularization by regarding the regularization parameter as an additional variable in the minimization [**?**], and manual choice. See for instance also [**?** **?**] for additional examples. Some of these methods require the Tikhonov functional to be minimized for several different values of the regularization parameter. In our case, minimizing the Tikhonov functional is very expensive in terms of computations, as we will see below. This makes methods which calls for repeated minimizations inappropriate for our purposes.

Of other methods, we have considered manual choice, and iterative regularization. The details are presented in Paper V, where we present a comparison, from which we conclude that the choice of regularization parameter does not appear to be the most critical factor in our approach to the inverse scattering problem under consideration.

2.2. Construction of initial approximation. Not only should we determine an appropriate value of the regularization parameter, we must also find a suitable initial approximation ε_0 for the dielectric permittivity. From the theoretical point of view, any ε_0 will do as we let the noise level drop to zero and choose regularization parameter accordingly. But, again, in practice, we cannot let the noise level drop to zero, and so the choice of ε_0 matters for the concrete reconstruction.

The simplest meaningful choice would be to assume that ε_0 is constant, with a value corresponding to that of the background medium. In absence of any other information, any other choice would require additional justification, and so the constant background initial approximation is commonly used. This is the case in Papers IV, and V. However, at the cost of performing additional computations, it is possible to use the observed data to construct a better initial approximation. In Papers I and II, this has been done using a so-called approximately globally convergent method. The full exposition of this method can be found in the book [**?**], and a brief overview is presented in Paper II of this thesis. For this introduction, we shall restrict ourselves to summarizing the key points of this method, in particular as applied to the construction of an initial approximation for the problem studied in this thesis.

The method is originally derived for reconstruction of the wave speed in the acoustic wave equation. It does not rely on any optimization algorithm, but uses the structure of the differential operator to obtain a sequence of approximations of the wave speed. The elements of this sequence are computed successively, with the first element computed from only the problem domain and the observed data.

In the framework of scalar waves, it can be shown that this sequence can be constructed in such a way that the elements converge to a wave speed which is close to, but not necessarily equal to, the ideal wave speed in Tikhonov's concept for the reconstruction problem [**?** **?**]. It is from this property the method derives its name. In addition, for correct choice of the regularization parameter, it can be shown that the corresponding Tikhonov functional is strongly convex in a neighbourhood of the obtained initial approximation, and that this neighbourhood contains the minimizer (Theorem 3.1 of [**?**]).

These important properties are theoretically and numerically established for reconstruction of speeds of scalar waves. For electromagnetic waves, where the dielectric permittivity ε corresponds to the wave

speed in the scalar case, the dependence of the electric field $\mathbf{E}_\varepsilon$ upon ε is more complicated than the dependence of a scalar wave on the wave speed. Hence, for electromagnetic waves, the approximately globally convergent method remains a heuristic, and the techniques used in the case of scalar waves cannot be used to derive a corresponding version of the approximately globally convergent method for electromagnetic waves. However, in view of numerical results in the appended papers, we will assume that the strong convexity property holds for F_α with the initial approximation ε_0 obtained as discussed in this section.

3. Modelling of the electric field

Although Tikhonov's method presented in Section 2, with the aid of the approximately globally convergent method presented in Section 2.2, shows us the way to proceed with the inverse scattering problem, we are by no means ready yet. In order to minimize the Tikhonov functional F_α of (2.1), we need at the very least to be able to evaluate it, and its gradient. This is where the proper mathematical model for the scattering comes into play. We recall that, in the approximately globally convergent method, the scalar wave equation was employed for this purpose. This equation is too unsophisticated to truly capture the nature of electromagnetic waves, though, and instead we should resort to Maxwell's equations

$$\begin{aligned} \mu\frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} &= \mathbf{0}, \\ \varepsilon\frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{H} &= -\sigma\mathbf{E}, \\ \nabla \cdot (\varepsilon\mathbf{E}) &= \rho, \\ \nabla \cdot (\mu\mathbf{H}) &= 0, \end{aligned} \tag{3.1}$$

where $\mathbf{E} = \mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H} = \mathbf{H}(\mathbf{x}, t)$ are electric and magnetizing fields, respectively, $\mu = \mu(\mathbf{x})$ is magnetic permeability, $\varepsilon = \varepsilon(\mathbf{x})$ is dielectric permittivity as before, $\sigma = \sigma(\mathbf{x})$ is conductivity, and $\rho = \rho(\mathbf{x}, t)$ is charge density. Here, and for the remaining part of this introduction, $\mathbf{x} \in \mathbb{R}^3$ denotes the spatial coordinates, and $t > 0$ denotes the time variable.

In contrast to the wave equation model used in the approximately globally convergent method, this full system of Maxwell's equations, involving six unknown functions (three components each of the two

unknown fields $\mathbf{E}$ and $\mathbf{H}$), may be too complicated and hence impractical for our computations. Therefore, we observe that we are, for the time being, concerned with non-magnetic and non-conductive media in absence of free charges, allowing us to put $\mu \equiv 1$, $\sigma \equiv \rho \equiv 0$, and at the cost of taking additional derivatives obtain the following system for $\mathbf{E}$,

$$
\begin{aligned}
&\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \nabla \times (\nabla \times \mathbf{E}) = \mathbf{0} && \text{in } \mathbb{R}^3 \times (0, T], \\
&\nabla \cdot (\varepsilon \mathbf{E}) = 0 && \text{in } \mathbb{R}^3 \times (0, T], \\
&\mathbf{E}(\cdot, 0) = \mathbf{0}, \quad \frac{\partial \mathbf{E}}{\partial t}(\cdot, 0) = \mathbf{I}\delta(\cdot - \mathbf{x}_0) && \text{in } \mathbb{R}^3,
\end{aligned}
\tag{3.2}
$$

which we here present as an "ideal" model, set in the whole space $\mathbb{R}^3$, and for the relevant time interval $[0, T]$, completed with initial conditions describing a pulse, parallel to $\mathbf{I}$, emanating at the point $\mathbf{x}_0 \in \mathbb{R}^3$ at time $t = 0$. Here, δ is the Dirac delta.

The system (3.2) is an ideal model in the sense that it captures more of the physical nature of electromagnetic waves, while it avoids too much complexity by not including any quantities which are not directly related to what we measure or seek to image. However, we cannot solve the system explicitly unless the permittivity ε is very simple, for instance constant. Moreover, being set in the whole space $\mathbb{R}^3$, the system does not easily lend itself well to numerical approximation by conventional methods. Thus, in order to proceed, we need to truncate the computational domain.

Suppose that the region of interest, that is, the region where we in the end will reconstruct the permittivity ε, is contained in a bounded – and for simplicity we may assume rectangular, hence convex – domain $\Omega \subset \mathbb{R}^3$, such that the permittivity has a constant background value close to the boundary as well as outside of Ω, and $\mathbf{x}_0 \notin \Omega$. The constant background value of the permittivity close to the boundary, together with the divergence condition, implies that the first equation of (3.2) decouples into three independent wave equations in that region. This fact helps us to make sense of the boundary conditions we apply, and we will also take advantage of it in our numerical scheme (see Subsection 4.2).

We describe the systems (3.2) in Ω by adding boundary conditions which emulate the free space around Ω, and the initial pulse, which by the time it reaches Ω is approximated by a plane wave. We end up

with the following system:

$$
\begin{aligned}
&\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \nabla \times (\nabla \times \mathbf{E}) = \mathbf{0} && \text{in } \Omega \times (0,\, T],\\
&\nabla \cdot (\varepsilon \mathbf{E}) = 0 && \text{in } \Omega \times (0,\, T],\\
&\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = \mathbf{0} && \text{on } \Gamma_{\mathrm{N}} \times (0,\, T],\\
&\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = -\frac{\partial \mathbf{E}}{\partial t} && \text{on } \Gamma_{\mathrm{A}} \times (0,\, T],\\
&\mathbf{E} = \mathbf{p} && \text{on } \Gamma_{\mathrm{I}} \times (0,\, t_1),\\
&\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = -\frac{\partial \mathbf{E}}{\partial t} && \text{on } \Gamma_{\mathrm{I}} \times [t_1,\, T],\\
&\mathbf{E}(\cdot,\, 0) = \frac{\partial \mathbf{E}}{\partial t}(\cdot,\, 0) = \mathbf{0} && \text{in } \Omega,
\end{aligned}
\tag{3.3}
$$

where $\mathbf{n}$ denotes the outward unit normal vector on the boundary $\Gamma = \Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{A}} \cup \Gamma_{\mathrm{I}}$ of Ω.

