DOE/ID/13566

An Adaptive B-Spline Method for Low-order Image Reconstruction Problems

Final Report – 09/24/1997 – 09/24/2000

X. Li
E. L. Miller
C. Rappaport
M. Silevich

April 2001

Work Performed Under Contract No. DE-FG07-97ID13566

For
U.S. Department of Energy
Assistant Secretary for
Energy Efficiency and Renewable Energy
Washington, DC

By
Northeastern University
Boston, MA

DOE/ID/13566

AN ADAPTIVE B-SPLINE METHOD FOR LOW-ORDER IMAGE RECONSTRUCTION PROBLEMS

FINAL REPORT
09/24/1997 – 09/24/2000

X. Li
E. L. Miller
C. Rappaport
M. Silevich

April 2001

Work Performed Under Contract No. DE-FG07-97ID13566

Prepared for the
U.S. Department of Energy
Assistant Secretary for
Energy Efficiency and Renewable Energy
Washington, DC

Prepared by
Northeastern University
Boston, MA

Final Report for DOE Contract DE-FG07-97ID13566

Eric L. Miller, Carey Rappaport, and Michael Silevich
235 Forsyth Building, Northeastern University, Boston, MA 02115
Tel: (617) 373-8386 FAX: (617) 373-8627 email: elmiller@ece.neu.edu

April 11, 2001

1 Introduction

Extracting information about the internal structure of a medium from noisy observations of scattered electromagnetic energy obtained along the boundary is a problem arising in a wide range of environmental monitoring and remediation applications. In many cases the processing goal is the characterization of an anomalous area such as the presence of a pollution plume in the overall region of interest. Often, this is approached by discretizing the medium, specifically the space-varying complex electrical permittivity, into a large number of small pixels. An estimate is obtained via the solution of a high dimensional, often non-linear optimization problem [1].

Such an approach is difficult for a number of reasons. First, because there can be anywhere from thousands to hundreds of thousands of pixels or voxels in the region, the computational burden of solving the optimization problem can be prohibitive. Second, these problems are highly underdetermined. Generally the number of pixels or voxels in the region is far less than the number of measurements collected at the boundary. Thus, there are many permittivity profiles which provide an adequate fit to the data but whose basic structures are physically unrealistic. For example, there may be negative permittivity values or the structure may be characterized by high frequency, large amplitude oscillations. Stabilizing the solution to these inverse problems (also known as regularizing them) to obtain physically plausible reconstructions is a difficult and subtle task [2–4].

Under this contract, we have considered two methods for approaching this inverse medium problem which attempt to address these two fundamental issues. The first approach is based on the idea that if one is looking for an object against a fairly well-behaved background then it should be possible to avoid the high price of using a fine scale pixelation. Rather, a low-dimensional parameterization of the unknown should be possible with fine scale variation used sparingly in areas of marked interest such as near boundaries between the object and the background or between layers in the background. The technical issue here is in adaptively determining where such detail should be placed as the algorithm progresses. To address this problem we represent the unknown permittivity profile using a sparse collection of refinable B-spline functions. Beginning with a coarse scale reconstruction of the unknown we exploit the mathematical structure of the B-splines to alternatively add detail in regions of large variability and then remove unneeded degrees of freedom from areas where there is little "interesting" behavior. By controlling the degrees of freedom (DOF) like this, we obtain an optimization problem with hundreds rather than thousands of unknowns. The controlled use of the B-splines results in a high degree of flexibility allowing us to recover geometric and contrast structure of anomalous regions against unknown backgrounds with a rather compact representation of the unknown profile. We have examined the performance of this

approach using a 2D crosswell radar tomography problem in which a simple, linearized physical model was employed. We obtained comparable results relative to a single, fine scale reconstruction using only about 12% of the unknowns in about half of the time. A full discussion of this method is provided in the Master's Thesis attached to this document.

In addition to our B-spline method, we have developed a new imaging method for the problem where one knows both the contrast of the objects as well as that of the background and is only seeking to determine the number, shapes, and locations of these perturbations. Here we have drawn heavily on recent work in the image processing community to derive a level-set-based curve evolution method which essentially "shrink-wraps" a curve around the objects of interest. In addition a very efficient algorithm using a so-called "source-type" formulation of the relevant physics has been synthesized to determine an initial contour for the level-set technique. We have explored the performance of these methods on a two dimensional, cross-well radar tomography problem using a full physical model so that the unknown permittivity was related non-linearly to the scattered field data. Our results indicate that this method is highly robust both to the presence of noise in the data as well as modeling errors in the assumed known background medium electrical properties. A detailed discussion of the mathematics underlying our approach and its evaluation on simulated data is provided in reprints of two papers attached to the end of this report.

The results of these two efforts are encouraging. The B-spline work indicates the utility of adaptive methods for reducing the complexity and ill-posedness of these complex inverse problems. The highly robust performance of the level-set approach points to the utility of parameterizing these problems in terms of the geometry of the unknown. In the future, we intend on exploring the following issues related to these methods

1. On the algorithmic side, the performance of the curve evolution method can be traced in part to the assumptions underlying the algorithm regarding knowledge of the background structure and contrast of the objects. Developing level-set type techniques in which these assumptions are removed is a key area of future work.

2. The method by which we adaptively refine the B-spline representation of the permittivity requires specification of threshold parameter. Currently, that parameter is set by hand; however to make the approach both reliable and automatic, it is important develop a more rigorous approach. Similarly, the manner in which we prune degrees of freedom from the representation can be improved. Currently, a simple greedy-type approach is used to solve what is essentially a combinatorial optimization problem. It would be useful to consider alternate approximation methods for such problems as well as approaches from the subset selection literature.

3. In reality, these two methods complement each other and should ultimately be combined. The B-spline approach provides rough geometric information regarding areas where anomalies might be located as well as the contrast structure of the anomaly and background contrast. Such information can be of significant use to the level-set technique. Similarly, the boundary information from the level-set approach can be used to delineate disjoint regions over which low order B-spline models can be used to parameterize the contrast. Determining how to make optimal joint use of these two modeling methods is at the heart of this type of work.

4. Both techniques need to be extended to consider fully 3D problems and verified on real sensor data. Currently, we are collaborating with INEEL on just such a project. The initial work has been in extending the level-set approach to a three dimensional, cross-borehole electrical resistivity problem. The physics here are far more tractable than is the case for

higher frequency radar imaging so we can concentrate our effort more on the information extraction problem. Over the next coupe of years, we anticipate validation of the level-set type approach and ultimately its fusion with something like the adaptive B-spline method as a means of dealing with unknown contrast in the background medium and anomalies.

The remainder of this final report is composed of reprints of the three major technical papers arising from this contract:

1. *An Adaptive B-Spline Method for Low-order Image Reconstruction Problems* by Xin Li. A Master's Thesis completed under the supervision of Prof. Miller at Northeastern University.

2. *A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets* by Oliver Dorn, Eric L. Miller and Carey Rappaport which appeared in Inverse Problems special issue on Electromagnetic Imaging and Inversion of the Earth Subsurface, October 2000.

3. *Shape reconstruction in 2D from limited-view multifrequency electromagnetic data* by Oliver Dorn, Eric L. Miller, and Carey M. Rappaport which will be published in the AMS Contemporary Mathematics Series.

References

[1] Y. M. Wang and W. C. Chew. An iterative solution to the two-dimensional electromagnetic inverse scattering problem. *International Journal of Imaging Systems and Technology*, 1(1):100–108, 1989.

[2] M. Bertero, C. De Mol, and E. R. Pike. Linear inverse problems with discrete data, II, Stability and regularisation. *Inverse Probl.*, 4:573–594, 1988.

[3] Murat Belge, Eric Miller, and Misha Kilmer. Simultaneous multiple regularization parameter selection by means of the l-hypersurface with applications to linear inverse problems posed in the wavelet transform domain. In *SPIE International Symposium on Optical Science, Engineering, and Instrumentation: Bayesian Inference for Inverse Problems*. SPIE, July 1998.

[4] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4):561–580, December 1992.

An Adaptive B-Spline Method for Low-order Image Reconstruction Problems

Xin Li

April 11, 2001

Contents

Abstract **6**

1 Introduction **8**

1.1 Contribution 10

1.2 Organization 12

2 Background **14**

2.1 Cross-well Tomography Problem 14

2.1.1 Problem Setup 14

2.1.2 Electromagnetic Inverse Scattering 15

2.1.3 Born Approximation 18

2.2 Regularization of Linear Inverse Problems 19

2.2.1 Linear Inverse Problems 19

2.2.2 Nature of Ill-posed Problems 19

2.2.3 SVD Analysis 20

2.2.4 L-Curve 22

2.2.5 General Regularizations 23

2.3 Introduction to B-Spline Functions 24

2.3.1 B-Spline Basics 25

2.3.2 Knots Distribution and Smoothness 28

2.4 Geometries of Splines 32
2.4.1 Curvature of Space Curves 32
2.4.2 Curvature of Surface Curves 35
2.4.3 Gaussian and Absolute Curvature 36

3 Edge-preserving Regularizer 39
3.1 Deficiency of Tikhonov Regularization 39
3.2 Regularization Functions 41
3.3 Edge-preserving Regularizer 42
3.4 Example 43

4 Adaptive B-Spline Reconstruction 49
4.1 Motivation 49
4.2 Inverse Model using B-Splines Basis 50
4.3 Optimum Knots Distribution 51
4.4 Knots Refining 53
4.5 Knot Pruning 55
4.5.1 Weight of Knots 56
4.5.2 Knots Deletion 57
4.6 Overall Algorithm 60
4.7 Examples 62
4.8 Monte-Carlo Simulation 63

5 Conclusions and Future Work 71
5.1 Conclusions 71
5.2 Future Work 73

Bibliography 75

List of Figures

2.1 Cross-well tomography problem. 15

2.2 Cross-well tomography problem. 16

2.3 A simple 1-d inverse problem, with a Gaussian function as the kernel. 21

2.4 Illustration of L-curve 23

2.5 Tikhonov regularization effects for the 1-d problem. 24

2.6 B-Splines $N_{i,k+1}$ with different order $(k+1)$ and their first and second derivatives. ($i = 3$, knots sequence $\lambda = [0123456]$) 27

2.7 Examples of tensor product B-Splines and corresponding knots distribution. 29

2.8 Illustration of using cubic B-Splines to approximate Titanium Heat data. (a)5 equally-spaced knots; (b)10 equally- spaced knots; (c)20 equally-spaced knots. (d)12 selected knots. 30

2.9 Distant knots leads to B-Spline of large support (flat) and vice versa; Certain multiple knots leads to discontinuity. 31

2.10 Arc length reparameterize of curve $\mathbf{x}(t)$ 34

2.11 Franet Frame and Osculating Circle 34

2.12 Frame for surface 37

2.13 Example of only need to add knots in one direction 38

3.1 Illustrations of regularization functions (a)Quadratic (Tikhonov) (b)1-norm (c)Edge-preserving ($\varphi(t) = \frac{t^2}{1+t^2}$) (d)Edge-preserving ($\frac{\varphi'(t)}{2t} = \frac{1}{(1+t^2)^2}$) 40

3.2 Example: (a),(b) objects; (c),(d) Tikhonov reconstructions; (e),(f) Edge-preserving reconstructions. Transmitters and receivers are placed along y dimension ($x = 0$ and $x = 15$). 46

3.3 Example: (a),(b) L-surfaces; (c),(d) Curvature of the L-surfaces. . . . 47

3.4 Example: (a),(b) CRB of Tikhonov regularizer; (c),(d) CRB of edge-preserving regularizer. 48

4.1 Illustration of Tree Structure for Knots Searching (1-d case) 53

4.2 Illustration of weight of knot. (a)10 equally-spaced knots to approximate the data(dotted line); (b)Approximation of deleting the 7th knot; (c)Approximation of deleting the 3rd knot; (d)Cost(Weight) of each knot. 57

4.3 Illustration of vicinity effects of weight measurement. (a)Initial knots distribution and approximation curve; (b)Weight measurement; (c)Weight of each knot after deleting the 12th knot; (d)Weight of each knot after deleting the 12th and 13th knot. 58

4.4 Flow chart of knot optimization algorithm. 60

4.5 Using knots refining and pruning algorithm on the Titanium data fitting problem. 61

4.6 Example: Reconstruction obtained by Tikhonov regularization and edge-preserving regularization method. (a)'+' shaped object, Tikhonov regularization (MSE = 5.7393) (b)rectangular object, Tikhonov regularization (MSE = 5.2943) (c)'+' shaped object, edge-preserving regularization (MSE = 3.3930) (d)rectangular object, edge-preserving regularization (MSE = 2.9372) 65

4.7 Example: Reconstruction and knots distribution using B-Spline based edge-preserving regularizer. (a),(b) '+' shaped object; (MSE = 7.8461) (c),(d) rectangular object. (MSE = 5.7867) 66

4.8 Example: B-Spline reconstruction for 3-layer problem (a)Object; (b)Reconstruction; (c)Knots distribution; (d)Convergence of cost (Each iteration consists of knot inserting and knot pruning procedure.) 67

4.9 Monte-Carlo simulation for the 3-layer problem. (SNR = 20dB) (a)Average pixel-based reconstruction (MSE = 2.0874); (b)Average B-Spline reconstruction (MSE = 5.8922); (c)Total knots distribution; (d)Scaled knots distribution. 68

4.10 Knot selection by using separate 1D histograms. (a)Histogram of knot distribution along x direction. (b)Histogram of knot distribution along y direction. (c)Reconstruction based on histogram analysis. (d)Corresponding knot distribution. 69

4.11 Variance of the Monte-Carlo simulation (2 different viewpoints) . . . 70

Abstract

A common problem in signal processing is to estimate the structure of an object from noisy indirect measurements linearly related to the desired image. These problems are broadly known as *inverse problems*. A key feature which complicates the solution to such problems is their ill-posedness. That is, small perturbations in the data arising e.g. from noise can and do lead to severe, non-physical artifacts in the recovered image. The process of stabilizing these problems is known as regularization of which Tikhonov regularization is one of the most common. While this approach leads to a simple linear least squares problem to solve for generating the reconstruction, it has the unfortunate side effect of producing smooth images thereby obscuring important features such as edges. Therefore, over the past decade there has been much work in the development of edge-preserving regularizers. This technique leads to image estimates in which the important features are retained, but computationally they require the solution of a nonlinear least squares problem, a daunting task in many practical multi-dimensional applications.