The boundary conditions, equations three through six in system (3.3) should be explained in some detail. See also Figure 2 for illustration. On Γ_{N}, which correspond to the faces of Ω which are not intersected by the line through the center of Ω and the point $\mathbf{x}_0$, we prescribe homogeneous Neumann boundary conditions to simulate free space stretching away in the directions perpendicular to the incident plane wave. On Γ_{A}, at the "back" of Ω from the point $\mathbf{x}_0$, we prescribe the first order absorbing boundary conditions of Engquist and Majda [**?**], to absorb transmitted signals. Finally, on Γ_{I} we prescribe an incident plane wave of profile $\mathbf{p} = \mathbf{p}(t)$ simulating the signal generated by the initial pulse in (3.2) far from the source at $\mathbf{x}_0$, until it has passed at time $t = t_1$ for some $t_1 \in (0,\, T)$, whence we switch to absorbing boundary conditions to absorb reflected signals.

The model (3.3) is the one that we actually use in computations, with the domain decomposition finite element-finite difference method of [**?**] for discretization. We shall return to this method below. Before we close this section, in order to simplify further presentation, we introduce a reduced version of (3.3), with simplified boundary conditions on a shrunken domain Ω. We also incorporate the divergence condition, the second equation of (3.3), as stabilization term of the form $-s\nabla(\nabla \cdot (\varepsilon \mathbf{E}))$, to the left hand side of the first equation, see [**?**

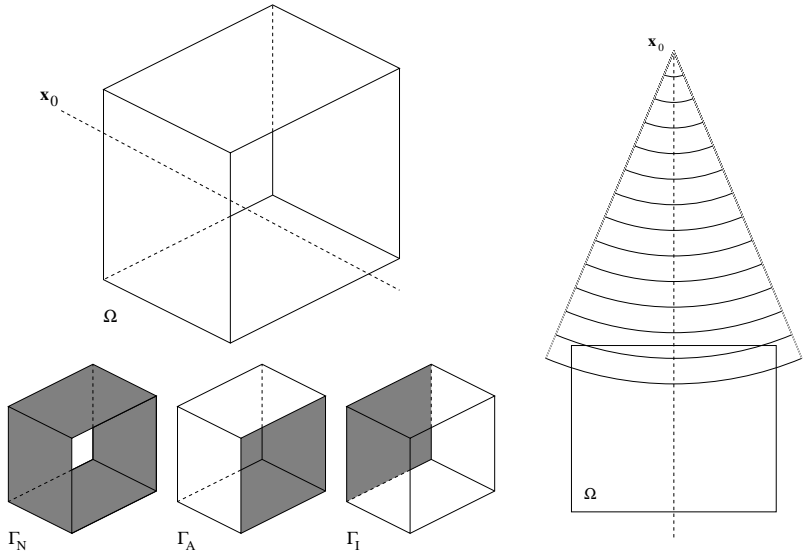

FIGURE 2. Schematic illustration of the truncated computational domain Ω (upper left), its boundary consisting of three parts $\Gamma = \Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{A}} \cup \Gamma_{\mathrm{I}}$ (lower left), and viewed from above with incident approximately plane wave (right).

?]. We choose $s = 1$ since this facilitates the derivation of the numerical scheme, and that is the value we typically use in computations. The stabilization term does not change the solution of the problem, but induces stability in the numerical approximation scheme which will be outlined in Section 4 (see [**?**]). It should also be mentioned that this numerical scheme may produce convergence to the wrong result in the case of non-convex domains, but this problem does not occur here, since we may always assume that Ω is convex in our problem framework.

$$
\begin{aligned}
&\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \nabla \times (\nabla \times \mathbf{E}) - \nabla(\nabla \cdot (\varepsilon \mathbf{E})) = \mathbf{0} && \text{in } \Omega \times (0,\, T], \\
&\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = \mathbf{P} && \text{on } \Gamma \times (0,\, T], \\
&\mathbf{E}(\cdot,\, 0) = \frac{\partial \mathbf{E}}{\partial t}(\cdot,\, 0) = \mathbf{0} && \text{in } \Omega.
\end{aligned}
\tag{3.4}
$$

Here $\mathbf{P} = \mathbf{P}(\mathbf{x}, t)$ is a given function defined on the entire $\Gamma \times (0, T]$.

4. Finite element approximation

With the models (3.3) as well as (3.4), we no longer face the problem of dealing with an infinite spatial domain which was the case in (3.2). But, there are still no explicit formulas available for the solution $\mathbf{E}_\varepsilon$ for a given ε, except in some trivial cases. Hence, we must consider numerical approximations. Our main approximation scheme is the finite element method (in fact hybridized with a finite difference scheme, as we will return to at the end of this section) which we outline below. In this section, we will illustrate this method as applied to the system (3.4), for good overviews, we recommend [**? ? ? ? ?**].

The idea of this method is the following: First, we should reformulate the original problem in variational form. This form makes the problem easier to analyze than it was in the original form, and one can show that the solutions to the two formulations coincide, under suitable assumptions on the problem domain and the data. The variational form is, in principle, no easier to actually solve than the original form, but it shows a way to proceed systematically with constructing approximations to the solution. The way to do it is to restrict the original function space over which the variational formulation was stated to a finite-dimensional subspace consisting of simple functions. Thus, we obtain a finite approximate problem, the finite element formulation. This problem can be solved systematically (with a computer), but the solution is no longer the same as to the original or variational problem, except in trivial cases. However, by selecting the finite dimensional subspace appropriately, we can assure that the solutions are close to each other, and that they become even closer if we increase the dimension of the finite-dimensional subspace.

For an example, let us consider the problem (3.4). To elaborate, we should really state it in the following way: *Given a permittivity ε in a set U^ε of admissible permittivities, and a function $\mathbf{P} \in C(\Gamma \times [0, T])$, determine a function $\mathbf{E} \in C^2(\Omega \times [0, T]; \mathbb{R}^3)$, such that these functions satisfy* (3.4).

The set U^ε is described in detail in for instance Paper IV. Here we merely note that we should have ε continuous, positive, bounded from above and below, and constantly equal to 1 on Γ. We have used the notation $C^k(X; Y)$ for k times continuously differentiable functions

over the domain X taking values in the space Y, leaving out k if it is zero and Y if it is $\mathbb{R}$.

Whether or not the function spaces stated in this problem are the least restrictive ones, and whether the problem is well posed in this formulation can be analyzed, but such analysis is beyond the scope of this thesis. Instead we refer to, for example, [**?** **?**], and assume for the time being that a solution does exist, and for this solution, we multiply the first equation of (3.4) by a smooth $\mathbb{R}^3$-valued function $\boldsymbol{\phi}$ defined in $\Omega \times [0,\, T]$, and integrate. This gives

$$\begin{aligned}
&\int_0^T \int_\Omega \varepsilon(\mathbf{x}) \frac{\partial^2 \mathbf{E}}{\partial t^2}(\mathbf{x},\, t) \cdot \boldsymbol{\phi}(\mathbf{x},\, t)\, \mathrm{d}\mathbf{x}\, \mathrm{d}t \\
&+ \int_0^T \int_\Omega \nabla \times (\nabla \times \mathbf{E}(\mathbf{x},\, t)) \cdot \boldsymbol{\phi}(\mathbf{x},\, t)\, \mathrm{d}\mathbf{x}\, \mathrm{d}t \\
&- \int_0^T \int_\Omega \nabla(\nabla \cdot (\varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x},\, t))) \cdot \boldsymbol{\phi}(\mathbf{x},\, t)\, \mathrm{d}\mathbf{x}\, \mathrm{d}t = 0.
\end{aligned}$$

Seeing that these expressions tend to become long and bulky, it is appropriate to simplify the notation by introducing the L_2-inner product:

$$\langle u,\, v \rangle_X \coloneqq \textstyle\int_X (u,\, v),$$

where $(\cdot,\, \cdot)$ denotes the standard inner product on $\mathbb{R}^n$, for appropriate $n \in \mathbb{N}$, and the integral is with respect to Lebesgue measure on X. This allows us to rewrite the above equation as

$$\langle \varepsilon \tfrac{\partial^2 \mathbf{E}}{\partial t^2},\, \boldsymbol{\phi} \rangle_{\Omega_T} + \langle \nabla \times (\nabla \times \mathbf{E}),\, \boldsymbol{\phi} \rangle_{\Omega_T} - \langle \nabla(\nabla \cdot (\varepsilon \mathbf{E})),\, \boldsymbol{\phi} \rangle_{\Omega_T} = 0 \tag{4.1}$$

with $\Omega_T \coloneqq \Omega \times (0,\, T)$.

We proceed to write this identity in a more symmetric form, and lower the burden of derivatives on a single function, by applying the following integration by parts formulas

$$\begin{aligned}
\int_0^T u'(t) v(t)\, \mathrm{d}t + \int_0^T u(t) v'(t)\, \mathrm{d}t &= u(T)v(T) - u(0)v(0), \\
\int_\Omega \nabla u(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_\Omega u(\mathbf{x}) \nabla \cdot \mathbf{v}(\mathbf{x})\, \mathrm{d}\mathbf{x} &= \int_\Gamma u(\mathbf{x}) \mathbf{n} \cdot \mathbf{v}(\mathbf{x})\, \mathrm{d}S
\end{aligned}$$

and the identity $\nabla \times (\nabla \times \mathbf{u}) = -\Delta \mathbf{u} + \nabla(\nabla \cdot \mathbf{u})$. Using these formulas in (4.1) we get

$$\begin{aligned} &- \langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}, \tfrac{\partial \boldsymbol{\phi}}{\partial t} \rangle_{\Omega_T} + \langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}(\cdot, T), \boldsymbol{\phi}(\cdot, T) \rangle_{\Omega} - \langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}(\cdot, 0), \boldsymbol{\phi}(\cdot, 0) \rangle_{\Omega} \\ &\quad + \langle \nabla \mathbf{E}, \nabla \boldsymbol{\phi} \rangle_{\Omega_T} - \langle \tfrac{\partial \mathbf{E}}{\partial \mathbf{n}}, \boldsymbol{\phi} \rangle_{\Gamma_T} - \langle \nabla \cdot \mathbf{E}, \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} + \langle \nabla \cdot \mathbf{E}, \mathbf{n} \cdot \boldsymbol{\phi} \rangle_{\Gamma_T} \\ &\quad + \langle \nabla \cdot (\varepsilon \mathbf{E}), \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} - \langle \nabla \cdot (\varepsilon \mathbf{E}), \mathbf{n} \cdot \boldsymbol{\phi} \rangle_{\Gamma_T} = 0, \end{aligned}$$

with $\Gamma_T \coloneqq \Gamma \times (0, T)$.

Here we observe that $\frac{\partial \mathbf{E}}{\partial \mathbf{n}} = \mathbf{P}$ and $\varepsilon \equiv 1$ on Γ, that $\frac{\partial \mathbf{E}}{\partial t}(\cdot, 0) \equiv \mathbf{0}$, and prescribe $\boldsymbol{\phi}(\cdot, T) \equiv \mathbf{0}$ to arrive at

$$\begin{aligned} &- \langle \varepsilon \tfrac{\partial \mathbf{E}}{\partial t}, \tfrac{\partial \boldsymbol{\phi}}{\partial t} \rangle_{\Omega_T} + \langle \nabla \mathbf{E}, \nabla \boldsymbol{\phi} \rangle_{\Omega_T} \\ &\quad - \langle \nabla \cdot \mathbf{E}, \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} + \langle \nabla \cdot (\varepsilon \mathbf{E}), \nabla \cdot \boldsymbol{\phi} \rangle_{\Omega_T} - \langle \mathbf{P}, \boldsymbol{\phi} \rangle_{\Gamma_T} = 0. \end{aligned} \tag{4.2}$$

We now pause to make some observations. For equation (4.2) to make sense, we do not really have to impose such strong restrictions on $\mathbf{E}$, $\mathbf{P}$, and $\boldsymbol{\phi}$ as we originally did. It is enough to have at most one derivative in either space or time on either of $\mathbf{E}$ and $\boldsymbol{\phi}$, and none of the functions need to be pointwise defined, as long as the integrals are well defined. Strictly speaking, this means that they need not be functions at all, but for convenience, we will continue to refer to them as such. In fact, it is enough to have $\mathbf{E}$, $\boldsymbol{\phi} \in H^1(\Omega_T; \mathbb{R}^3)$ and $\mathbf{P} \in L_2(\Gamma_T; \mathbb{R}^3)$, where

$$\begin{aligned} L_2(\Gamma_T; \mathbb{R}^3) &\coloneqq \{\mathbf{u} : \Gamma_T \to \mathbb{R}^3 : \|\mathbf{u}\|_{\Gamma_T} < \infty\}, \\ H^1(\Omega_T; \mathbb{R}^3) &\coloneqq \{\mathbf{u} : \Omega_T \to \mathbb{R}^3 : \|\mathbf{u}\|_{\Omega_T}, \|\nabla \mathbf{u}\|_{\Omega_T}, \left\|\tfrac{\partial \mathbf{u}}{\partial t}\right\|_{\Omega_T} < \infty\} \end{aligned}$$

with $\|\cdot\|_X \coloneqq \sqrt{\langle \cdot, \cdot \rangle_X}$, and the derivatives are understood in the weak sense. That is, $\frac{\partial \mathbf{u}}{\partial t}$ is defined by the relation

$$\langle \tfrac{\partial \mathbf{u}}{\partial t}, \boldsymbol{\phi} \rangle_{\Omega_T} = -\langle \mathbf{u}, \tfrac{\partial \boldsymbol{\phi}}{\partial t} \rangle_{\Omega_T}$$

for all $\boldsymbol{\phi} \in C^\infty(\Omega_T; \mathbb{R}^3)$ with compact support in Ω_T, that is, vanishing for $\mathbf{x}$ close to Γ and t close to 0 and T. Similar relations define $\nabla \mathbf{u}$, $\nabla \times \mathbf{u}$, and so on.

Putting what we have done so far together leads us to the following variational formulation of the problem 3.4: *Given a permittivity $\varepsilon \in U^\varepsilon$, and a function $\mathbf{P} \in L_2(\Gamma_T; \mathbb{R}^3)$, determine a function $\mathbf{E} \in H^1(\Omega_T; \mathbb{R}^3)$ such that $\mathbf{E}(\cdot, 0) \equiv \mathbf{0}$ and (4.2) is satisfied for these functions and all $\boldsymbol{\phi} \in H^1(\Omega_T; \mathbb{R}^3)$ with $\boldsymbol{\phi}(\cdot, T) \equiv \mathbf{0}$.*

Again, careful analysis may show that, for instance, the assumption $\mathbf{P} \in L_2(\Gamma_T; \mathbb{R}^3)$ is stronger than strictly necessary, but this assumption is sufficient for our purposes here.

The variational formulation enables us to analyze the problem easier in the framework of functional analysis, but it does not in itself tell us how to actually compute a solution. By making a finite dimensional approximation, we obtain a problem for which is relatively straight-forward to compute the solution, but this solution is then only an approximation to solution of the variational problem.

We will concentrate on approximation in the space $H^1(\Omega_T; \mathbb{R}^3)$ since we consider $\mathbf{P} \in L_2(\Gamma_T; \mathbb{R}^3)$ to be known exactly. Our aim is to construct a subspace of simple functions, which can approximate any function in $H^1(\Omega_T; \mathbb{R}^3)$, and for which we can easily find a basis which will allow us to reduce the variational problem, restricted to the finite dimensional subspace, to a problem of linear algebra.

Most methods for constructing such a subspace are based on dividing the underlying space, in our case Ω_T into small subregions – cells – and prescribe that the functions of the subspace should belong to some simple class on each cell. In the computations reported in the appended papers of this thesis, we have used polynomials of degree one on each cell, as we will shortly detail. The main reason for this is that the obtained space will allow domain decomposition and hybridization with the simpler but less flexible finite difference method in the manner of [**?**]. Polynomials of higher or lower degrees are also commonly used, as are special classes of polynomials or functions, such as in the edge element method of Nédélec [**?** **?**], which is widely used for Maxwell's equations and related systems. Those polynomials remove the need of a stabilization term, but is not as clear how to construct a domain decomposition method similar to [**?**] for the corresponding finite element method and a finite difference method. There are also certain methods which do not rely (directly) on a division of the problem domain into cells, such as radial basis function methods [**?**], or extended finite element methods [**?**], but these methods often produce less well-conditioned systems of equations than the cell-based methods, hence we do not consider such methods here.