In this thesis we explore low-order models for reducing the complexity of the reconstruction process. Specifically, B-Splines are used to approximate the object. If a "proper" collection B-Splines are chosen so that the object can be efficiently represented using a few basis functions, the dimensionality of the underlying problem will be significantly decreased. Consequently, an optimum distribution of splines needs to be determined. Here, an adaptive refining and pruning algorithm is developed to

solve the problem. The refining part is based on curvature information, in which the intuition is that a relatively dense set of fine scale basis elements should cluster near regions of high curvature while a sparse collection of basis vectors are required to adequately represent the object over spatially smooth areas. The pruning part is a greedy search algorithm to find and delete redundant knots based on the estimation of a weight associated with each basis vector. The overall algorithm iterates by inserting and deleting knots and end up with much fewer knots than pixels to represent the object, while the estimation error is within a certain tolerance. Thus, an efficient reconstruction can be obtained which significantly reduces the complexity of the problem.

In this thesis, the adaptive B-Spline method is applied to a cross-well tomography problem. The problem comes from the application of finding underground pollution plumes. Cross-well tomography method is applied by placing arrays of electromagnetic transmitters and receivers along the boundaries of the interested region. By utilizing inverse scattering method, a linear inverse model is set up and furthermore the adaptive B-Spline method described above is applied. The simulation results show that the B-Spline method reduces the dimensional complexity by 90%, compared with that of the pixel-based method, and decreases time complexity by 50%, without significantly degrading the estimation.

Chapter 1

Introduction

A common problem in signal processing is to estimate the structure of an object from noisy indirect measurements. These problems are broadly known as *inverse problems* with many applications such as radar/medical imaging, mine detecting and image reconstruction [1] [2] [3] [4]. Among these, a common application is the cross-well tomography problem, where we need to locate and estimate the buried objects by observations of scattered electromagnetics radiation taken along the boundaries [6] [7] .

In this thesis, we will first introduce the cross-well tomography problem, then set up the linear inverse model based on inverse scattering theory [8] [18]. The difficulty of solving the linear inverse problem lies in the ill-posed nature of the degradation kernel. The classical solutions tend to have large high-frequency artifacts which contaminate the solution. To make the solution more stable, i.e. smooth out those high-frequency components, a smoothing constraint is added to the underlying optimization problem, a procedure called *regularization*. Many regularization methods have been widely explored [19] [20], among which the Tikhonov regularization is the most common [21].

Tikhonov regularization actually adds a smoothing constraint on the solution so

that the high-frequency artifacts are filtered out. However, at the same time, it also blurs the edges of the object, which are frequently the most interesting areas to us. Therefore, more specific constraint should be added to preserve the edges. Edge-preserving regularization using non-quadratic regularizers is such an idea [24] [25] [26] [27] . The assumption is that, in the reconstruction, large changes in the parameters we are estimating are likely to be true edges which we want to preserve, while smaller ones are mostly artifacts which we want to smooth out. Use of such a regularizer leads to non-linear least squares problem, which can be solved by the Gauss-Newton method [28] or other iterative methods.

In practical multi-dimensional cases, the Gauss-Newton method needs to produce an estimate of all the pixels at each iteration, thus the complexity is very large. One idea is to represent the object by efficiently using some basis functions to aggregate the pixels. A straightforward solution is to use shape-based template if the shape of the objects are known [29]. However, in most cases, the structure of the object is complex and unknown to us, so a general basis which can be suitable to a wide variety of objects should be used.

B-Splines have been thoroughly explored and proved to be very useful in approximation and have been widely used in computer aided designs [30] [32] [31]. The B-Spline basis has several good properties and can be easily manipulated by dealing with knots. Furthermore, some efficient algorithms of construction, evaluation, and knot insertion/deletion exist [36] [37] [38] [39], so that not too much effort is needed to deal with B-Splines themselves.

If proper B-Splines are chosen so that the object can be efficiently represented using few basis functions, the dimension of the problem will be significantly decreased. To select those proper B-Splines, we should find an appropriate number of B-Splines to be used and a suitable distribution of these B-Splines. The number of basis functions to be used is the number of unknowns in the revised problem, thus determining the

complexity of the B-Spline-based problem. The placements of these knot represent where to put those B-Splines, which affects how close the object can be approximated by using these fixed number of knots. The intuition is to place few knots at flat regions and more knots at rough regions with more details [40] [41] [42]. In this thesis, an adaptive knot refining and pruning algorithm is developed to solve the problem. The knot refining part efficiently inserts knots based on curvature information. The knot pruning part is a greedy search algorithm to find and delete redundant knots based on an estimate of the weight of each knot. The overall effect is to select proper number of knots to be used and, at the same time, relocate these knots to find a suitable distribution. And as the algorithm converges, the final result uses many fewer knots than pixels to represent the object while keeping the error of estimate within certain tolerance. A 2D cross-well tomography example will be provided to test the algorithm. The simulation results show that the B-Spline method reduces the dimensional complexity by 90%, compared with that of the pixel-based method, and decreases time complexity by 50%, without significantly degrading the estimation.

1.1 Contribution

In this thesis, we propose using B-Splines to represent the object to be reconstructed. The goal is to decrease the large complexity brought by the edge-preserving regularizer, and adaptively find the solution. Our contributions are mainly in the following aspects:

- We make a thorough analysis of the edge-preserving regularizer, especially the CRB(Cramér-Rao Bound) analysis. The result is very interesting. The edge-preserving regularizer does preserve edges better compared with Tikhonov regularization. As to the variance of estimate, however, the edge-preserving regularizer performs worse than Tikhonov one. This partly reflects the difference

between the two regularization methods. Tikhonov regularization add equally constraint on all areas, while an edge-preserving regularizer puts more effort on locating the edges. The price is that the variance of estimate is increased.

- We use B-Splines to represent the object and remodel the inverse problem. By finding the optimum knot distribution, the result has a significant decreasing in dimension. More efforts are made on how to adaptively find the optimum knot distribution. A knot refining and pruning algorithm is developed to automatically find the edges as well as select proper B-Splines.

- For knot refining, a curvature based knot insertion method is developed to intelligently insert knots. The intuition is that the optimum knot distribution should reflect the underlying smoothness structure of the object. Specifically, more knots should cluster at regions with large curvatures, while fewer knots are needed for flat regions. Therefore, simply doubling the number of knots will bring overhead for the insertion, i.e. most knots inserted in flat regions will have to be deleted later. Assuming a current coarse estimate is a good hint for further refined structures, we can use the curvature information as a guide to only insert new knots in areas of potential edges. By doing so, the knots inserted will be mostly useful and need not to delete them later. In the 2D case, the idea can be further refined by using principle directions used in computing curvatures. The principle directions are directions where the minimal and maximal curvature lie on a surface. If at a point, the large curvature only happens along x dimension, and in the y dimension, the curvature is very small, we will only want to increase resolution along the x dimension. By using the principle directions, we will be able to determine whether to insert knot in one dimension or both.

- For knot pruning, we should find which knots need to be deleted. A greedy

search algorithm is developed using the estimate of the weight of each knot. The weight of each knot is measured as the change in reconstruction when deleting the knot. The more the change, the larger the weight. Knots with small weight are considered redundant, i.e. they were previously placed in regions where enough knots had already been assigned. With the weight information, we will be able to sort knots in the order of their importance and delete them one at a time until the error of resulting reconstruction (computed by the cost function of the edge-preserving regularizer) exceeds a certain tolerance. We also notice the vicinity effects of each knot, i.e. deleting a knot changes the actual weight of adjacent knots. An efficient solution is to avoid deleting adjacent knots.

1.2 Organization

This thesis is organized into five chapters; the main body is Chapter 2, 3 and 4. Chapter 2 introduces related background of the problem, including linear inverse problem and Tikhonov regularization, B-Splines basis and geometric computations. Chapter 3 investigates the edge-preserving regularization method and makes some analysis. Chapter 4 presents the developments of the adaptive B-Splines methods incorporated with edge-preserving regularizer solving the image reconstruction problem.

Chapter 2 introduces some related background knowledge of this thesis. Firstly, the cross-well tomography problem is set up. The electromagnetic inverse scattering theory is introduced to solve the problem, and the Born approximation is used to get the linear inverse problem model. Secondly, the Tikhonov regularization method is introduced to solve the problem. The idea of edge-preserving regularization is proposed after analyzing the smoothing effects of Tikhonov regularization. Thirdly, we introduce the basics of B-Spline basis. With the good properties it possesses, we

propose using it to decrease the complexity of the problem. Lastly, some advanced geometric computation of B-Splines are introduced. The most important is computing the curvature of B-Spline surfaces, which will be helpful to our knot refining algorithm developed in Chapter 4.

Chapter 3 analyzes the deficiency of Tikhonov regularization. It then introduces the edge-preserving regularizer using non-quadratic regularization functions. The common properties of regularization functions are introduced. Furthermore, a CRB analysis is provided to examine the estimation properties of the edge-preserving regularizer. Finally, we apply the edge-preserving regularizer to solve the cross-well tomography problem.

Chapter 4 first points out the large complexity the edge-preserving regularization brings, then B-Spline basis is proposed to be used. By representing the object by B-Splines, the inverse problem is remodeled. It is shown that if the B-Splines are carefully selected, we can obtain a good reconstruction while significantly decreasing the dimension of the problem. An adaptive knot refining and pruning algorithm is developed to automatically find a suitable knot distribution. The knot refining part involves knot insertion under the guidance of curvature information; the knot pruning part involves knot deletion based on the estimate of weight of each knot. As the algorithm converges, which is generally within 10 iterations, the result will be the fewest possible number of knots that can be used to make an estimate within a certain error. Also, the 2D example used in Chapter 3 is revisited here to illustrate the algorithm.

Chapter 5 discusses some conclusions of this thesis and some possible ideas for the future work.

Chapter 2

Background

2.1 Cross-well Tomography Problem

In this section, we will set up the cross-well tomography problem which is the motivation of this thesis. We will revisit the problem as an example later in the following chapters to examine our approach.

2.1.1 Problem Setup

The problem comes from the application of finding underground pollution plumes. It is proposed to use a cross-well tomography method: one array of electromagnetic transmitters and one array of receivers are placed in two wells along the boundaries of the region of interest(Fig. 2.1). By measuring the scattered electromagnetic waves at receivers, we can determine the physical properties, such as conductivity, permittivity and permeability, of the medium in the region. Since the pollution plumes will have different electromagnetic properties than common soils, we will be able to determine the existence and shape of the plumes.

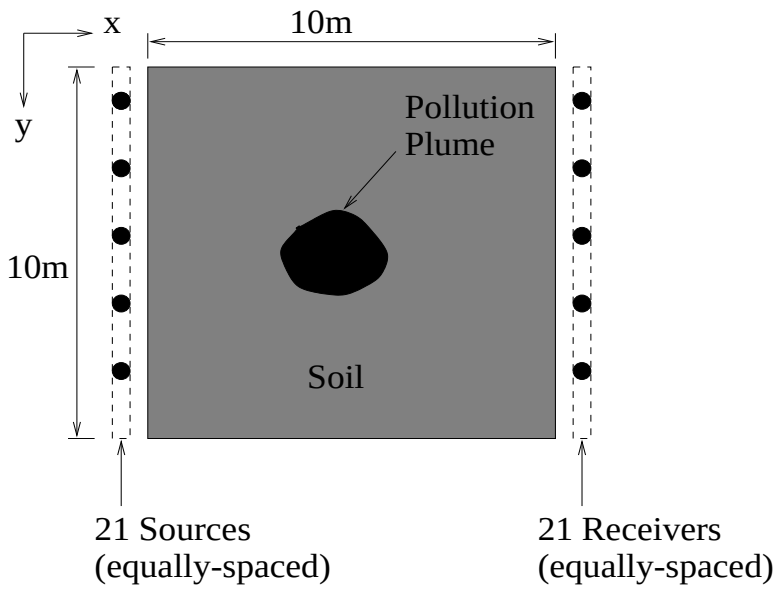


Figure 2.1: Cross-well tomography problem.

For the work in this thesis, as shown in Fig. 2.1, 21 equally-spaced dipole transmitters and receivers are placed along the boundaries respectively. The frequency of electromagnetic transmitted is $\omega = 5 \times 10^7 Hz$. To represent the electromagnetic properties of the region, we use typical settings for wet soils, where electric permittivity $\varepsilon = 20$, magnetic permeability $\mu = \mu_0$ is the permeability of free space, and electric conductivity $\sigma = 0.01 s/m$.

2.1.2 Electromagnetic Inverse Scattering

Electromagnetic scattering theory has played an important role in many applications, which are concerned with the effect an inhomogeneous medium has on an incident wave. More specifically, it involves two aspects of the problem, one is *direct scattering problem*, which is to determine the scattered field from the incident field; another is *inverse scattering problem* we will be focusing on, which is to determine the

nature of the inhomogeneity from the knowledge of the scattered field. Lots of work has been done in this region and many different approaches exist [8] [9] [10]. The inverse scattering problem is important in cases where details about the structure or composition of an object are required but can not be measured directly, instead, the measurements are often taken remotely without affecting the object. Thus, inverse scattering has been applied in many areas such as radar/sonar, medical diagnostics, and nondeconstructive testing [11] [12] [13].

The cross-well tomography problem we have set up can be shown, from the wave scattering point of view, in Fig. 2.2.