To begin with the actual construction of finite dimensional subspaces, we let $\mathcal{T}_h = \{K\}$ be a division of Ω into tetrahedra K. This division is associated to a mesh function $h = h(\mathbf{x})$ with $h(\mathbf{x}) \coloneqq \operatorname{diam}(K)$, $\mathbf{x} \in K$, where $\operatorname{diam}(K)$ is the diameter, the largest distance between

two points, of K. We also partition $[0, T]$ uniformly into subintervals $\mathcal{I}_\tau \coloneqq \{I_k\}$ with $I_k = [t_{k+1}, t_k]$, $k = 0, \ldots, N_\tau$, for $0 = t_0 < \ldots < t_{N_\tau} = T$ where $t_{k+1} - t_k = \tau \coloneqq T/N_\tau$. We choose N_τ such that $\tau \propto \min_{\mathbf{x}\in\Omega} h(\mathbf{x})$ to fulfil a so-called Courant-Friedrichs-Lewy condition [**?**], which ensures stability of the numerical scheme we are constructing.

We now define the subspace S_h as the space of those functions in $H^1(\Omega_T; \mathbb{R}^3)$ which on each tetrahedron K and each time interval I_k are described by an $\mathbb{R}^3$-valued polynomial of degree one in the spatial variable, times a scalar polynomial of the time variable. In other words,

$$S_h \coloneqq \{\mathbf{u} \in H^1(\Omega_T; \mathbb{R}^3) : \mathbf{u}|_{K\times I_k} \in P^1(K; \mathbb{R}^3) \times P^1(I_k) \quad \forall K \in \mathcal{T}_h,\, I_k \in \mathcal{I}_\tau\},$$

where

$$P^1(K; \mathbb{R}^3) \coloneqq \{\mathbf{u} \in C(K; \mathbb{R}^3) : \mathbf{u}(\mathbf{x}) = A\mathbf{x} + \mathbf{b},\, A \in \mathbb{R}^{3\times 3},\, \mathbf{b} \in \mathbb{R}^3\},$$
$$P^1(I_k) \coloneqq \{u \in C(I_k) : u(t) = at + b,\, a,\, b \in \mathbb{R}\}.$$

The approximation properties of such a subspace is reflected in the typical interpolation error estimate

$$\|\mathbf{u} - i_h\mathbf{u}\|_{\Omega_T} \leq C \, \|h^m D^m \mathbf{u}\|_{\Omega_T}$$

where $i_h : H^1(\Omega_T; \mathbb{R}^3) \to S_h$ is an appropriate interpolation operator, such as the Scott-Zhang interpolant [**?**], and D^m denotes derivatives of order m. Thus, in the worst case, the interpolation error decreases linearly ($m = 1$) with the mesh size, for $\mathbf{u} \in H^1(\Omega_T; \mathbb{R}^3)$. Interpolation estimates like these play an important role in error analysis for finite element methods, which we will see examples of below.

We can now state the finite element formulation of the problem just as we stated the variational formulation, except that we replace the space $H^1(\Omega_T; \mathbb{R}^3)$ by its subspace S_h. That is: *Given $\varepsilon \in U^\varepsilon$, and a function $\mathbf{P} \in L_2(\Gamma_T; \mathbb{R}^3)$, determine $\mathbf{E}_h \in S_h$ with $\mathbf{E}_h(\cdot, 0) \equiv \mathbf{0}$, such that* (4.2) *holds for $\mathbf{E}_h$ and for all $\boldsymbol{\phi} \in S_h$ with $\boldsymbol{\phi}(\cdot, T) \equiv \mathbf{0}$.* (Note that this formulation does not rely on our explicit choice of S_h, and would be valid for other subspaces if required.)

By introducing a basis for S_h, we can reduce this finite element formulation to a system of linear algebraic equations. For a basis, there is a straight-forward choice, which yields good numerical properties. The basis is constructed as follows: Let $\{\mathbf{x}_i\}_{i=1}^{N_h}$ be the set of nodes on

$\mathcal{T}_h$ and for $i = 1, \ldots, N_h$ define $\varphi_i = \varphi_i(\mathbf{x})$ by

$$\varphi_i|_K \in P^1(K) \coloneqq \{\mathbf{a} \cdot \mathbf{x} + b,\, \mathbf{a} \in \mathbb{R}^3,\, b \in \mathbb{R}^3\}\ \forall K \in \mathcal{T}_h,$$

$$\varphi_i(\mathbf{x}_j) \coloneqq \delta_{i,j} \coloneqq \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad j = 1, \ldots, N_h.$$

Similarly, we define $\chi_k = \chi_k(t)$ for $k = 0, \ldots, N_\tau$ by

$$\chi_k|_{I_l} \in P^1(I_l)\ \forall I_l \in \mathcal{I}_\tau,$$
$$\chi_k(t_l) = \delta_{k,l},\ l = 0, \ldots, N_\tau.$$

It is now easy to verify that $\{\mathbf{e}_m \varphi_i \chi_k\}$, where $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is the standard basis for $\mathbb{R}^3$, is a basis for S_h. This means that we may expand

$$\mathbf{E}_h(t, \mathbf{x}) = \sum_{n=1}^{3} \sum_{j=1}^{N_h} \sum_{l=1}^{N_\tau} E_n^{j,l} \mathbf{e}_n \varphi_j(\mathbf{x}) \chi_l(t),$$

where $\{E_n^{j,l}\}$ are the weights that we seek. Moreover, it is sufficient to use $\boldsymbol{\phi} = \mathbf{e}_m \varphi_i \chi_k$ for $m = 1, 2, 3$, $i = 1, \ldots, N_h$, $k = 0, \ldots, N_\tau - 1$ in the finite element formulation.

By doing so, we reduce the finite element formulation for (4.2) to the following system of linear equations

$$A(\varepsilon)\frac{\mathbf{E}^{k+1} - 2\mathbf{E}^k + \mathbf{E}^{k-1}}{\tau^2} + B(\varepsilon)\frac{\mathbf{E}^{k+1} + 4\mathbf{E}^k + \mathbf{E}^{k-1}}{6} = \tau^2 \mathbf{P}^k, \tag{4.3}$$

for $k = 1, \ldots, N_\tau - 1$, with the slightly modified system

$$A(\varepsilon)\frac{\mathbf{E}^{k+1} - \mathbf{E}^k}{\tau^2} + B(\varepsilon)\frac{\mathbf{E}^{k+1} + 2\mathbf{E}^k}{6} = \tau^2 \mathbf{P}^k, \tag{4.4}$$

for $k = 0$. Here $A(\varepsilon)$ and $B(\varepsilon)$ are matrices defined by

$$A(\varepsilon) \coloneqq \begin{bmatrix} M(\varepsilon) & 0 & 0 \\ 0 & M(\varepsilon) & 0 \\ 0 & 0 & M(\varepsilon) \end{bmatrix},$$

$$B(\varepsilon) \coloneqq \begin{bmatrix} S + D_{1,1}(\varepsilon) & D_{1,2}(\varepsilon) & D_{1,3}(\varepsilon) \\ D_{2,1}(\varepsilon) & S + D_{2,2}(\varepsilon) & D_{2,3}(\varepsilon) \\ D_{3,1}(\varepsilon) & D_{3,2}(\varepsilon) & S + D_{3,3}(\varepsilon) \end{bmatrix},$$

$$M(\varepsilon) \coloneqq (m_{i,j}(\varepsilon)) \coloneqq (\textstyle\int_\Omega \varepsilon(\mathbf{x}) \varphi_j(\mathbf{x}) \varphi_i(\mathbf{x})\, \mathrm{d}\mathbf{x}),$$

$$S \coloneqq (\textstyle\int_\Omega \nabla \varphi_j(\mathbf{x}) \cdot \nabla \varphi_i(\mathbf{x})\, \mathrm{d}\mathbf{x})$$

$$D_{m,n}(\varepsilon) \coloneqq (\textstyle\int_\Omega \nabla \cdot ((\varepsilon(\mathbf{x}) - 1)\varphi_j(\mathbf{x})\mathbf{e}_n) \nabla \cdot (\varphi_i(\mathbf{x})\mathbf{e}_m)\, \mathrm{d}\mathbf{x}),$$

and

$$\mathbf{E}^k \coloneqq \begin{bmatrix} \mathbf{E}_1^k \\ \mathbf{E}_2^k \\ \mathbf{E}_3^k \end{bmatrix}, \quad \mathbf{E}_m^k \coloneqq \begin{bmatrix} E_m^{1,\,k} \\ \vdots \\ E_m^{N_h,\,k} \end{bmatrix}, \quad \mathbf{E}^0 = \mathbf{0},$$

$$\mathbf{P}^k \coloneqq \begin{bmatrix} \mathbf{P}_1^k \\ \mathbf{P}_2^k \\ \mathbf{P}_3^k \end{bmatrix}, \quad \mathbf{P}_m^k \coloneqq \begin{bmatrix} p_m^{1,\,k} \\ \vdots \\ p_m^{N_h,\,k} \end{bmatrix},$$

$$p_m^{i,\,k} \coloneqq \int_{\Gamma_T} P_m(t,\,\mathbf{x})\varphi_i(\mathbf{x})\chi_k(t)\,\mathrm{d}S\,\mathrm{d}t.$$

Thus, if we are given ε and $\mathbf{P}$, we can compute the matrices $A(\varepsilon)$ and $B(\varepsilon)$, as well as the vectors $\mathbf{P}^k$, and hence compute the vectors $\mathbf{E}^k$, $k = 1,\,\ldots,\,N_\tau$ by (4.3) and (4.4), starting from the known $\mathbf{E}^0 = \mathbf{0}$. We also note that one particular advantage of the basis we used is that the matrices $M(\varepsilon)$, S, and $D_{m,\,n}(\varepsilon)$ become sparse and structured, symmetric in the case of $M(\varepsilon)$ and S, since any two basis functions φ_i and φ_j overlap only if the nodes $\mathbf{x}_i$ and $\mathbf{x}_j$ are adjacent, otherwise the product of the basis functions, or their derivatives, are zero in all of Ω.

Although the system (4.3) is sparse, solving it repeatedly for each time step may be very time consuming, especially since we eventually compute the electric field not once but several times when solving the inverse problem. There are ways to circumvent this, and in our work we have considered the following two: making additional approximations to obtain an explicit formula for $\mathbf{E}^{k+1}$ in terms of $\mathbf{E}^k$ and $\mathbf{E}^{k-1}$, and using parallel computations. For the former, we use so-called mass lumping [**?**], approximating

$$A(\varepsilon) \approx A_{\mathrm{L}}(\varepsilon) \coloneqq \begin{bmatrix} M_{\mathrm{L}}(\varepsilon) & 0 & 0 \\ 0 & M_{\mathrm{L}}(\varepsilon) & 0 \\ 0 & 0 & M_{\mathrm{L}}(\varepsilon) \end{bmatrix}$$

where $M_{\mathrm{L}}(\varepsilon)$ is diagonal, with the row-sums of $M(\varepsilon)$ as diagonal entries, and

$$\frac{\mathbf{E}^{k+1} + 4\mathbf{E}^k + \mathbf{E}^{k-1}}{6} \approx \mathbf{E}^k.$$

With these approximations, we get

$$\mathbf{E}^{k+1} = 2\mathbf{E}^k - \mathbf{E}^{k-1} + \tau^2 A_{\mathrm{L}}(\varepsilon)^{-1}(\mathbf{P}^k - B(\varepsilon)\mathbf{E}^k),$$

which is explicit, since the inverse of the diagonal matrix $A_{\mathrm{L}}(\varepsilon)$ is just another diagonal matrix, with reciprocal diagonal entries.

For parallel computations, the idea of which is to divide the full problem into smaller sub-problems which are solved in parallel on several computer processors, we use the PETSc library [**?**] as implemented in the software package WavES [**?**].

4.1. Adaptivity. Regardless of the practicalities of the implementation, having stated a variational problem and its corresponding finite element formulation, where the solution to the latter approximates the solution to the former, leads to two natural questions:

(1) How large is the error, the difference between the finite element solution which we compute, and the solution to the variational problem, which we ideally would like to have? And,
(2) how do we efficiently make this error as small as we require?

The first of these two questions is answered through a priori error analysis. In such analysis, one relates the error to the mesh size h and the problem data (Ω, T, ε, and $\mathbf{P}$ for the problem above). Typical results show how the error decreases as some power of h, with a constant related to the data, as h tends to zero.

In answering the second question, that of efficiently reducing the error, another kind of analysis, a posteriori error analysis, plays an important role. Here, the idea is to estimate local (in space or time) contributions to the error, in terms of quantities which are actually computed, such as $\mathbf{E}_h$ in the problem described above. Since the approach we use to the inverse problem is based on improving the accuracy of an initial approximation, a posteriori error analysis is of primary interest. In the remainder of this section, we will describe the principle of how to use such analysis to efficiently reduce errors, with so-called adaptive error control (see for instance [**?** **?**]). For the precise estimates we use, we refer to the next section as well as the appended papers, in particular Paper IV.

To illustrate the principle, suppose that we have a mesh $\mathcal{T}_h$ as above, and that we on this mesh have computed a quantity q_h, approximating the true quantity q. Suppose also that through a posteriori error analysis, we have obtained an estimate on the form $\|q - q_h\| \leq C \|R_{h,\,q_h}\|$, where $C > 0$ is some constant, and $R_{h,\,q_h} = R_{h,\,q_h}(\mathbf{x})$ is some known non-negative function, depending only on the listed computed quantities, such that, in the worst case $|q(\mathbf{x}) - q_h(\mathbf{x})| \approx C R_{h,\,q_h}(\mathbf{x})$.

The idea is now to use a balancing principle, reducing large contributions to the error while not wasting additional computational efforts on reducing errors which already are small. Thus, as we have bounded the error in terms of $R_{h,\,q_h}$, and as we know that a finer mesh (with smaller mesh size h) leads to a smaller error, our strategy for reducing the error is to make the mesh finer where $R_{h,\,q_h}$ is close to attaining its maximum value. In other words, we have the following step by step procedure:

0: Make an initial coarse mesh $\mathcal{T}_h$.
1: Compute q_h on the current mesh $\mathcal{T}_h$.
2: Compute $R_{h,\,q_h}$ on the current mesh, using q_h from Step 1.
3: Check some stopping criterion, such as the size of $R_{h,\,q_h}$, if the criterion is satisfied, stop computing and accept the latest approximation q_h, otherwise proceed to Step 4.
4: Refine the current mesh by splitting cells in which $R_{h,\,q_h}(\mathbf{x}) \geq \theta \max_{\mathbf{x}} R_{h,\,q_h}(\mathbf{x})$, where $\theta \in (0,\,1)$ is a chosen tolerance. Then return to Step 1 with the refined mesh as the current mesh.

We illustrate one cycle of this procedure schematically for a mesh in one dimension in Figure 3.

4.2. Hybridization with finite differences. We conclude this section with some words about the hybridization of the finite element method with the finite difference method, which was mentioned previously. The details can be found in [**?**]. Here, we begin with a brief introduction to the finite difference method.

As with the finite element method, we start with a subdivision of the problem domain. Here, however, we are not concerned with what happens inside cells, but focus on the nodes from the beginning. As a consequence, we must be more restrictive with what types of subdivisions that we allow, namely, we require that the subdivision produces a rectangular grid of nodes (see Figure 4).

On this rectangular grid, we approximate the differential equation at each node using difference quotients for the derivatives. For example, if we consider the scalar wave equation

$$\frac{\partial^2 u}{\partial t^2}(\mathbf{x},\,t) - \Delta u(\mathbf{x},\,t) = 0,$$

we get the following equation at each node $\mathbf{x}_{i\,j\,k}$ (see Figure 5)

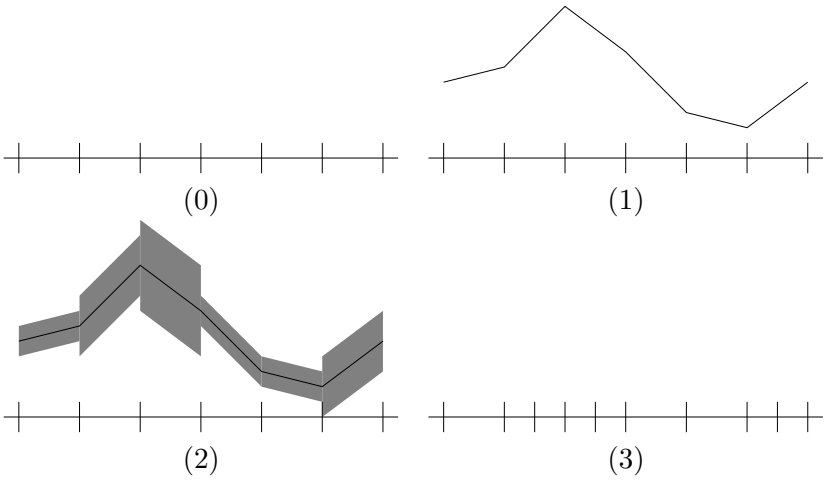

FIGURE 3. One dimensional schematic illustration of adaptivity: (0) The mesh at the start of the cycle. (1) The mesh with computed approximation. (2) Estimated errors, indicated by the width of the gray box around the approximation. (3) The refined mesh, for a certain tolerance.

$$\begin{aligned}\frac{\partial^2 U_{i,j,k}(t)}{\partial t^2} &+ \frac{U_{i-1,j,k}(t) - 2U_{i,j,k}(t) + U_{i+1,j,k}(t)}{h^2} \\ &+ \frac{U_{i,j-1,k}(t) - 2U_{i,j,k}(t) + U_{i,j+1,k}(t)}{h^2} \\ &+ \frac{U_{i,j,k-1}(t) - 2U_{i,j,k}(t) + U_{i,j,k+1}(t)}{h^2} = 0,\end{aligned}$$

where $U_{i,j,k}(t)$ denotes the nodal value of the approximation of the solution at node $\mathbf{x}_{i,j,k}$ time t, and h is the distance between two adjacent nodes, assumed for simplicity to be constant. It is convenient to denote the last three terms in the left hand side of the above equation by $-\Delta_h U_{i,j,k}(t)$.