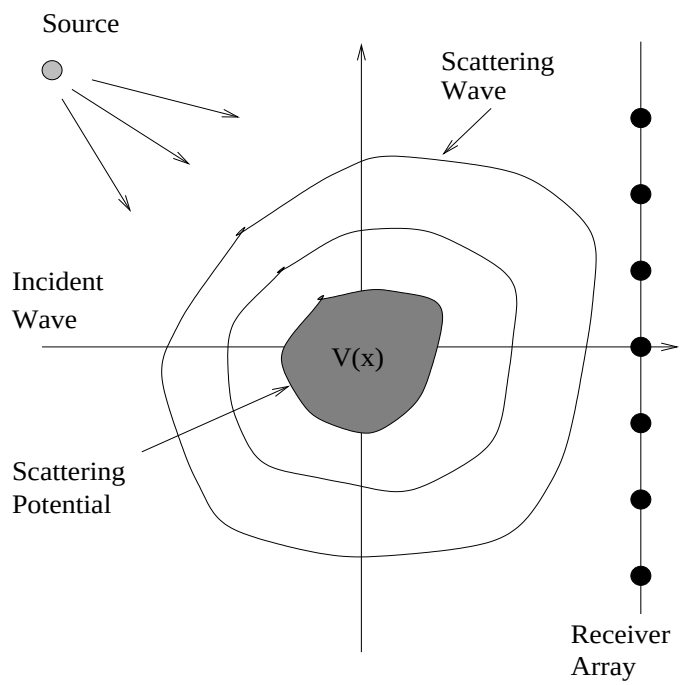


Figure 2.2: Cross-well tomography problem.

where we consider a finite scattering region mathematically described by a scattering potential $V(x)$ embedded in a homogeneous medium of permittivity ε_1. Thus the distribution of permittivity through the scatterer can be written as:

$$\varepsilon(\mathbf{x}) = \varepsilon_1 + \triangle\varepsilon(\mathbf{x}) \tag{2.1}$$

where $\triangle\varepsilon(\mathbf{x}) = 0$ outside the scatterer. With k^2 represents the wave number, we have the non-homogeneous Helmholtz equation [8]:

$$\nabla^2\Psi + k^2(\varepsilon_1 + \triangle\varepsilon(\mathbf{x}))\Psi = 0 \tag{2.2}$$

where θ is the incident angle of the plane wave and ∇^2 is the square of the gradient operator:

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.3}$$

Now we introduce the Green's function which will translate equation (2.2) into an integral form. The Green's function $g(\mathbf{x} \mid \mathbf{x}')$ is defined as the solution to the following equation [14]:

$$\nabla^2 g(\mathbf{x}, \mathbf{x}') + \mathbf{k^2}\varepsilon_\mathbf{1}\mathbf{g}(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}') \tag{2.4}$$

where $\mathbf{x}'$ is the coordinate of source and $\mathbf{x}$ is the position to be measured with respect to $\mathbf{x}'$. And it has been proved [8] if the Green's function satisfies certain boundary conditions, it possesses an important property:

$$g(\mathbf{x}', \mathbf{x}'') = \mathbf{g}(\mathbf{x}'', \mathbf{x}') \tag{2.5}$$

Thus, equation (2.2) can be rewritten as:

$$\begin{cases} \Psi(\mathbf{x}) = \Psi^{inc}(\mathbf{x}) + \int k^2 \triangle\varepsilon(\mathbf{x}')g(\mathbf{x}, \mathbf{x}')\Psi(\mathbf{x}')d\mathbf{x}' & \text{(a)} \\ \Psi^{scat}(\mathbf{x}) = k^2 \int g(\mathbf{x}, \mathbf{x}')\triangle\varepsilon(\mathbf{x}')\Psi(\mathbf{x}')d\mathbf{x}' = \text{data} & \text{(b)} \end{cases} \tag{2.6}$$

The above equation can be solved provided all the boundary conditions are satisfied by both $\Psi(\mathbf{x}')$ and $g(\mathbf{x}, \mathbf{x}')$ are given. And the field Ψ here is actually the combination of the local incident field Ψ^{inc} and Ψ^{scat}, i.e.

$$\Psi = \Psi^{inc} + \Psi^{scat} \tag{2.7}$$

2.1.3 Born Approximation

In practice the integral equations in (2.6) bring into a non-linear structure, because Ψ in (2.6)(b) depends on $\triangle\varepsilon$ in (2.6)(a). It usually does not have close form solution. Thus, some approximate solution should be considered, among which the Born approximation [15] [16] has become the prevalent technique because of its easy implementation and simple physical interpretation. It replaces $\Psi(\mathbf{x})$ in (2.6)(b) with $\Psi^{inc}(\mathbf{x})$ which yields to:

$$\Psi^{scat}(\mathbf{x}) = \int_{\tau'} k^2 g(\mathbf{x}, \mathbf{x}') \triangle\varepsilon(\mathbf{x}') \mathbf{\Psi^{inc}}(\mathbf{x}') \mathbf{d}\tau' \tag{2.8}$$

The approximation has been proved to be good if $\triangle\varepsilon(x)$ is small. So the Born approximation results in the form of a first-kind Fredholm integral equation, which can be discretized furthermore to a finite dimension matrix problem of the form [5]:

$$y = Af + n \tag{2.9}$$

where each element of vector y is the scattered field observations obtained along the receiver array; A is a matrix associated with the integral kernel; f corresponds to the conductivity perturbation and n represents additive noise.

Therefore, the cross-well tomography problem has been translated into the form of a typical linear inverse model. Further details of such model will be discussed in the following section.

2.2 Regularization of Linear Inverse Problems

In this section, we will investigate the ill-posed property of discretized linear inverse problems and the energy-based regularization method as a solution. Based on these discussions, we will be able to address the edge-preserving regularization techniques in Chapter 5.

2.2.1 Linear Inverse Problems

Generally, we can describe a linear inverse problem as a linear process with a certain distortion model which has the following form:

$$y = Af + n, \tag{2.10}$$

where A is a known degradation process; n is additive, Gaussion noise; y is the degraded signal. Our problem is to obtain an estimate of the real object f, which is usually mapping of some physical properties, from the data vector y.

2.2.2 Nature of Ill-posed Problems

A first and obvious solution to the equation (2.10) we established above would be least squares solution:

$$f_{LSQ} = \left[A^T A\right]^{-1} A^T y = \left[A^T A\right]^{-1} AA^T f + \left[A^T A\right]^{-1} A^T n \tag{2.11}$$

but for practical problems, the above approach does not make sense, i.e. f_{LSQ} tends to be contaminated by large amplitude high-frequency artifacts, and this is much different from the true object. Those non-physical artifacts of the solution f_{LSQ} comes from the ill-conditioned nature of degradation kernel A, i.e. the condition

number of A is always very large in practice which means the columns of A are nearly linearly dependent. Though we can replace A by a well-conditioned matrix derived from A, we still can not get a useful solution. In another word, we need to filter out the high-frequency artifacts to get useful solution. While in practice, generally A is a degradation process which is a kind of lowpass process, thus $\left[A^T A\right]^{-1}$ is a highpass filter which would only amplify the high-frequency noise in the data.

2.2.3 SVD Analysis

To gain in-depth understanding, we will analyze the ill-posed problems and regularization methods from the SVD (Singular Value Decomposition) point of view.

The Singular Value Decomposition of A is as the following:

$$A = U\Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T \tag{2.12}$$

where $\Sigma = diag(\sigma_1, ...\sigma_n), \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$

Thus, the condition number of A is equal to σ_1/σ_n. For ill-posed problems, the singular values σ_i typically decay gradually to zero, and u_i and v_i tend to have more sign changes as σ_i decreases. This shows the nature of ill-posed problem in two aspects: 1. In ill-posed problems, A is highly ill-conditioned, i.e. A has large condition number σ_1/σ_n. The vectors v_i associated with the small (close to zero) σ_i are numerical null-vectors of A and have many sign changes. 2. Af is a smooth process of f and the inverse problem amplifies the high-frequency oscillations in the right-hand side y. It is difficult to get an acceptable solution, because small errors in data may cause large errors in the solution. A simple 1D example is shown in Fig. 2.3.

We can also draw the same conclusion from the following mathematical analysis:

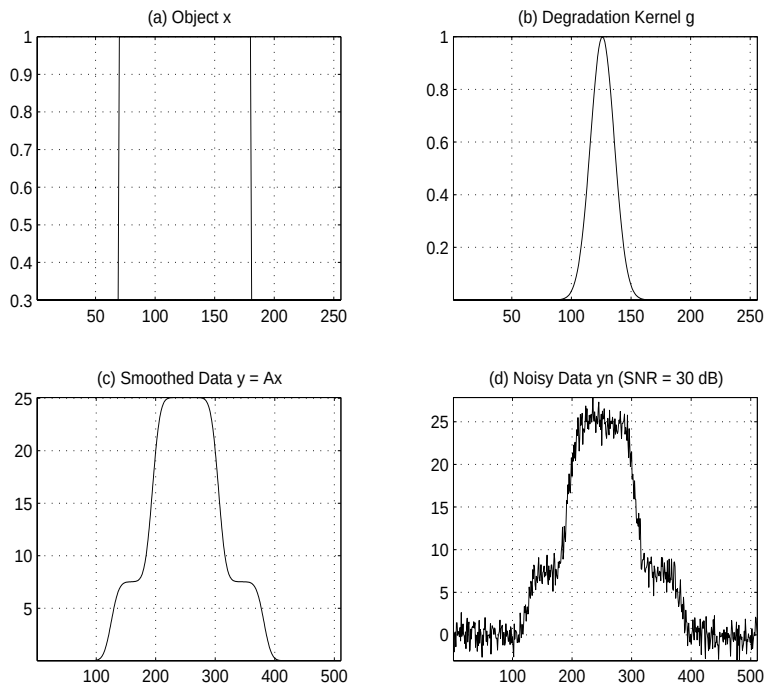


Figure 2.3: A simple 1-d inverse problem, with a Gaussian function as the kernel.

Consider linear least-squares problems:

$$\min_f \|Af - y\|_2, A \in \mathcal{R}^{m\times n} m > n \tag{2.13}$$

from the SVD of A in equation (2.12), the least-squares solution will be:

$$f_{LSQ} = \sum_{i=1}^{n} \frac{u_i^T y}{\sigma_i} v_i \tag{2.14}$$

Apparently, since u_i, v_i has many sign changes with σ_i degrades to zero, the solution f_{LSQ} has many large components of high-frequencies. Therefore, f_{LSQ} appears to be random and is not acceptable.

From the above analysis, the idea of regularization can be naturally induced.

Specifically, the purpose of regularization would be to filter out the contributions of the small singular values to the solution. One well-known regularization method is Tikhonov regularization [21]. The idea is to define the regularized solution f_λ as the minimizer of the following combination of the residual norm and the side constraint:

$$f_{Tik} = \underset{f}{argmin} \|Af - y\|_2^2 + \lambda^2 \|L(f - f^*)\|_2^2, \tag{2.15}$$

where the *regularization parameter* λ controls the weight of compromise of fidelity of solution to the residual norm and to the side constraint norm. Generally, if $L = I_n$ and $f^* = 0$,

$$f_{Tik} = \sum_{i=1}^{n} g_i \frac{u_i^T y}{\sigma_i} v_i \tag{2.16}$$

where, $g_i = \sigma_i^2/(\sigma_i^2 + \lambda^2)$. Hence the filter factor g_i counters the contributions of small σ_i.

2.2.4 L-Curve

In 2.15, the regularization parameter λ is critical and need to be wisely chosen. L-curve is an convenient way to find proper λ [23]. Given 2.15, we can draw out a L-curve as shown in Fig. 2.4.

The vertical part of the L-curve corresponds to solutions where $||Lf_{Tik}||_2$ is very sensitive to changes in the regularization parameter λ. The horizontal part of the L-curve corresponds to solutions where it is the residual norm $||Af_{Tik} - b||_2$ that is most sensitive to the regularization parameter λ. Thus, we can find the optimal regularization parameter by locating the corner of this L-curve. This is the general procedure of solving the Tikhonov regularization problem. Using the previous 1D example, some results corresponding to different regularization parameters are shown in Fig. 2.5.

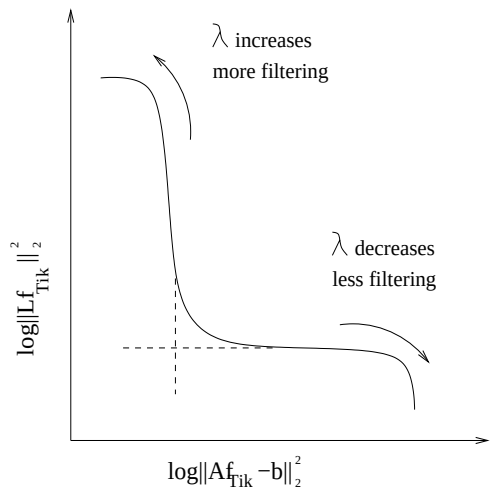


Figure 2.4: Illustration of L-curve

2.2.5 General Regularizations

In general, regularization method can be written as the following:

$$f_{reg} = \underset{f}{argmin} \|Af - y\|_2^2 + \lambda\Omega(f) \tag{2.17}$$

In Tikhonov regularization, $\Omega(f)$ is of the form of 2-norm or derivative of it. This quadratic regularization function adds large constraint on large frequency components of the solution, thus leads to a 'smooth' solution. Therefore, Tikhonov regularization is not good for reconstruction of objects with sharp edges which are critical to describe the shape and internal structures of the objects for our cross-well problem. Since in practice sharp edges of an object usually have very large frequency components, it is natural to think of applying non-quadratic regularization functions which can smooth out relatively small perturbations while preserving larger ones, thus preserve edges. In Chapter 5, we will discuss the edge-preserving regularization in details.

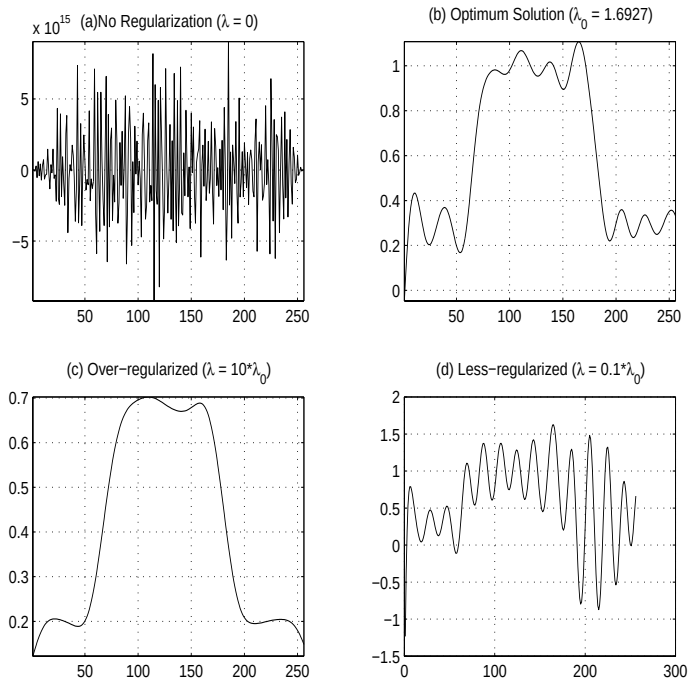


Figure 2.5: Tikhonov regularization effects for the 1-d problem.

2.3 Introduction to B-Spline Functions

In this section, we will introduce some background material about B-Spline functions. First, we will define the B-Spline basis to represent a certain function or object. Then we will talk about how the knot distribution affect the properties of the representation, which is the motivation of using B-Splines in this problem. Also we will expand it to 2D case by introducing tensor-product B-Splines to fit in our application. Lastly, we will discuss the geometries of B-Splines, which are required in calculating the differentials and curvatures.

2.3.1 B-Spline Basics

B-Splines are widely used in many aspects of numerical analysis (statistical data interpolation, data smoothing, numerical solution of inverse problem with integral model, computer aided geometric design, etc. [32] [33] [34]). There exists a set of well-developed theories on B-Splines, their properties and applications, including some efficient algorithms for computation using B-Splines [36] [37] [38] [39]. Also, B-Splines are a strong candidate for use in our inverse problem, especially considering their local support and smoothness property which we will discuss later.

Definition 1 [32] A function $s(x)$, defined on a finite interval $[a, b]$, is called a *spline function* of *degree* $k > 0$ (*order* $k + 1$), having *knots* as the strictly increasing sequence $\lambda_j, j = 0, 1, ..., n(\lambda_0 = a, \lambda_n = b)$, if the following conditions are satisfied:

1. $s(x) \in \mathcal{P}_k,$ for $x \in [\lambda_j, \lambda_{j+1}],\ j = 0, 1, ..., n.$
2. $s(x) \in \mathcal{C}^{k-1}[a, b].$

where $\mathcal{P}_k$ is the set of all polynomial functions of k orders, $\mathcal{C}^{k-1}$ is the set of all $k - 1$ continuous functions.