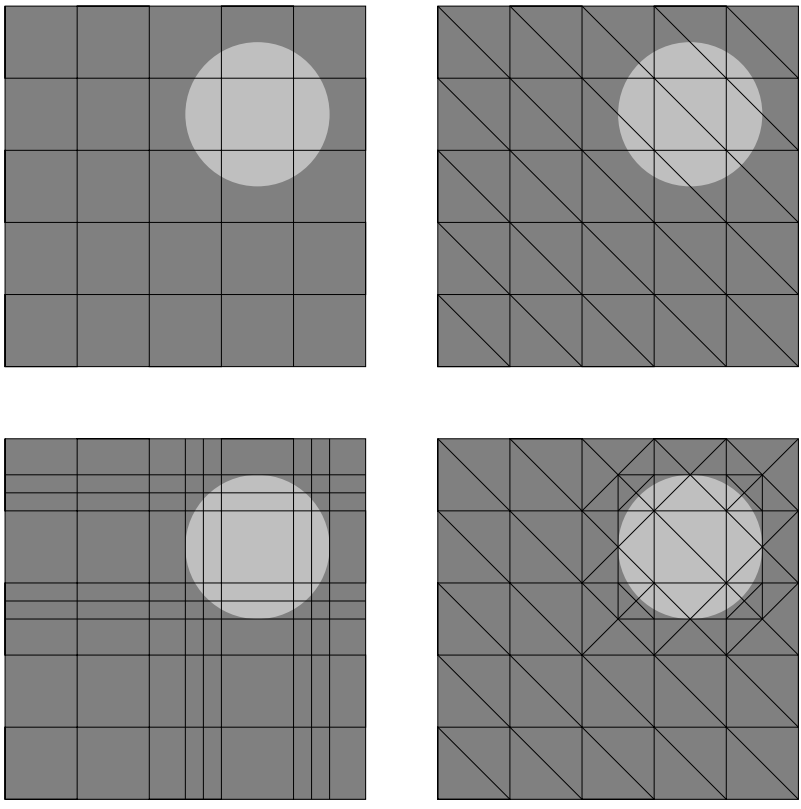

FIGURE 4. Finite difference meshes (left), and finite element meshes (right). The meshes in the upper part have the same 36 nodes. The lower figures show the meshes adaptively refined to fit the lighter circle to approximately the same accuracy. Then the finite difference mesh has 100 nodes, while the finite element mesh has 56 nodes.

Making a similar discretization in the time variable, we finally obtain

$$\frac{-U^{l+1}_{i,j,k} + 2U^{l}_{i,j,k} - U^{l-1}_{i,j,k}}{\tau^2} - \frac{\Delta_h U^{l+1}_{i,j,k} + 4\Delta_h U^{l}_{i,j,k} + \Delta_h U^{l-1}_{i,j,k}}{6} = 0,$$

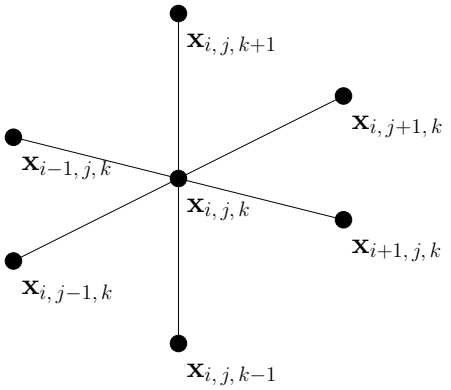

FIGURE 5. Local configuration of nodes in a finite difference mesh.

where $U^l_{i,\,j,\,k}$ denotes the approximate solution at node $\mathbf{x}_{i,\,j,\,k}$ and time $t = t_l$ for a (again for simplicity) uniform time partition. This system can be solved iteratively, just as (4.3) in the finite element method.

To connect this to our previous discussion on Maxwell's equations, and the finite element method, we make the following observations: Wherever we have a constant background value of $\varepsilon \equiv 1$, the expansion

$$\nabla \times (\nabla \times \mathbf{E}) = -\Delta \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}),$$

and the divergence condition $\nabla \cdot \mathbf{E} = 0$ implies that the differential equation of (3.2) (or (3.3), or (3.4)) decouples into three scalar wave equations, one for each component of the electric field $\mathbf{E}$. Moreover, if we consider the scalar wave equation with constant coefficients and source term, then the discrete system of equations produced by the finite difference method coincides with the one produced by the piecewise linear continuous finite element formulation, if we use a structured mesh, subdivided for the finite element method as in Figure 6. To show this is a straight forward generalization of the one dimensional case presented in Chapter 0 of [**?**].

In view of these observations, we can replace the more flexible finite element method with the more quick-assembled finite difference method in regions of our computational domain which are occupied only by the background medium, in other words close to the boundary of Ω. This allows us to propagate waves and implement boundary

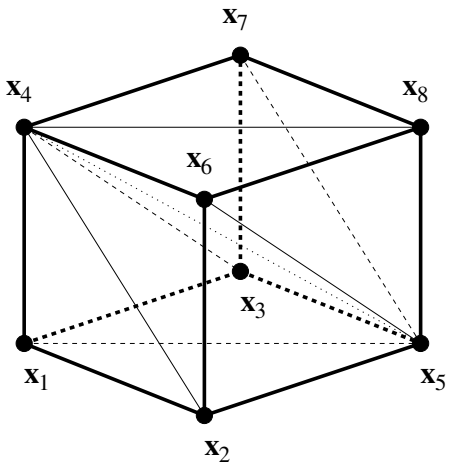

FIGURE 6. A subdivision of a cube with nodes $\mathbf{x}_i$, $i = 1, \ldots, 8$, into six tetrahedra $\triangle_{1254}$, $\triangle_{1345}$, $\triangle_{2465}$, $\triangle_{3457}$, $\triangle_{4586}$, $\triangle_{4578}$, where the tetrahedron $\triangle_{ijkl}$ has nodes $\mathbf{x}_i$, $\mathbf{x}_j$, $\mathbf{x}_k$, and $\mathbf{x}_l$.

conditions more efficiently, while maintaining flexibility to describe complicated variations in the dielectric permittivity.

5. Minimization, the Lagrangian approach

We have now presented methods for obtaining an initial approximation, and for approximately solving the direct problem of computing the electric field. Thus we can construct and evaluate the Tikhonov functional (2.1), and we may turn to the actual minimization.

As a continuous optimization problem, we expect this method to require the computation of gradients in order to achieve good rates of convergence in any iterative method used to solve it. Computing any kind of gradient of the Tikhonov functional directly is a very complicated task, however, since its dependence on the permittivity is implicit through the solution of a partial differential equation. Using techniques of optimal control theory presented in [**?**], we can circumvent this difficulty.

With this approach, we first reformulate the minimization problem as a problem of two formally independent variables, where a constraint ensures the correct dependence. That is, we consider the problem to

minimize

$$F_\alpha(\varepsilon,\, \mathbf{E}) \coloneqq \frac{1}{2}\,\|\mathbf{E} - \mathbf{E}_{\text{obs}}\|^2_{\Gamma_T} + \frac{\alpha}{2}\,\|\varepsilon - \varepsilon_0\|^2_{\Omega}$$

for $\varepsilon \in U^\varepsilon$, $\mathbf{E} \in H^1(\Omega_T;\, \mathbb{R}^3)$, $\mathbf{E}(\cdot,\, 0) \equiv \mathbf{0}$, under the constraint that

$$\mathcal{D}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\phi}) = 0 \;\forall \boldsymbol{\phi} \in H^1(\Omega_T;\, \mathbb{R}^3) : \boldsymbol{\phi}(\cdot,\, T) \equiv \mathbf{0},$$

where $\mathcal{D}$ is the functional of the variational formulation of the direct problem.