Definition 2 The *B-Spline* $N_{i,k+1}$ of degree k with knots $\lambda_i, ..., \lambda_{i+k+1}$ is defined as :

$$N_{i,k+1}(x) = (\lambda_{i+k+1} - \lambda_i) \sum_{j=0}^{k+1} \frac{(\lambda_{i+j} - x)_+^k}{\prod_{l=0, l\neq j}^{k+1} \lambda_{i+j} - \lambda_{i+l}} \tag{2.18}$$

where

$$(\lambda - x)_+^k = \begin{cases} (\lambda - x)^k, & \text{if } \lambda > x; \\ 0, & \text{otherwise.} \end{cases} \tag{2.19}$$

Based on the above definition, B-Splines have the following important properties:

1. *Positivity and Local Support :*

$$\begin{aligned} &N_{i,k+1}(x) \geq 0 \text{ for all } x. \\ &N_{i,k+1}(x) = 0 \text{ if } x \notin [\lambda_i, \lambda_{i+k+1}]. \end{aligned} \tag{2.20}$$

2. *Boundary values :*

$$N_{i,k+1}^{(l)}(\lambda_i) = N_{i,k+1}^{(l)}(\lambda_{i+k+1}) = 0, l = 0, ..., k-1. \tag{2.21}$$

where $(\cdot)^{(l)}$ stands for the l order derivative of the function.

3. *Minimal support :* If a piecewise polynomial with the same smoothness property over the same knot vector has less support than N_i^{k+1}, it must be the zero function.

4. *Linear Independence :*

$$\sum_{i=0}^{n-1} c_i N_{i,k+1}(x) = 0 \text{ implies } c_i = 0 \text{ for all } i. \tag{2.22}$$

Some simple examples of B-Splines are shown in Fig. 2.6

The *B-Splines* form a *basis* for the space consist of piecewise polynomial functions; every piecewise polynomial function $s(x)$ over $[a, b]$ has a unique representation:

$$s(x) = \sum_{i=-k}^{n-1} c_i N_{i,k+1}(x), \tag{2.23}$$

where $N_{i,k+1}(x)$ for $i = -k, ..., -1$ and $i = n-k, ...n-1$ are generated as before by introducing 'arbitrary' additional knots $\lambda_{-k} \leq \lambda_{-k+1} \leq ... \leq \lambda_0 = a$, and $b = \lambda_n \leq \lambda_{n+1} \leq ... \leq \lambda_{n+k}$. The reason of using these additional knots is that each B-Spline has support on several adjacent knots (As shown in Fig. 2.6). Thus, the B-Splines defined on these additional knots also have support on the knots at or near the ends. Different choices of these additional knots will bring different ending effects on the curves/surfaces[32].

As we shall see in the next chapter, these basic properties make B-Splines a good

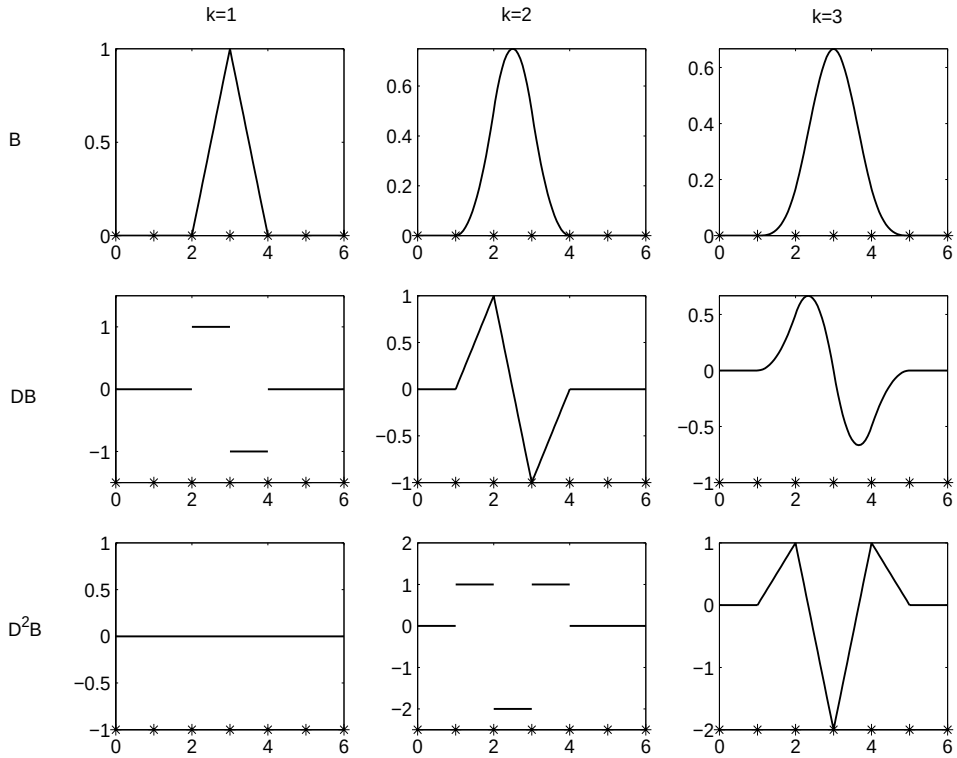


Figure 2.6: B-Splines $N_{i,k+1}$ with different order $(k+1)$ and their first and second derivatives. ($i = 3$, knots sequence $\lambda = [0123456]$)

candidate to be used in our inverse problem to obtain multiscale solutions. Furthermore, there are efficient algorithms [36] for evaluating the values and derivatives of B-Splines.

To fit into our 2D problem, we need to define Bivariate B-Splines. There are different ways in extending the univariate splines to 2D case. *Tensor product B-Splines* are the most widely used due to their simplicity.

Definition 3 Having strictly increasing sequences :

$$\begin{aligned} a = \lambda_0 < \lambda_1 < ... < \lambda_m = b, \\ c = \mu_0 < \mu_1 < ... < \mu_n = d, \end{aligned} \tag{2.24}$$

The function $s(x, y)$ on $R = [a, b] \times [c, d]$, of degree $k > 0$ in x and $l > 0$ in y, with knots $\lambda_i, i = 0, 1, ..., m$ in the x-direction and knots $\mu_j, j = 0, 1, ..., n$ in the y-direction, is called a *Tensor Product Spline*, if the following conditions are satisfied:

$$\begin{aligned} &1.\ s(\lambda_i, \mu_j) \in \mathcal{P}_k \otimes \mathcal{P}_l, i = 0, 1, ..., m; j = 0, 1, ..., n. \\ &2.\ \frac{\partial^{i+j} s(x, y)}{\partial x^i y^j} \in \mathcal{C}(R), i = 0, 1, ..., k-1; j = 0, 1, ..., l-1. \end{aligned} \tag{2.25}$$

Similarly, let the $N_{i,k+1}$ and $M_{j,l+1}$ are the normalized univariate B-Splines defined as before, on the knot sequences λ and μ respectively, we can uniquely represent every spline $s(x, y) \in (\lambda_0, ..., \lambda_m; \mu_0, ..., \mu_n)$ as the following :

$$s(x, y) = \sum_{i=-k}^{m-1} \sum_{j=-l}^{n-1} c_{i,j} N_{i,k+1}(x) M_{j,l+1}(y), \tag{2.26}$$

Some examples of tensor product B-Splines are shown in Fig. 2.7.

Tensor product B-Splines retain all the good properties of univariate B-Splines. Especially, since the basis corresponding to x-direction and y-direction are separable, it can be easily handled and the efficient algorithms in 1D case can be directly employed. A certain drawback would be that only rectangular approximation patches can be applied thus the refinement procedure are tied to existing rectangular grid. There are other approaches using triangular patches which bring more freedom and adaptation [49] [50]. We will leave this effort to future work.

2.3.2 Knots Distribution and Smoothness

As stated before, given a certain knot sequence, we can uniquely define a set of B-Spline basis which are piecewise polynomials of degree k. We can combine the B-Spline curves/surfaces to approximate a wanted curve/surface. It can be proved [35]

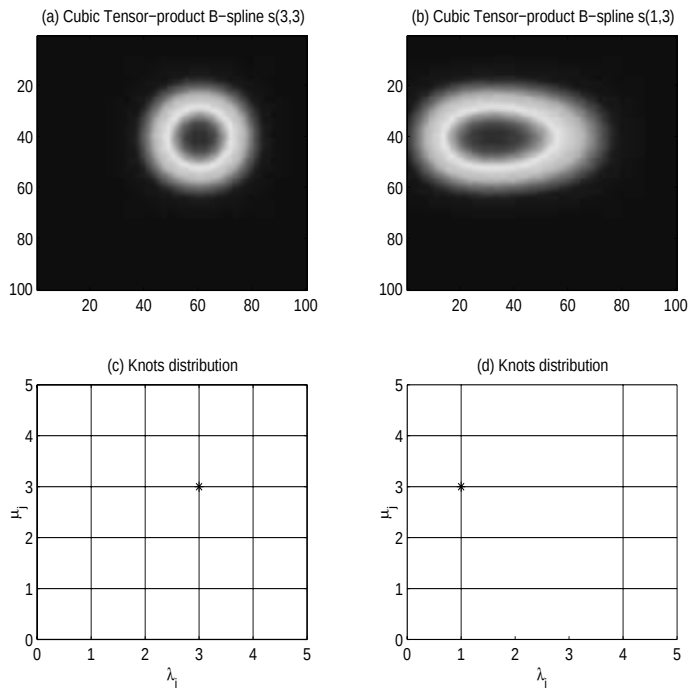


Figure 2.7: Examples of tensor product B-Splines and corresponding knots distribution.

that if sufficiently many knots are inserted into the knot sequence, the resulted approximation will be arbitrarily close to the curve/surface. Thus instead of describing the curve/surface directly, a set of knots would be sufficient. This is the whole basis of using B-Splines to approximate curves/surfaces. In practice, we are interested in finding the 'best' approximation which uses as few knots as possible within a certain error tolerance. A simple example is given in Fig. 2.8.

The sample data are Titanium Heat data which gives a certain property of titanium as a function of temperature. It has been used extensively as test data for Spline fitting. As the number of knots increases, the resulted Spline curves are closer to real data in the means of Mean Square Error(MSE). In 2.8(d), 12 non-equally-spaced and well-selected knots are used which leads to a less MSE than that of Fig. 2.8(c) where

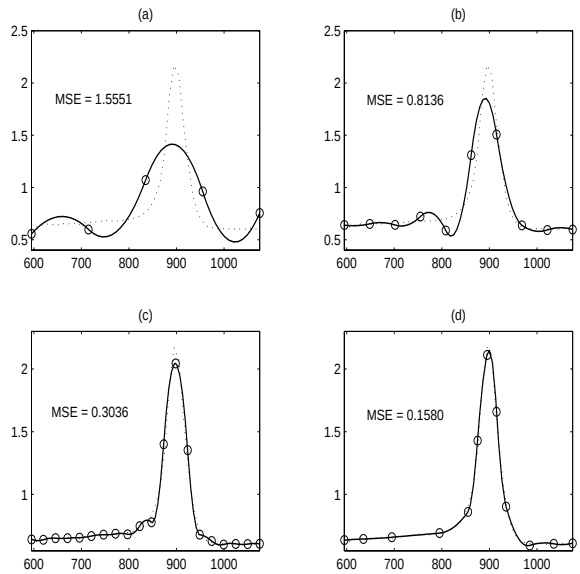


Figure 2.8: Illustration of using cubic B-Splines to approximate Titanium Heat data. (a)5 equally-spaced knots; (b)10 equally- spaced knots; (c)20 equally-spaced knots. (d)12 selected knots.

20 equally-spaced knots were used. This is a very interesting and helpful property of B-Splines approximation and is the motivation of using B-Splines in our adaptive reconstruction approach to solve the inverse problems.

Another important property of B-Splines is their local support. It allows the easy alteration of a complex curve/surface in one region without affecting the remote portion of the curve. Since each B-Spline is defined over a certain knot sequence $\lambda_0, ..., \lambda_n$, it is natural to see that if the knots are set apart, the B-Spline support will be large, consequently we will get relatively flat basis. On the contrary, if we set the knots to be very near to each other, the B-Spline will have very narrow support which lead to more local details (Fig. 2.9). Since the knots alteration will only locally affect the approximated function and remote part of the curve will remain the same, it is natural to assign a large density of knots to areas with significant fine scale variations.

Only a 'sparse' distribution of knots are needed to well approximate the flat part of the curve.

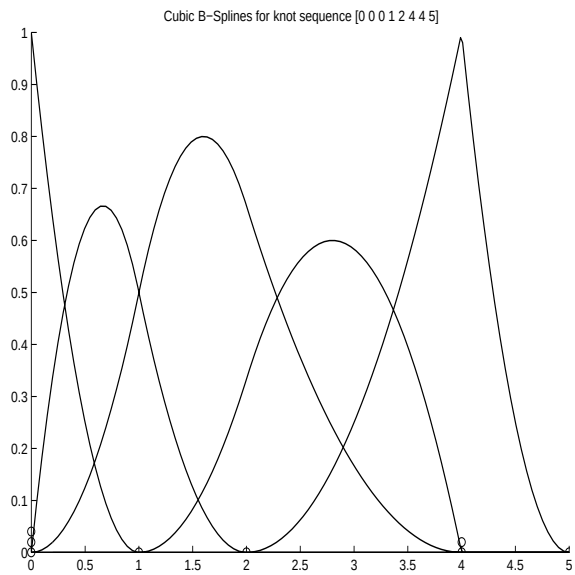


Figure 2.9: Distant knots leads to B-Spline of large support (flat) and vice versa; Certain multiple knots leads to discontinuity.

In Fig. 2.9, we also note that to construct cubic B-Splines, additional knots have been added at end points. At one end, we use multiple knots ([0 0 0]), which brings high-order discontinuity there; at the other end, we extend knots out of the interested region, i.e. add one knot at position 5, which makes the reconstruction at that end point more smooth. But the shortcoming is that we will have leakage outside the focused region.

As stated before, we are interested in efficiently representing a surface, i.e. a function of two variables. Our approach will be to use B-Splines with the fewest possible knots to best approximate the surface within a certain error tolerance. In this case, the approximation will essentially be adapted to the smoothness of the surface. Thus, the problem can be stated as to find a suitable set of knots based on

which the B-Splines can approximate the surface efficiently.

To find an adaptive procedure of achieving this goal, we will need to reallocate the knots to better reflect the underlying smoothness of the curve/surface. Curvature information will be a good hint to this further redistribution of knots. Specifically, if the knots is equally spaced, then, given a certain error tolerance, we will need to insert more knots to the regions of large curvatures and delete redundant knots at regions of small curvatures. The computation of knot insertion and deletion have been thoroughly discussed in many papers [37] [38] [39]. And in the following section we will focus on the geometries needed to compute spline curvature.

2.4 Geometries of Splines

As stated before, we will need to measure the smoothness of a reconstructed B-Spline curve/surface, then use it as a guide in further refinement. Specifically, we want to redistribute the knots in such a way that large density of knots are assigned to regions of large curvatures, and fewer knots to relative flat regions. In this section, we will introduce the curvature computing of B-Spline curves/ surfaces.

2.4.1 Curvature of Space Curves

Before discussing surface curvature, we introduce the geometric properties of curves.

In general, it is helpful to represent a space curve (in $\mathcal{R}^3$) by the following parametric form:

$$\mathbf{x} = \mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}, t \in [a, b] \subset \mathcal{R} \tag{2.27}$$

Furthermore, we can reparametrize the curve by:

$$s(t) = \int_{t_0}^{t} ds = \int_{t_0}^{t} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} dt = \int_{t_0}^{t} \|\dot{\mathbf{x}}(t)\| dt \tag{2.28}$$

where the dots denote derivatives with respect to t, and s is called *arc length* reparameterization of the curve [44](an example is shown in Fig. 2.10). From 2.28, we have $\dot{s}(t) = \|\dot{\mathbf{x}}(t)\|$, which is called *arc element* of the curve. Furthermore, it indicates that the variation rate of arc length with respect to curve parameter is the speed of the curve parameterization. For B-spline curves, where the parameterization is regular, i.e. $\|\dot{\mathbf{x}}(t)\| \neq 0$ [44], we may write $dt/ds = 1/\|\dot{\mathbf{x}}(t)\|$. Let $\alpha(s) = \mathbf{x}(t(s))$ be the arc length parameterized curve, we have:

$$\|\frac{d\alpha}{ds}\| = \|\frac{d}{dt}(\mathbf{x}(t(s)))\frac{dt}{ds}\| = \|\dot{\mathbf{x}}(t(s))\frac{1}{\|\dot{\mathbf{x}}(t(s))\|}\| = 1 \tag{2.29}$$

Hence, arc length reparameterization has unit speed for regular curves.

Now, we introduce a local orthonormal coordinate system, *the Frenet Frame* $\mathbf{t}, \mathbf{m}, \mathbf{b}$[44], which is very helpful to provide the local behavior of curves. By observing the changes of the Frenet Frame as a function of parameter t, we can get the curvature information of space curves. The Frenet Frame is the frame formed by *tangent vector* $\mathbf{t}$, *main normal vector* $\mathbf{m}$ and *binormal vector* $\mathbf{b}$, where :

$$\mathbf{t} = \frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|}, \quad \mathbf{b} = \frac{\dot{\mathbf{x}} \times \ddot{\mathbf{x}}}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|}, \quad \mathbf{m} = \mathbf{b} \times \mathbf{t}. \tag{2.30}$$

Given the Frenet Frame, we can construct the *osculating circle* [45] which lies in the osculation plane formed by $\mathbf{t}$ and $\mathbf{m}$, and has the same first and second derivative vectors as the curve. The inverse of the osculating circle radius ρ is called *curvature* κ.

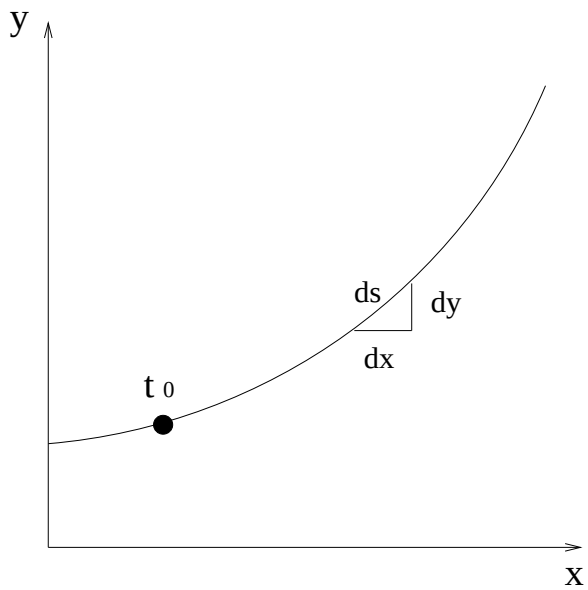


Figure 2.10: Arc length reparameterize of curve $\mathbf{x}(t)$

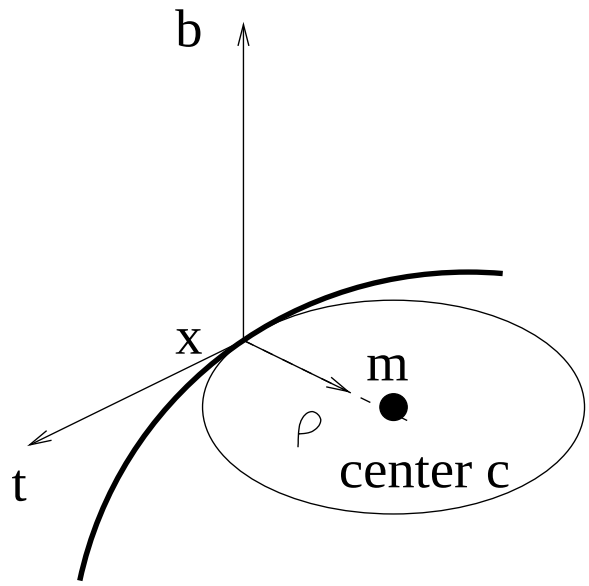


Figure 2.11: Franet Frame and Osculating Circle

By applying the arc length reparametrization, we have the following results:

$$\kappa = \kappa(s) = \|\mathbf{x}''\| \tag{2.31}$$

where the prime denotes derivative with respect to arc length s. In terms of the actual parameter t, we have:

$$\kappa = \kappa(t) = \frac{1}{\rho} = \frac{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|}{\|\dot{\mathbf{x}}\|^3} \tag{2.32}$$

The geometric interpretation of curvature is straightforward: consider the angle $\triangle\theta$ between two tangent vectors $\mathbf{t}$ and $\mathbf{t}(s + \triangle s)$, thus $\kappa = d\theta/ds$.

2.4.2 Curvature of Surface Curves

The above discussion can be extended to surfaces. Assume a regular parametric surface:

$$\mathbf{x} = \mathbf{x}(\mathbf{u}) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} \; ; \; \mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix} \in [\mathbf{a}, \mathbf{b}] \subset \mathcal{R}^{\mathbf{2}} \tag{2.33}$$

with

$$\mathbf{x}_u \times \mathbf{x}_v \neq 0 \text{ for } \mathbf{u} \in [\mathbf{a}, \mathbf{b}] \tag{2.34}$$

Consider a regular curve $\mathbf{x}(\mathbf{u}(t))$ on the surface, the squared arc element (*first fundamental form*) is defined as the following [44]:

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2 \tag{2.35}$$

where

$$E = \mathbf{x}_u \cdot \mathbf{x}_u, \; F = \mathbf{x}_u \cdot \mathbf{x}_v, \; G = \mathbf{x}_v \cdot \mathbf{v}_v \tag{2.36}$$

Analogous to the space curves, we can also define a frame $\mathbf{x}_u, \mathbf{x}_v, \mathbf{n}$ for the surface. The partial $\mathbf{x}_u$ and $\mathbf{x}_v$ at a point $\mathbf{x}$ span the tangent plane to the surface. The normalized normal $\mathbf{n}$ [45] to the surface is:

$$\mathbf{n} = \frac{\mathbf{x_u} \times \mathbf{x_v}}{\|\mathbf{x_u} \times \mathbf{x_v}\|} \tag{2.37}$$

Let $\mathbf{u}(t)$ be a curve on the surface $\mathbf{x}(\mathbf{u})$ in the direction of $\mathbf{t}$ as defined in 2.30. In 2D case, consider $\mathbf{x}_u = d\mathbf{x}/du, \mathbf{x}_v = d\mathbf{x}/dv$, the direction of $\mathbf{t}$ can be represented by $\lambda = dv/du$) (Fig. 2.12), applying the above definition of space curve curvature to the surface curve, the normal curvature at point $\mathbf{x}$ is given by:

$$\kappa(\lambda) = \frac{L + 2M\lambda + N\lambda^2}{E + 2F\lambda + G\lambda^2} \tag{2.38}$$

where

$$L = \mathbf{n} \cdot \mathbf{x}_{uu},\ M = \mathbf{n} \cdot \mathbf{x}_{uv},\ N = \mathbf{n} \cdot \mathbf{x}_{vv} \tag{2.39}$$

Generally, κ always changes as λ changes, thus for a given point $\mathbf{x}$ on the surface, the normal curvatures corresponding to different surface curves which pass the same point $\mathbf{x}$ are different. This result provides useful information on the smoothness of surface along different directions.

2.4.3 Gaussian and Absolute Curvature

The definition of normal curvature of a surface, $\kappa(\lambda)$ is of rational quadratic form. Thus the extreme values κ_1 and κ_2 of $\kappa = \kappa(\lambda)$, which are called *principal curvatures* of the surface, are the roots of

$$det \begin{bmatrix} \kappa E - L & \kappa F - M \\ \kappa F - M & \kappa G - N \end{bmatrix} = 0 \tag{2.40}$$

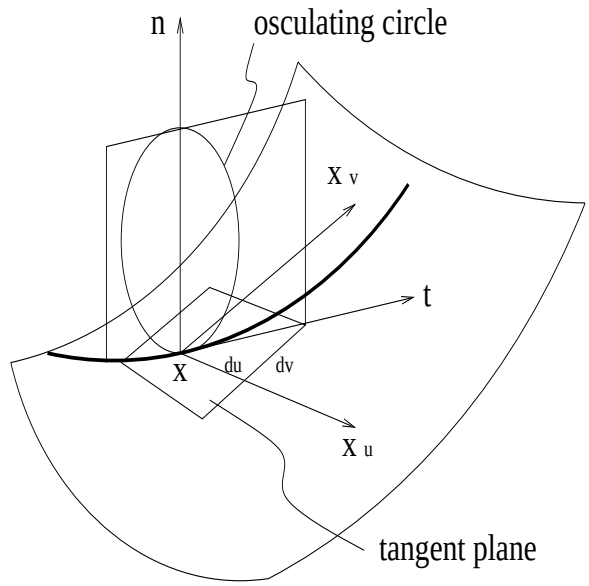


Figure 2.12: Frame for surface

And the corresponding quantities λ_1 and λ_2 define the *principal directions*, which are the surface curve directions in the tangent plane.

Having the principal curvatures, we have two important definitions of the surface curvatures on point $\mathbf{x}$. One is *Gaussian Curvature*, which is

$$K_{Gaussian} = \kappa_1 \kappa_2 = \frac{LN - M^2}{EG - F^2} \tag{2.41}$$

The other is *absolute curvature*:

$$K_{abs} = |\kappa_1| + |\kappa_2| \tag{2.42}$$

In our approach to the 2D inverse problem, we want to add new knots at places of large curvatures and remove knots at places of small curvatures. By using tensor-product B-Splines, we can add or remove knots independently in two directions. To

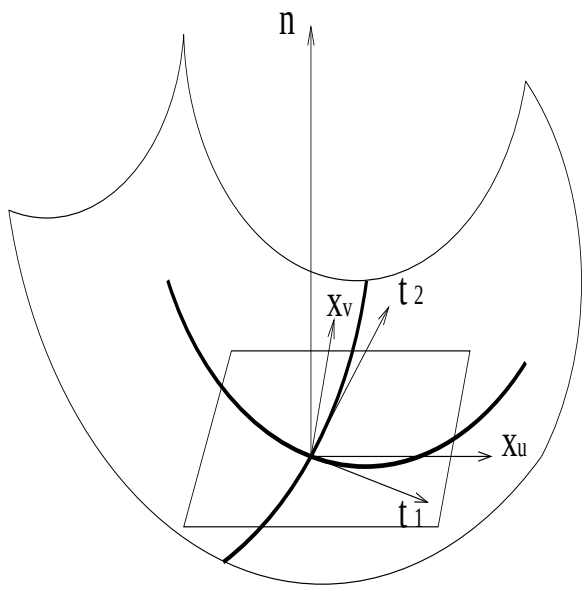


Figure 2.13: Example of only need to add knots in one direction

make the knot insertion/deletion procedure more efficient, we should decide whether to add or remove knots on a specific direction. For example, if we have a surface bending along u-direction and is relatively flat along v-direction (Fig. 2.13), the Gaussian or absolute curvature maybe large so that we would be tempted to add new knots in both u and v-directions. Apparently, there will be a waste for all the knots inserted on u-direction. To solve the problem, we think of using the principal directions, which are two directions correspond to the largest and smallest curvature respectively. If the principal direction of the largest curvature κ_1 is almost along the u-direction (t_1 in Fig. 2.13), we will only consider adding knots in the u-direction. Because it can be induced from the principal direction information that the measured surface is relatively flat at v-direction, thus no knots need to be added there. Further details will be discussed in Chapter 4.

Chapter 3

Edge-preserving Regularizer

In this chapter, we will continue the discussion on regularization methods in Chapter 2. A thorough discussion will show that the Tikhonov regularization tends to bring smoothing effects to the reconstruction thus is not good at preserving the edges. Then the idea of edge-preserving regularization will be introduced and analyzed. Some examples and comparisons will be provided at the end of this chapter.

3.1 Deficiency of Tikhonov Regularization

As introduced in Chapter 2, the general regularization method can be written in the following form:

$$f_{reg} = \underset{f}{argmin} \|Af - b\|_2^2 + \lambda\Omega(f) \tag{3.1}$$

In Tikhonov regularization, $\Omega(f)$ is of the 2-norm form : $\Omega(f) = \|Lf\|_2^2$, where L is a derivative operator. Thus the regularization term is defined as a sum of derivatives to the object f, which is a quadratic function (Fig. 3.1(a)) of Lf. In some cases, L is the first order derivative operator to f[19]. The idea is to counter the high-frequency artifacts by adding large constraint on them. The resulting reconstruction will be a

rather smooth one. The problem is that in many applications, the objects always have sharp edges which are very important for accurately determining the location and precisely describing the shape of discrete structures in the overall scene. In this case, the Tikhonov reconstruction will not be satisfying in that the smoothing effect blurs the edges as well as filters out the high-frequency artifacts. Note the different scales used in the four plots of Fig. 3.1, we can see the quadratic function brings significantly more constraints on large components of Lf, which usually correspond to sharp edges of the object. Therefore, we need to find other methods preserving edges better than Tikhonov regularization.