We may solve this problem by determining the stationary point to the corresponding Lagrangian

$$L(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda}) \coloneqq F_\alpha(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda}) + \mathcal{D}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda}).$$

Since the three variables of the Lagrangian may be varied independently, calculating derivatives of the Lagrangian is much more straight forward than doing so for the Tikhonov functional. Indeed, we find the partial Fréchet derivatives of L, acting on $\bar{\varepsilon} \in U^\varepsilon$, $\bar{\mathbf{E}} \in H^1(\Omega_T;\, \mathbb{R}^3) : \bar{\mathbf{E}}(\cdot,\, 0) \equiv \mathbf{0}$, and $\bar{\boldsymbol{\lambda}} \in H^1(\Omega_T;\, \mathbb{R}^3) : \bar{\boldsymbol{\lambda}}(\cdot,\, T) \equiv \mathbf{0}$, respectively, to be

$$\frac{\partial L}{\partial \varepsilon}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda};\, \bar{\varepsilon}) = \alpha\langle \varepsilon - \varepsilon_0,\, \bar{\varepsilon}\rangle_\Omega - \langle \tfrac{\partial \mathbf{E}}{\partial t} \cdot \tfrac{\partial \boldsymbol{\lambda}}{\partial t},\, \bar{\varepsilon}\rangle_{\Omega_T} + \langle \nabla \cdot \boldsymbol{\lambda},\, \nabla \cdot (\bar{\varepsilon}\mathbf{E})\rangle_{\Omega_T},$$

$$\frac{\partial L}{\partial \mathbf{E}}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda}; \bar{\mathbf{E}}) = \mathcal{A}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda},\, \bar{\mathbf{E}}),$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}}(\varepsilon,\, \mathbf{E},\, \boldsymbol{\lambda};\, \bar{\boldsymbol{\lambda}}) = \mathcal{D}(\varepsilon,\, \mathbf{E},\, \bar{\boldsymbol{\lambda}}),$$

where $\mathcal{A}$ is the functional of a variational formulation of an adjoint problem to the direct problem, where $\mathbf{E}$ acts as boundary data.

At the stationary point, all of these partial derivatives are zero, or in other words, their action on each admissible $\bar{\varepsilon}$, $\bar{\mathbf{E}}$, and $\bar{\boldsymbol{\lambda}}$ is zero. We observe that putting $\partial L/\partial \boldsymbol{\lambda}$ to zero in this manner implies that $\mathbf{E}$ solves the variational form of the direct problem, and putting $\partial L/\partial \mathbf{E}$ to zeros implies that $\boldsymbol{\lambda}$ solves the adjoint variational problem, determined by $\mathcal{A}$. Thus, we can ensure that these two partial derivatives remain zero by letting $\mathbf{E}$ and $\boldsymbol{\lambda}$ solve the correct variational problems, or in practice, the corresponding finite element problems.

Unlike the case of $\partial L/\partial \mathbf{E} = 0$ and $\partial L/\partial \boldsymbol{\lambda} = 0$, there is no obvious way to achieve $\partial L/\partial \varepsilon = 0$ by solving a system of partial differential equations, and so we instead apply the conjugate gradient method. We then arrive at the following algorithm:

0. Let ε_0 be the initial approximation for the permittivity, set $L_2(\Omega) \ni d_{-1} = 0$, and let $k = 0$.
1. Compute $\mathbf{E}_k$ by solving $\mathcal{D}(\varepsilon_k, \mathbf{E}_k, \cdot) = 0$.
2. Compute $\boldsymbol{\lambda}_k$ by solving $\mathcal{A}(\varepsilon_k, \mathbf{E}_k, \boldsymbol{\lambda}_k, \cdot) = 0$.
3. Compute the gradient $g_k = \frac{\partial L}{\partial \varepsilon}(\varepsilon_k, \mathbf{E}_k, \boldsymbol{\lambda}_k)$.
4. Update the permittivity by the conjugate gradient rule:

$$d_k = -g_k + \frac{\|g_k\|_\Omega^2}{\|g_{k-1}\|_\Omega^2} d_{k-1},$$
$$\varepsilon_{k+1} = \varepsilon_k + \beta_k d_k$$

   for step size $\beta_k > 0$ such that

$$\left(\frac{d}{d\beta} L(\varepsilon_k + \beta d_k, \mathbf{E}_k, \boldsymbol{\lambda}_k) \right) \Big|_{\beta=\beta_k} = 0.$$

5. Check stopping criterion. If satisfied, accept the current ε_k, otherwise increase k by 1 and return to step 1.

In practice, we replace the variational problems of steps 1 and 2 of this algorithm by the corresponding finite dimensional problems over some mesh. Consequently, we also obtain an approximate update of the permittivity in step 5, and an approximate permittivity ε_h as output.

At this point, we would like to apply the adaptive algorithm of Section 4.1, with $q = \varepsilon$ being the stationary point to the Lagrangian, and $q_h = \varepsilon_h$ being the output of the algorithm above. In order to do so, we require an a posteriori error estimate for $\varepsilon - \varepsilon_h$. Such an estimate is derived in Paper IV, and we here outline the main steps of the derivation.

As a first step, we use the strong convexity as discussed in Section 2.2. This implies that there is some constant $C > 0$ such that

$$\|\varepsilon - \varepsilon_h\|_\Omega^2 \leq C \left| F'_\alpha(\varepsilon; \varepsilon - \varepsilon_h) - F'_\alpha(\varepsilon_h; \varepsilon - \varepsilon_h) \right|,$$

where $F'_\alpha(\cdot, \varepsilon - \varepsilon_h)$ denotes the action of the Fréchet derivative of F_α on $\varepsilon - \varepsilon_h$. Since ε minimizes F_α, the first term in the right hand side vanishes, leaving us with

$$\|\varepsilon - \varepsilon_h\|_\Omega^2 \leq C \left| F'_\alpha(\varepsilon_h; \varepsilon - \varepsilon_h) \right|.$$

We can relate this derivative of F_α to $\partial L / \partial \varepsilon$ to get

$$\|\varepsilon - \varepsilon_h\|_\Omega^2 \leq C \left| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_{\varepsilon_h}, \boldsymbol{\lambda}_{\varepsilon_h}; \varepsilon - \varepsilon_h) \right|,$$

where $\mathbf{E}_{\varepsilon_h}$ and $\boldsymbol{\lambda}_{\varepsilon_h}$ are the exact solutions to the direct and adjoint variational problems, respectively, using the approximate permittivity ε_h.

A Cauchy-Schwartz inequality argument now gives, after cancellation of a common factor $\|\varepsilon - \varepsilon_h\|_\Omega$,

$$\|\varepsilon - \varepsilon_h\|_\Omega \leq C \left\| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_{\varepsilon_h}, \boldsymbol{\lambda}_{\varepsilon_h}) \right\|.$$

Since we cannot evaluate $\mathbf{E}_{\varepsilon_h}$ and $\boldsymbol{\lambda}_{\varepsilon_h}$, we need some further manipulations to obtain an a posteriori estimate. We add and subtract $\frac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h)$, where $\mathbf{E}_h$ and $\boldsymbol{\lambda}_h$ are finite element approximations of $\mathbf{E}_{\varepsilon_h}$ and $\boldsymbol{\lambda}_{\varepsilon_h}$, which we have computed on steps 1 and 2 of the minimization algorithm. With the triangle inequality, this gives

$$\begin{aligned} \|\varepsilon - \varepsilon_h\| \leq C \left\| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_{\varepsilon_h}, \boldsymbol{\lambda}_{\varepsilon_h}) - \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \right\| \\ + C \left\| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \right\|. \end{aligned} \tag{5.1}$$

The second term of the right hand side in (5.1) has already been calculated (up to the constant C) as part of the minimization algorithm, while the first term must be estimated further. To this end, we linearize the difference as

$$\begin{aligned} &\frac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_{\varepsilon_h}, \boldsymbol{\lambda}_{\varepsilon_h}) - \frac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \\ &\quad \approx \frac{\partial}{\partial \varepsilon} \left(\frac{\partial L}{\partial \mathbf{E}}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h; \mathbf{E}_{\varepsilon_h} - \mathbf{E}_h) + \frac{\partial L}{\partial \boldsymbol{\lambda}}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h; \boldsymbol{\lambda}_{\varepsilon_h} - \boldsymbol{\lambda}_h) \right). \end{aligned}$$