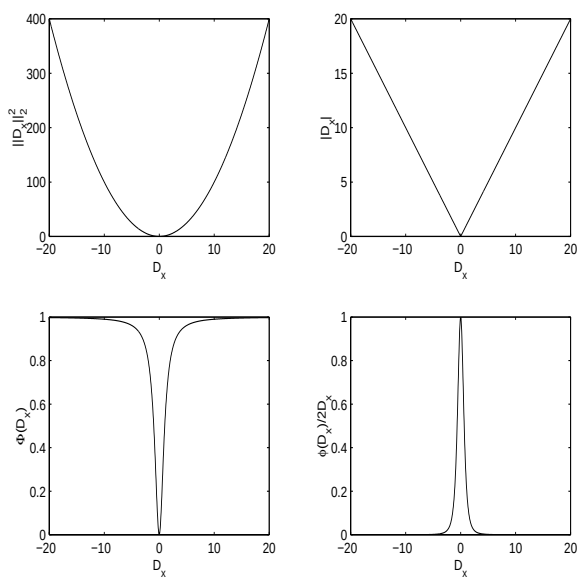


Figure 3.1: Illustrations of regularization functions (a)Quadratic (Tikhonov) (b)1-norm (c)Edge-preserving ($\varphi(t) = \frac{t^2}{1+t^2}$) (d)Edge-preserving ($\frac{\varphi'(t)}{2t} = \frac{1}{(1+t^2)^2}$)

3.2 Regularization Functions

A natural thought is to preserve larger fluctuations which are more likely to be edges of the object, and eliminate the smaller ones which are mostly to be high-frequency artifacts. To achieve this goal, we may want to apply 1-norm (Fig. 3.1(b)) instead of 2-norm to the regularization term, which will bring relatively smaller constrains on large components of Lf. A even better approach would be using the function as shown in Fig. 3.1(c) as the regularization term. In this case, only those regions with fluctuations exceeding a certain threshold will be considered to be edges and relatively little constraint will be applied there. The edge-preserving regularization can be written as:

$$\Omega_{ep}(f) = \sum_{i,j} \varphi[(D_x f)_{i,j}] + \sum_{i,j} \varphi[(D_y f)_{i,j}] \tag{3.2}$$

where D_x and D_y are first derivatives:

$$\begin{aligned} (D_x f)_{i,j} &= (f_{i,j+1} - f_{i,j}) \\ (D_y f)_{i,j} &= (f_{i+1,j} - f_{i,j}). \end{aligned} \tag{3.3}$$

It has been shown [24] that $\varphi(t)$ should have the following special properties in order to preserve the edges:

$$\begin{aligned} &\text{i) } \varphi'(t)/2t \text{ continuous and strictly decreasing on} [0, +\infty). \\ &\text{ii) } \lim_{t \to +\infty} \frac{\varphi'(t)}{2t} = 0 \\ &\text{iii) } \lim_{t \to 0^+} \frac{\varphi'(t)}{2t} = M, 0 < M < +\infty. \end{aligned} \tag{3.4}$$

And a group of candidates of function φ have been given in [24], in our application, we use $\varphi(t) = t^2/(1+t^2)$ (Fig. 3.1(c)).

3.3 Edge-preserving Regularizer

Given the edge-preserving regularization function $\varphi(t) = \frac{t^2}{1+t^2}$, we can rewrite the inverse problem in a new form with the following procedure.

First, for simplicity, we represent the 2D object in a 1D vector form by stacking all the elements column by column, i.e.

$$\mathbf{f} = [f_{11}, ..., f_{m1}, f_{12}, ..., f_{m2}, ..., f_{1n}, ..., f_{mn}]^T \tag{3.5}$$

Then, in 3.2 let $\mathbf{g} = D_x\mathbf{f}$ and $\Phi^T(\mathbf{g})\Phi(\mathbf{g}) = \sum_{i,j} \varphi[\mathbf{g}_{i,j}]$, thus

$$[\Phi(\mathbf{g})]_i = \sqrt{\varphi(\mathbf{g}_i)} \tag{3.6}$$

We can rewrite the problem as:

$$\mathbf{f}_{ep} = \underset{f}{argmin} \|A\mathbf{f} - b\|_2^2 + \lambda_x^2 \|\Phi^T(D_x\mathbf{f})\Phi(D_x\mathbf{f})\|_2^2 + \lambda_y^2 \|\Phi^T(D_\mathbf{f})\Phi(D_y\mathbf{f})\|_2^2. \tag{3.7}$$

Then we can deem it as a non-linear least squares problem which has the following form:

$$\min \|F(\mathbf{f})\|_2^2 = \mathbf{e}^T(\mathbf{f})\mathbf{e}(\mathbf{f}). \tag{3.8}$$

Our edge-preserving regularizer can be fitted into this form by defining:

$$\mathbf{e}(\mathbf{f}) = [(y - A\mathbf{f}), \lambda_x\Phi(D_x\mathbf{f}), \lambda_y\Phi(D_y\mathbf{f})]^T. \tag{3.9}$$

We can use the Tikhonov reconstruction as an initial guess then find the local minimum of the non-linear least squares problem by Gauss-Newton method [28].

In the above approach, two regularization parameters λ_x and λ_y are used instead of one. It is because in most applications, the degradation effects brought by kernel

A can vary sharply in different directions. And by specifically tune the regularization on both x and y directions, we can reconstruct the object more precisely.

3.4 Example

In this section, a 2D example will be used to compare Tikhonov and edge-preserving regularization methods. Further analysis about the complexity and the error of estimate(Cramér-Rao bound) will also be provided.

The examples are from the simulation of the cross-well tomography problem we set up in Chapter 2. Two ideal objects, one '+' shaped and one rectangular shaped (Fig. 3.4(a),(b)), are given and white Gaussian noise (SNR=20) is assumed. The Tikhonov reconstructions are shown in Fig. 3.4(c)(d), which have lots of small fluctuations in 'flat' area and smoothing effects on edges.

Now we apply the edge-preserving regularization to the problem. Firstly, notice that the Tikhonov reconstructions reflect the different properties of kernel A in x and y directions. Specifically, edges in x direction tends to be more difficult to restore than those in y direction because the transmitters and receivers are placed along y direction. Therefore, different regularization parameters λ_x and λ_y should be used in the two dimensions respectively. And intuitively, λ_x should greater that λ_y to add more regularization in the x direction. Secondly, by using Gauss-Newton method to solve the non-linear least squares problem, a initial guess of the object should be decided. Here, it is natural to apply the Tikhonov result as the initial value. The remaining difficulty would be to find the best regularization parameters.

For 1D problem, it has been shown in Chapter 2 that a L-curve can be used to determine the best regularization parameter λ. For the 2D problem here, though we have 2 distinct regularization parameters λ_x and λ_y, analogously they can be determined by finding the corner of the L-surface. The process is shown in Fig. 3.4.

The x and y dimension of a L-surface are the two regularization terms $\lambda_x \varphi(D_x \hat{f}_{ep}))$ and $\lambda_y \varphi(D_y(\hat{f}_{ep}))$ respectively. The z dimension is the residue $\|y - A\hat{f}_{ep}\|_2^2$. The best set of λ_x and λ_y would correspond to the corner of the L-surface, which has the largest curvature. In Fig. 3.4(c)(d), the curvature of the L-surface are computed, thus the optimum regularization parameters are determined by locating the largest curvature [22]. As what we expected, λ_x is greater than λ_y.

Then we can get the edge-preserving regularization result (Fig. 3.4(e)(f)). Compared with the Tikhonov result, it filters out the small high-frequency artifacts and at the same time preserves the edges well. As stated before, we have placed transmitters and receivers along y direction, thus we can get better resolution along y dimension which brings the asymmetry of the reconstructions between x and y directions.

However, the edge-preserving regularizer has one major backdrawing. By applying non-quadratic regularization functions, it turns into a non-linear least squares problem which significantly increases the complexity of the problem. Since the algorithm is pixel-based, and most applications have huge amount of pixels, in each iteration of the Gauss-Newton method all the pixel values need to be evaluated, the whole process will be very slow. Further discussions will be provided in Chapter 5 and a B-Spline basis solution will be given there.

Another interesting analysis is the Cramér-Rao Bound(CRB) evaluation. If the reconstruction is considered to be an estimate of the object, the CRB analysis in detection and estimation theory can be applied. Given the cost function as the estimation function $G(f)$, we can get the CRB for both Tikhonov and edge-preserving regularizer. (Appendix A). For our example, the CRB of Tikhonov regularizer is shown in Fig. 3.4(a)(b), the CRB of edge-preserving regularizer is shown in Fig. 3.4(c)(d). The CRB represents the upper bound on the variance of the estimation error for different cross-well regions(pixels). Comparing the results, Tikhonov regularizer put relatively equal efforts to the estimation of edges and other 'flat' areas.

Edge-preserving regularizer preserves the edges better, which means it locates the edges more accurately; thus as a compensation, the variance of the estimation error in such areas increases.

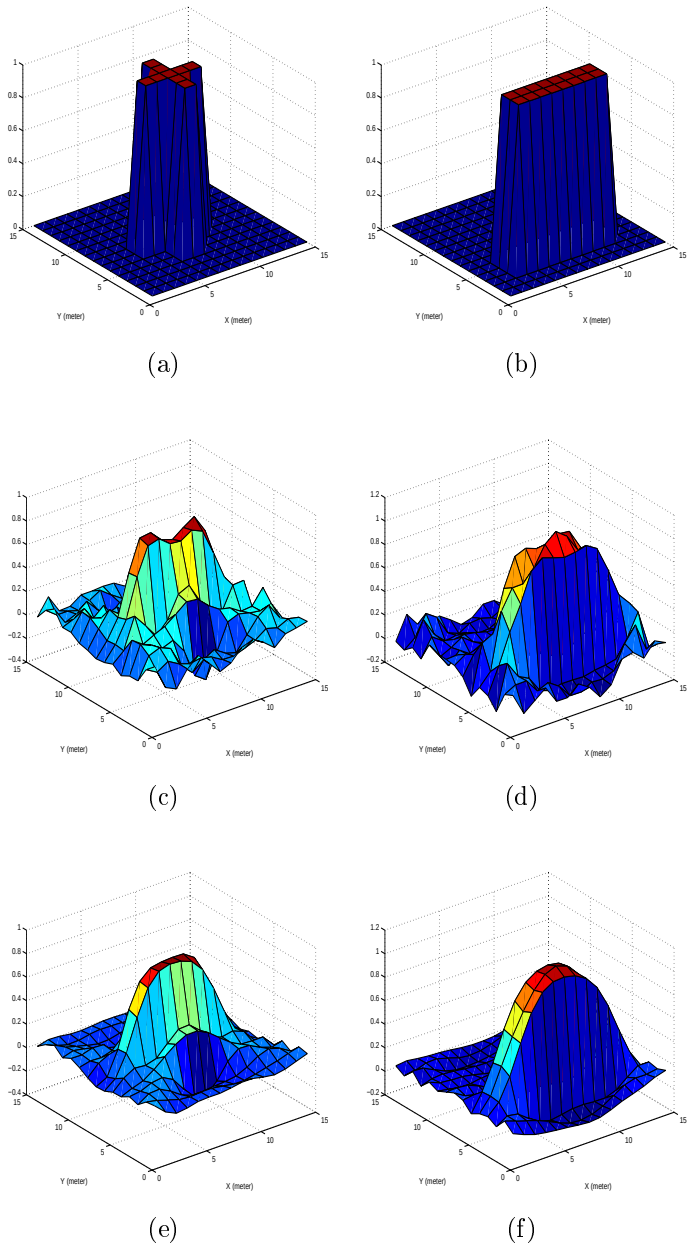


Figure 3.2: Example: (a),(b) objects; (c),(d) Tikhonov reconstructions; (e),(f) Edge-preserving reconstructions. Transmitters and receivers are placed along y dimension ($x = 0$ and $x = 15$).

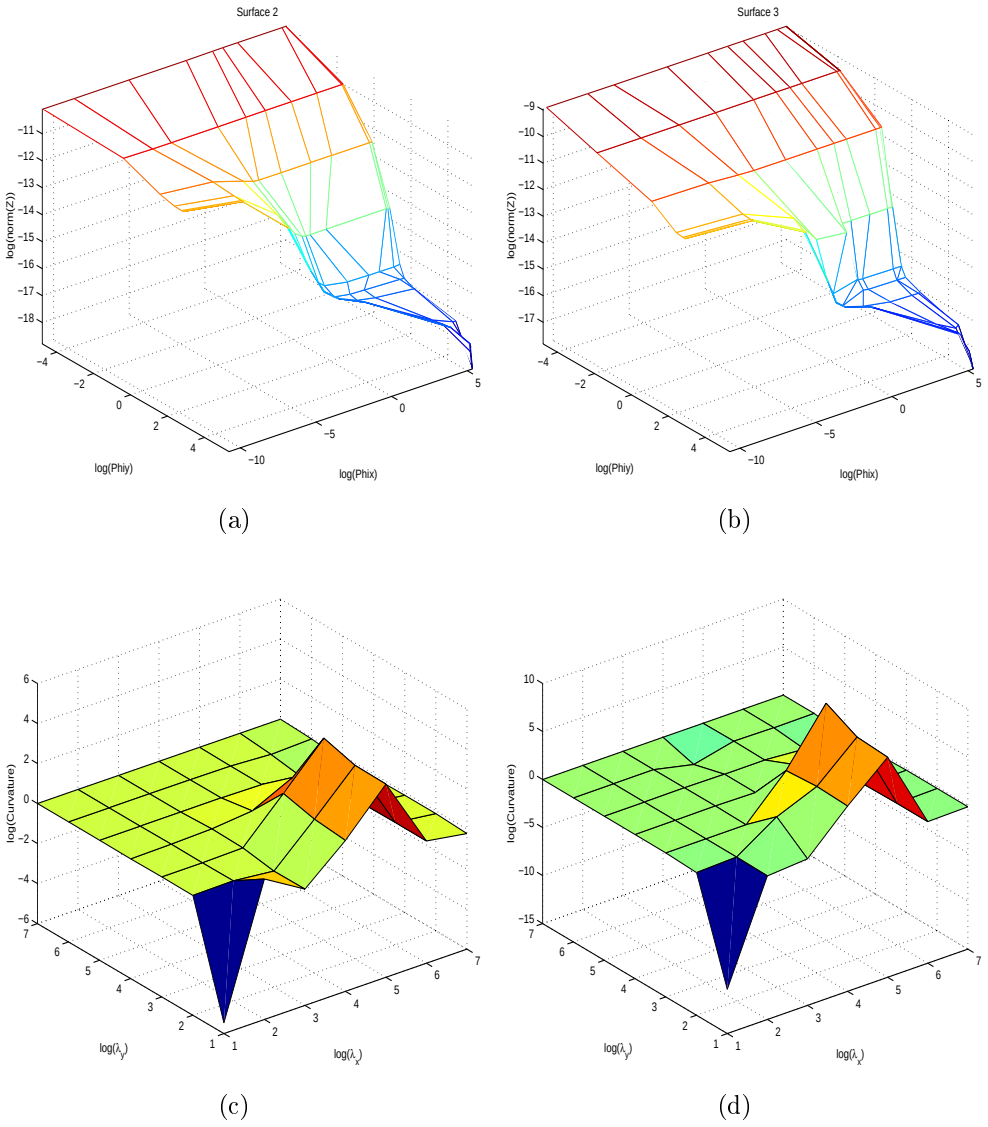


Figure 3.3: Example: (a),(b) L-surfaces; (c),(d) Curvature of the L-surfaces.

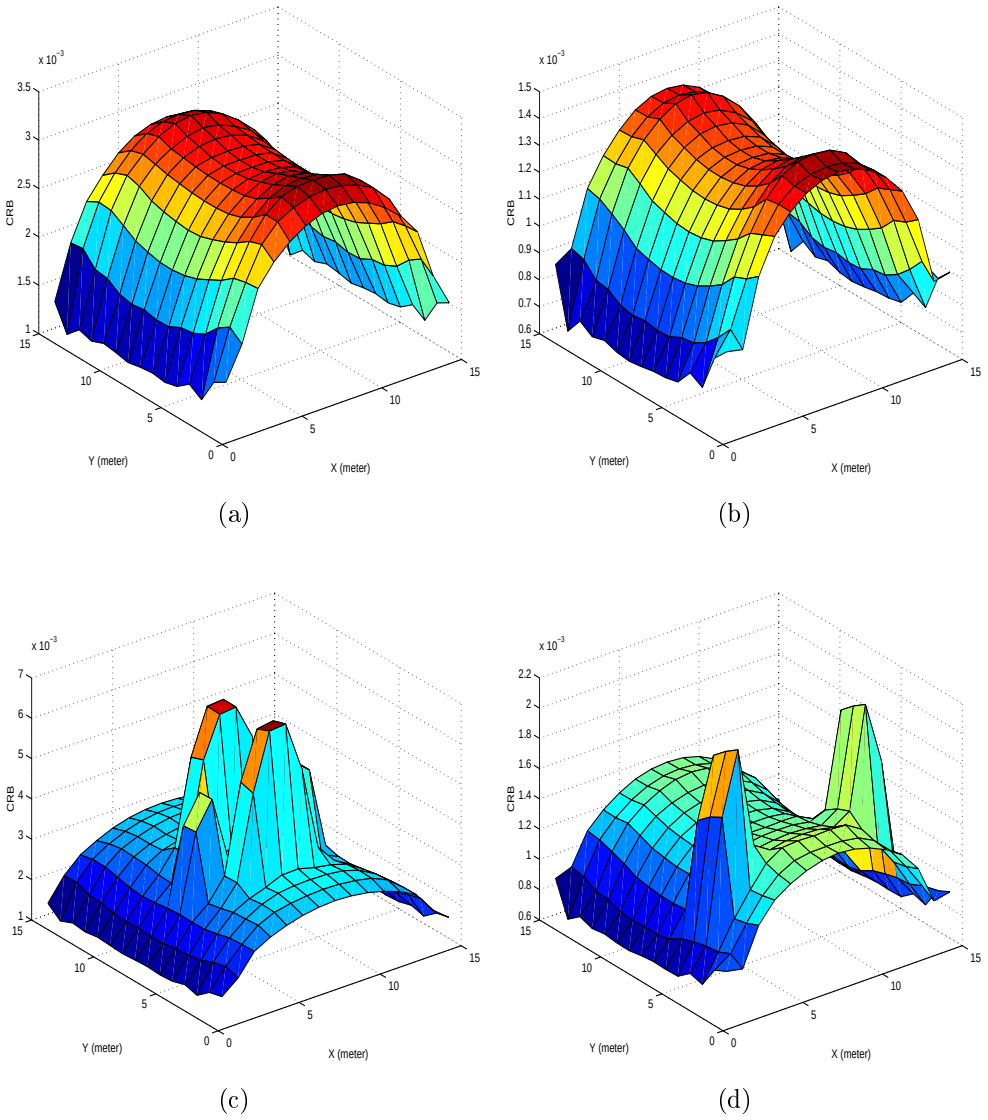


Figure 3.4: Example: (a),(b) CRB of Tikhonov regularizer; (c),(d) CRB of edge-preserving regularizer.

Chapter 4

Adaptive B-Spline Reconstruction

In Chapter 3, the edge-preserving regularization method was discussed and one major deficiency brought by solving the corresponding non-linear least squares problem is the significant complexity in applications. To solve the problem, a B-Spline based adaptive algorithm is developed. The cross-well 2D example will be revisited to show the improvement at the end of this chapter.

4.1 Motivation

For reconstruction methods previous introduced, both Tikhonov and edge-preserving regularization, are based on pixel-by-pixel estimate. In practical cases, especially in 2D or even 3D problems, the total number of pixels is always very large. Thus the reconstruction problems will need significant time and effort to obtain a solution. The situation will be even worse for edge-preserving regularizer, because the Gauss-Newton method used to solve the corresponding non-linear least squares problem requires evaluation of estimations of all the pixels in each iteration.

To solve the problem, a natural thought is that if we know the shape of the object, a proper shape-base template would be efficient to precisely describe the

object and significantly reduce complexities [29]. However, in practice, we don't have prior knowledge of the object shape and size, so a compromised way would to find a set of proper basis to represent the object. The basis should have the properties that flat regions can be represented by one or few coarse scale functions while more fine scale functions can be used to represent details of regions of interests , e.g. edges.

Based on the discussions in chapter 2, B-Spline basis would be a good candidate. First, it is natural to represent an unknown object by a set of k-order piecewise polynomial functions. Second, the multiscale property of B-Splines can be utilized to efficiently represent the object hence decrease the complexity of the problem. Third, efficient algorithms of evaluating and manipulating B-Splines exist which will not bring much extra efforts in dealing with the basis itself. Fourth, the smoothness and scale of B-Splines can be easily manipulated by knots distribution, which means an adaptive algorithm may exist for choosing a suitable set of B-Splines.

4.2 Inverse Model using B-Splines Basis

Using B-Spline basis instead of pixels to represent object f, the inverse problem can be remodeled as:

$$y = Af + n = ANa + n = Ba + n, \tag{4.1}$$

where N is the B-Splines basis and a is a vector of expansion coefficients for the B-Splines. Thus the problem can be replaced by a new inverse model whose degradation kernel will be $B = AN$, and the new object to be reconstructed is a. If proper B-Splines basis are chosen to represent f, i.e. the representation is efficient, the dimension would be significantly decreased compared with that of f. Furthermore, the edge-preserving regularization method can be used to solve the new problem.

In the 2D example, tensor-product B-Splines can be utilized because it is a simple extension of 1D B-Splines, and knots in x and y direction can be manipulated separately. The drawback is that the rectangular patches may lose some flexibility. Two examples of tensor-product B-Splines and corresponding knot distribution have been shown in Fig. 2.7.

4.3 Optimum Knots Distribution

Many sets of B-Spline basis can be chosen to represent a certain object, the best would be the one that yields the most efficient representation. In another word, the B-Spline basis needs to be distributed in the way that flat regions are covered by few coarse B-splines while a larger density of B-Splines cluster at rough regions. By utilizing such a set of B-Splines, the complexity of the inverse process can be decreased.

From the discussions of B-Splines in Chapter 2, the number of knots corresponds to the number of B-Splines being used, thus the number of unknowns in the revised model (4.1). The knot locations directly affect the support(scale) and distribution of B-Splines. Therefore, the problem is to find an optimum knots distribution of the B-Splines. It has two aspects: 1. How many B-Splines are sufficient to represent the object? 2. Given a fixed number of knots (B-Splines), what is the optimum distribution of those knots?

The second problem is straightforward that the knots distribution should reflect the shape/smoothness of the object. Specifically, large density of knots should cluster at rough regions, i.e. regions with large curvature. And only few knots would be enough to represent the remaining flat regions.

The first problem can be interpreted as the following: Given a certain tolerance of estimation error, what is the fewest number of knots/B-Splines can be used to make

the reconstruction? Roughly the process involves inserting new knots where the estimate is poor and deleting redundant knots where the estimate is over accurate than needed which will unnecessarily increase the complexity of the problem.

Based on the above discussions, an ideal solution would be able to find the optimum set of B-Splines, including the number of B-Splines and the placement of each B-Spline. It is difficult to accomplish because the searching effort will be significant, and even if we can find such solutions, the computational cost may be not worth the improvement gained compared with the pixel-based solutions. An alternate way would be to efficiently find a suitable set of B-Splines instead of an optimum one. The desired algorithm will be able to end up with a highly suitable knot distribution in the sense that the number of knots is much fewer than that of pixels, while the reconstruction does not degrade much.

Our processing approach can be represented in a tree structure [46] [47]. As shown in 4.1, for simplicity we use binary tree to show the cases of 1D problem, each level of the tree corresponds to different number of knots being used. Then, we can think of using an adaptive algorithm to automatically find the suitable knot distribution. First, the algorithm should be able to iterate between coarser and finer level of reconstruction to determine the most suitable number of knots to be used. If one knot is not good enough to cover a certain region, the algorithm should be able to split it into 2 knots on the next finer level and vice versa. Second, at the same time, the algorithm should be able to intelligently alter the locations of knots to make better reconstruction in a certain resolution level. Therefore, the placement of each knot is not fixed but varies during the iteration of the algorithm. The overall algorithm is to find a set of knots with much fewer knots than pixels based on which the estimate is within a certain error.

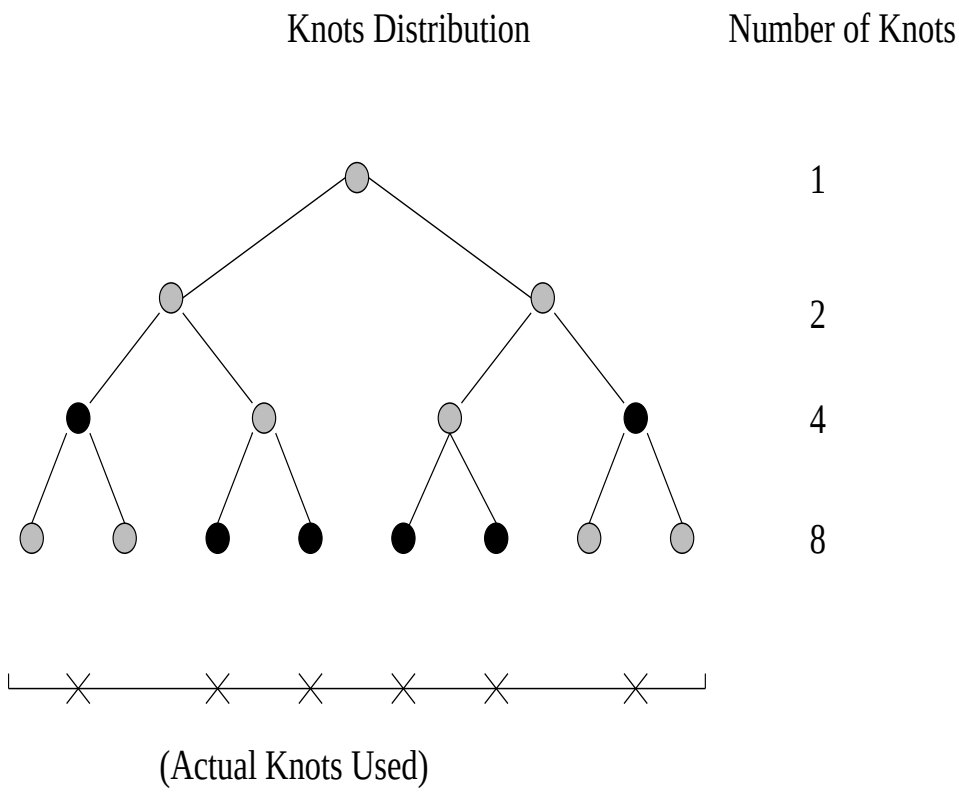


Figure 4.1: Illustration of Tree Structure for Knots Searching (1-d case)

4.4 Knots Refining

Given a coarser level of reconstruction, one may want to add more details in some interesting areas, e.g. edges. A simple way is to double the knots/B-Splines used in reconstruction. The process is to insert new knots in between all existing knots, thus the resolution of all areas are doubled. The drawback is that the total complexity is at least doubled. For example, in the 1D problem showed in Fig. 2.8, notice that many knots are inserted in uninteresting areas, e.g. flat areas where existing knots are sufficient to have an acceptable estimation Fig. 2.8(c). Therefore, a more adaptive way should be found to insert new knots. Specifically, knots should only be inserted as necessary, i.e. the algorithm should automatically know where the areas

of interests– potential edges are and only inserts new knots in such areas.

In 2.8(a), notice that though the coarser level(with only 5 knots) reconstruction is rough, we can still get some information of the object shape/smoothness. In another word, the coarser level reconstruction gives a hint to the shape/smoothness of the finer level reconstruction. Since we are interested in edges which have large roughness or discontinuity, we can use this hint as a guide of inserting new knots. The process is called *knot refining* because it is not a simple knot insertion procedure[41][43]. It also involves curvature computing (Chapter 3.3) based on current estimate. The knots refining process is described as the following:

Algorithm 1: *Knot Refining*

Input: Knot sequence $P_i, i = 1, 2, ..., m$

Corresponding reconstruction $\hat{f}(P)$

Threshold K_0

Output: New knot sequence $Q_j, j = 1, 2, ..., n$

Begin:

$K = CurvatureComputing(P, f(P))$;

for $i = 1$ to m,

if $(K_i > K_0)$ then

$KnotInsertion(K_i)$;

(K_i is the curvature at knot P_i)

end if

end for;

End

Specifically, in the above algorithm, the knot insertion function splits one knots into two equally space knots to cover the same region.

In the above algorithm, an important assumption is that coarser level reconstructions will provide correct guide to get finer resolutions. But the reconstruction may not be good enough to show the rough structure of the object if only too few B-Splines are used. It can be avoided, in practice, by always starting from a relatively fine level guess, and the knot deletion algorithm proposed in next section will guarantee not too many knots will be deleted.

Another important issue is to choose a proper threshold K_0 for inserting new knots. If K_0 is too large, i.e. only few knots will be inserted to the areas with very large curvature, the resulted reconstruction will be only focusing on several large abruptly changing regions. Another drawback is that the refining process may be misled by some false edges in coarser level. On the other hand, if K_0 is too small, too many knots will be inserted while most of them will bring no good to the reconstruction and have to be deleted later. How to automatically find the proper K_0 will need further investigation. In this thesis, the threshold K_0 is determined by experimenting different values and comparing the results.