Here, we can use the fact that the Fréchet derivatives are linear in the last argument, and the fact that $\mathbf{E}_h$ and $\boldsymbol{\lambda}_h$, as solutions to the finite element problems, are also valid test functions, to replace the last $\mathbf{E}_h$ and $\boldsymbol{\lambda}_h$ of the two terms above, by the corresponding interpolants $i_h \mathbf{E}_{\varepsilon_h}$ and $i_h \boldsymbol{\lambda}_h$, respectively. Thus we get

$$\begin{aligned} &\frac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_{\varepsilon_h}, \boldsymbol{\lambda}_{\varepsilon_h}) - \frac{\partial L}{\partial \varepsilon}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h) \\ &\qquad \approx \frac{\partial}{\partial \varepsilon} \left(\frac{\partial L}{\partial \mathbf{E}}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h; \mathbf{E}_{\varepsilon_h} - i_h \mathbf{E}_{\varepsilon_h}) \right. \\ &\qquad\qquad \left. + \frac{\partial L}{\partial \boldsymbol{\lambda}}(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h; \boldsymbol{\lambda}_{\varepsilon_h} - i_h \boldsymbol{\lambda}_{\varepsilon_h}) \right). \end{aligned}$$

Now, we can use the Cauchy-Schwartz inequality and interpolation estimates to get

$$\begin{aligned}
&\left\| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h,\, \mathbf{E}_{\varepsilon_h},\, \boldsymbol{\lambda}_{\varepsilon_h}) - \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\| \\
&\quad \leq C_{\mathrm{i}} \left\| \tfrac{\partial^2 L}{\partial \varepsilon \partial \mathbf{E}}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\| \left\| h \left[\tfrac{\partial \mathbf{E}_h}{\partial \mathbf{n}} \right] + \tau \left[\tfrac{\partial \mathbf{E}_h}{\partial t} \right] \right\| \\
&\qquad + C_{\mathrm{i}} \left\| \tfrac{\partial^2 L}{\partial \varepsilon \partial \boldsymbol{\lambda}}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\| \left\| h \left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial \mathbf{n}} \right] + \tau \left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t} \right] \right\|,
\end{aligned}$$

where $C_{\mathrm{i}} > 0$ is an interpolation constant, and $[\cdot]$ denotes jumps of discontinuous functions over faces in the triangulation of the computational domain, and nodes in the time partition.

Putting this last estimate back into (5.1), we obtain the final estimate

$$\begin{aligned}
\|\varepsilon - \varepsilon_h\|_\Omega \leq\; & C C_{\mathrm{i}} \left\| \tfrac{\partial^2 L}{\partial \varepsilon \partial \mathbf{E}}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\| \left\| h \left[\tfrac{\partial \mathbf{E}_h}{\partial \mathbf{n}} \right] + \tau \left[\tfrac{\partial \mathbf{E}_h}{\partial t} \right] \right\| \\
& + C C_{\mathrm{i}} \left\| \tfrac{\partial^2 L}{\partial \varepsilon \partial \boldsymbol{\lambda}}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\| \left\| h \left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial \mathbf{n}} \right] + \tau \left[\tfrac{\partial \boldsymbol{\lambda}_h}{\partial t} \right] \right\| \\
& + C \left\| \tfrac{\partial L}{\partial \varepsilon}(\varepsilon_h,\, \mathbf{E}_h,\, \boldsymbol{\lambda}_h) \right\|,
\end{aligned}$$

where the first term corresponds to the error incurred by the approximation of the electric field, the second term corresponds to the error incurred by the approximation of the solution to the adjoint problem, and the last term corresponds to the error incurred by the approximations in the computation of the dielectric permittivity.

6. Other methods

We have now presented the principles and ideas of our approach to the electromagnetic inverse problem. To round off this introduction we will here perform a brief comparison between the method described in this thesis, and other methods which are commonly used for solving related inverse problems.

We begin by mentioning methods of boundary maps which are widely studied in the mathematical community of inverse problems (see for example[**?** **?**] and many other works by Salo, Somersalo, Uhlmann, and co-workers). These methods usually rely on observations of scattering of several waves, and thus with over-determined inverse problems which are quite different from the single-measurement context that we deal with in this thesis.

In engineering, minimization based approaches – again deterministic or stochastic – are more common, also for imaging problems similar

to the one studied here. See for example [**?** **?** **?** **?** **?** **?**]. One major difference between the typical approach in engineering, and the method described in this thesis, is that we are dealing with observations and computations in time-domain, while the usual engineering approach is set in frequency-domain, that is, in mathematical terminology, on the Fourier side in the time variable. This removes the need of time stepping, but introduces a frequency parametrization and complex arithmetic instead. As such, this may or may not make the method less computationally demanding, depending on the application. By imposing restrictions and parametrizations of the incident waves, and the geometry of the scatterers, one can lower the computational complexity significantly, possibly even allowing the use of analytical solution formulas, usually in terms of Hankel and Bessel functions, for the direct problem. Naturally, this restricts the type of problems which the method may be applied to. Thus, we can conclude that the method we study in this thesis may be more computationally demanding, but is comparatively flexible in terms of geometrical configuration of the scatterers and the nature of the incident waves.

7. Outline of the appended papers

Papers I, II, and III were previously included in the licentiate thesis [**?**]. The first two papers demonstrate the performance of our method on experimental data from a laboratory at the University of North Carolina at Charlotte, USA. In Paper I, the data represents scatterers placed in air, while the data in Paper II is for the more challenging case of scatterers placed inside a box filled with dry sand. Good reconstructions are obtained in both cases, but with restriction on the depth at which the scatterer was placed below the surface of the sand in the studies of Paper II. A case of super-resolution, where two scatterers at a separation of less than one half of a wavelength were resolved separately, is observed in Paper II.

Paper III can be regarded as a pre-work to Paper IV, it concerns a simpler a posteriori error estimate, for the error in the value of the Lagrangian, $L(\varepsilon, \mathbf{E}, \boldsymbol{\lambda}) - L(\varepsilon_h, \mathbf{E}_h, \boldsymbol{\lambda}_h)$, and introduces an alternative stabilization term. This stabilization fits into the same general framework as the one outlined above, namely, we consider a general stabilization term of the form $-\nabla(s\nabla \cdot (\varepsilon\mathbf{E}))$, where we require $s = s(\mathbf{x}) \geq 1/\varepsilon(\mathbf{x})$ by Theorem 4.1 of [**?**]. In this introduction as

well as the remaining appended papers we take a constant s, while in Paper III we take $s = 1/\varepsilon$, which allows us to expand $\nabla \times (\nabla \times \mathbf{E})$ and cancel certain terms (see Paper III for additional details).

In Paper IV, we derive in full detail the a posteriori error estimate which was sketched in Section 5, and present additional numerical results from simulated data for small scatterers with low contrast.

Paper V is devoted to effect of different techniques for choosing the regularization parameter on the quality of the reconstructed dielectric permittivity. We find that the regularization parameter has little impact on the quality of the reconstruction for our inverse problem, and hence other aspects of the reconstruction procedure are more important for future studies.

Papers

Paper I

Reprinted from Larisa Beilina, Nguyen Trung Thành, Michael V. Klibanov and John Bondestam Malmberg, Reconstruction of shapes and refractive indices from backscattering experimental data using the adaptivity, *Inverse problems* 30:105007, 2014, with permission from the publisher.

Paper II

Reprinted from Larisa Beilina, Nguyen Trung Thành, Michael V. Klibanov and John Bondestam Malmberg, Globally convergent and adaptive finite element methods in imaging of buried objects from experimental backscattering radar measurements, *Journal of Computational and Applied Mathematics*, 289:371–391, 2015 with permission from Elsevier.

Paper III

Reprinted from John Bondestam Malmberg, *A posteriori error estimate in the Lagrangian setting for an inverse problem based on a new formulation of Maxwell's system*, volume 120 of *Springer Proceedings in Mathematics and Statistics*, pages 42–53, Springer, 2015, with permission from the publisher.

Paper IV

John Bondestam Malmberg, and Larisa Beilina. An Adaptive Finite Element Method in Quantitative Reconstruction of Small Inclusions from Limited Observations. Manuscript submitted to *Applied Mathematics & Information Sciences*.

Paper V

John Bondestam Malmberg, and Larisa Beilina. Iterative Regularization and Adaptivity for an Electromagnetic Coefficient Inverse Problem. Manuscript to appear in the *Proceedings of the 14th International Conference of Numerical Analysis and Applied Mathematics*.