4.5 Knot Pruning

After inserting a set of new knots, a refined reconstruction can be obtained. The new reconstruction should have different shape/smoothness and different curvature structures than the previous coarser one. Since the knots inserted were based on the coarser reconstruction and it was only a guess of the underlying refined details, some of the knots may be misplaced, i.e. knots were actually inserted in a flat region rather than the expected rough region. On the other hand, the regions considered to be flat may appear to have large curvatures. The *Knot Pruning* algorithm is used to prune the new set of knots, find and delete those improper and redundant ones.

4.5.1 Weight of Knots

The most important problem here is how to prune the knot sequence and find those redundant ones. Therefore, a measurement of each knot's importance should be provided to determine which ones are less important to the reconstruction and can be deleted.

The most important knot should be the one that contribute the most to the current reconstruction, i.e. by deleting it, the reconstruction will have the greatest change. The underlying thought is that only few knots are needed to well approximate a flat region, while a cluster of knots may needed in rough regions with lots of details to make an equally good approximation. So those knots appear to be less important to the reconstruction means there are too many of them than needed compared with the information in these areas, i.e. redundant. A natural measurement of the change in reconstruction is the underlying cost function for the problem. For edge-preserving regularizer, the cost function is:

$$\text{cost } = \|y - Ba\|_2^2 + \lambda_x^2 \varphi[D_x(Ba)] + \lambda_y^2 \varphi[D_y(Ba)]. \tag{4.2}$$

The edge-preserving regularization is to find the solution $\hat{a}$ minimizes the cost function. And apparently, the more knots used(or the larger dimension the a has), the better resolution the reconstruction should have. Thus the cost function increases each time a knot is deleted from current knot sequence. And the knots that cause the larger increase of the cost function are of the more importance, i.e. have larger weight in the reconstruction.

An illustration is shown in Fig. 4.2. Here a simple B-Spline approximation problem is used to get better understanding. The cost of deleting each knot is measured by Mean Squared Error(MSE). For the cross-well reconstruction problem, the situation is quite analogous, except that the cost function is measured by the edge-preserving

regularizer.

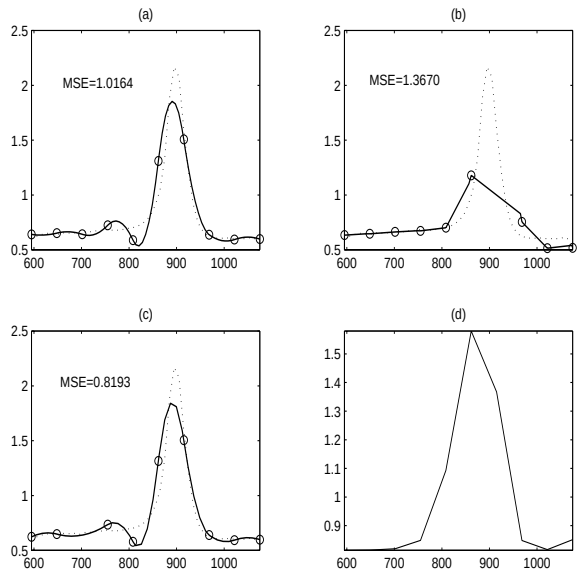


Figure 4.2: Illustration of weight of knot. (a)10 equally-spaced knots to approximate the data(dotted line); (b)Approximation of deleting the 7th knot; (c)Approximation of deleting the 3rd knot; (d)Cost(Weight) of each knot.

4.5.2 Knots Deletion

After measuring the weight of each knot, we can start to delete those redundant ones. First, sort the knots in the descendant order of weight. Then delete knots one by one in the order of their weights until the resulting reconstruction is coarse enough but not too much worse. The criteria is set to be that the cost function value after deletion does not exceed the one at the beginning of each iteration(i.e. before the knot insertion). The idea is that the whole knot insertion and deletion process is actually to improve the knot distribution. Thus, for each iteration of such process, the resulted cost should not exceed the initial one, otherwise, the knot distribution is not improved but worsened. The specific mechanics of deleting a single knot has

been discussed in many papers [38] [39].

Consider a simple 1D example in Fig. 4.3, notice the knot distribution here is not the same as of Fig. 4.2. More knots are assigned to the bump area to simulate the result of inserting cluster of knots in large curvature regions. In this case, the weight of each knot are shown in Fig. 4.3(b). If we delete knots in the order based on this weight measurement, we may delete the 12th knot first then the 13th knot because they have the smallest cost. However, as shown in Fig. 4.3(c), after deleting the 12th knot, the actual weights of some knots change, specifically for those adjacent knots. If we still try to delete the 13th knot obeying the deletion order given by the previous weight measurement, the resulted approximation error may be too large so that the deletion process will terminate. The result is far from satisfying, because actually we can delete more knots than we did without violating the tolerance.

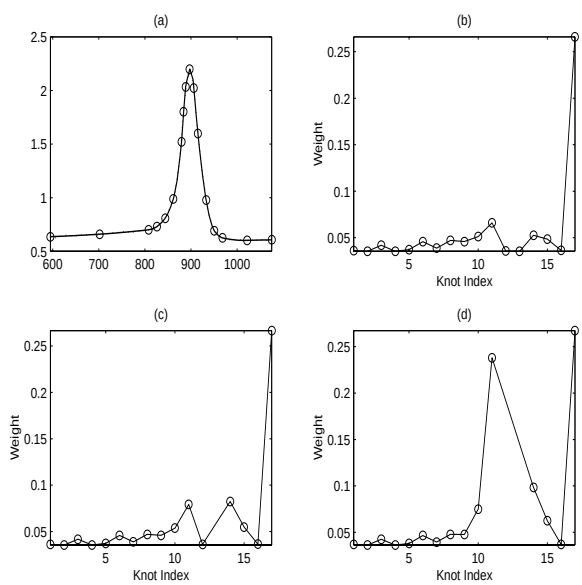


Figure 4.3: Illustration of vicinity effects of weight measurement. (a)Initial knots distribution and approximation curve; (b)Weight measurement; (c)Weight of each knot after deleting the 12th knot; (d)Weight of each knot after deleting the 12th and 13th knot.

The problem arises from the fact that B-Splines have supports leaking into adjacent knot intervals, thus we can not ignore their effects on vicinity regions. In another word, deleting a knot does have effects on its vicinity regions, and change the actual weight of adjacent knots. To solve the problem, one solution would be to recompute the weight of each knot in each step after a knot has been deleted. This will significantly increase the complexity of the algorithm. Considering the local support of B-Splines, deleting one knot will not bring much effect to remote regions. Thus, an alternative and much efficient way to avoid the vicinity effects may be continue using the weights we get before the deletion but simply not to delete adjacent knots to the one who has been just deleted. Therefore, we will only need to measure the weight of each knot once and choose and delete several knots at a time. Since we have avoided deleting adjacent knots, the measured weight of each knot would be a good estimate of the actual weight during the knot deletion procedure.

The knot pruning algorithm is described as the following:

Algorithm 2: *Knot Pruning*

Input: Knot sequence $Q_i, i = 1, 2, ..., m$

Corresponding reconstruction $f(Q)$

Threshold C_0

Output: New knot sequence $P_j^*, j = 1, 2, ..., n$

Begin:

$W = Weight(Q);$ **Compute weights(costs) **

$[W, index] = Sort(W);$ **Sort in ascending order**

$i = 0; cost = 0;$

$P^* = Q;$

While $cost \leq C_0$

{

$i = i + 1$

if $(Neighbour(Q(index(i)))) \in P^*$ then {

$\{P^*\} = \{P^*\} - Q(index(i))$ *Delete one knot at a time**

$cost = Weight(P^*)$;

}

}

End

4.6 Overall Algorithm

The knot refining and pruning algorithm can be integrated together to form the overall algorithm to find a suitable knot distribution. The flow chart of the overall algorithm is shown below in Fig. 4.4:

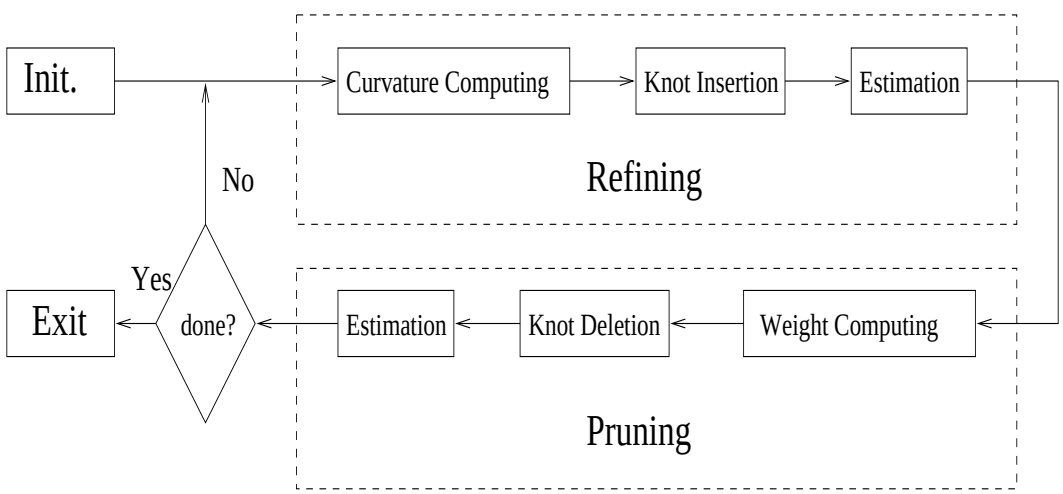


Figure 4.4: Flow chart of knot optimization algorithm.

The knot refining part is to efficiently insert new knots according to the curvature information of current coarse reconstruction. The knot pruning part is to find

and delete the redundant knots, either previously exist or newly inserted, based on a finer reconstruction. Therefore, the overall effect is to insert knots firstly then make correction to it by pruning. In each iteration, the resulted knot distribution is guaranteed to be better than the previous one in the sense of having less cost. As the algorithm converges, a more suitable knot distribution can be found, i.e. with much fewer number of knots than that of pixels and the estimation is within certain error. The algorithm will terminate while no more improvement has been obtained, i.e. the current cost function almost equals the last result.

Fig.4.5 shows the result of using our algorithm for the example in previous Chapters using the Titanium data. The resulting knot distribution makes the approximation with $MSE < 0.1$. Compared with Fig. 2.8, our algorithm gets a smaller approximation error while using fewer but well-distributed knots. In Fig. 2.8(d), 12 selected knots were used. Our algorithm only ends up with 10 knots and the approximate error is less in the means of MSE.

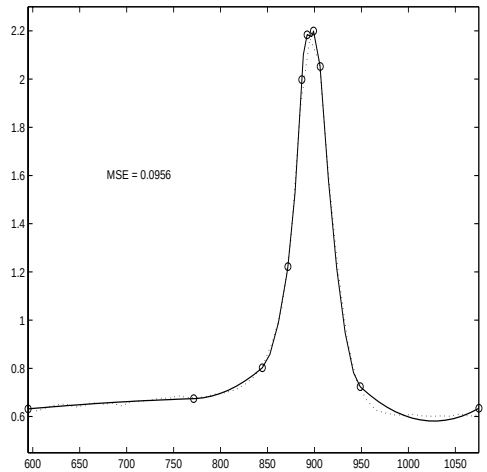


Figure 4.5: Using knots refining and pruning algorithm on the Titanium data fitting problem.

4.7 Examples

The 2D cross-well example used in Chapter 3 is revisited here. The previous Tikhonv reconstruction and edge-preserving recontruction results are shown in Fig. 4.6 as comparisons. The resulting reconstructions and corresponding knot distributions are shown in Fig. 4.7. The algorithm converges fast (Fig. 4.8). The knot distributions are what we expected, i.e. cluster around edges and only few on flat areas. The total number of B-Spline basis used are less than 50, compared with the total pixel number $21 \times 21 = 441$, approximately a 8 times or more reduction in complexity is achieved. Consider that the objects are large and have many details, especially for the '+' shaped one, compared with the total area, the results are satisfying. For applications with larger dimensions and relatively small objects, the decrease in complexity will be significant. Although the MSE almost doubled , we can still state by visionary obervations that the B-Spline based reconstructions do not degrade much compared with the pixel based ones.

Another example is the 3-layer structure we may be of interest in practice, the result is shown in Fig. 4.8. The algorithm works well, converges fast and only ends up with about 30 knots.

The only problem is that the saw shaped parts of the object are not well reconstructed. Only one knot is assigned in the vertical direction. The problem arises from the using of tensor-product B-Splines, the approximation patches should always be rectangular. For the object we have, in the vertical direction, the upper side of it has one significant edge, while the lower side of it has lots of details(saws). Since we were using rectangular patches, a compromise should be made for the knots distribution here. Using few knots will decrease the complexity but sacrifice the details at lower part. Using more knots are a waste for the upper side but are needed by the details at lower part. It's a difficulty for our approach. Better solutions may be using triangular

patches instead of tensor-product B-Splines[49] [50]. It will be considered in future work.

4.8 Monte-Carlo Simulation

To better understand the performance of the algorithm, we will analyze Monte-Carlo simulations in this section. The process is to repeatedly simulate the test by applying different noises at the same SNR level. With large number of such tests, we can get the statistics of how the algorithm performs under such SNR level.

In this section, we use a 3-layer example to simulation the actual situation of the cross-well problem stated before. The result in Fig. 4.9(b) is obtained by applying the B-Splines edge-preserving reconstruction method. The algorithm converges fast and the overall time of achieving the estimate is about half of the time of the pixel-based reconstruction.

In Fig. 4.9(a), the average pixel-based reconstruction of the 3-layer problem is shown. The average B-Spline reconstruction is shown in Fig. 4.9(b). The B-Spline result is quite good and does not degrade much than the pixel-based one, and at the same time, by utilizing much fewer basis, the complexity has been significant reduced. To understand the knot distributions of the Monte-Carlo simulations, a 2D histogram for all the knots ended in the 200 tests is shown in Fig. 4.9(c). It shows the knot distributions have certain relationship with the object structure as we expected. Specifically, there are higher possibilities for knots being place around edges of the object, and the density of knots are also larger there. To make it clear, a scaled version of knot distribution has been shown in Fig. 4.9(d). It only shows those knots whose frequencies of being chosen are larger than 0.1 during all the 200 tests. In another word, the knots shown here are essential and most efficient to represent the object.

To further determine the average knot distribution, since we used tensor-product B-Splines, separate the 2D histogram into two 1D histograms will be helpful. The results are shown in fig. Fig. 4.10.

In Fig.4.10(a)(b), the 1D histograms of knot distribution along x and y directions are shown. Since we use tensor-product B-Splines, the separation of knots in different dimensions is applicable. By choosing certain threshold, we can determine which knots are to be used. Here, a threshold of 0.13 has been applied to both x and y dimensions, and the resulted reconstruction and knot distribution are shown in 4.10(c)(d). Since we use tensor-product B-Splines, the overall knot distribution is abided by the rectangular grid. The variance of estimates over the 200 reconstructions are shown in Fig. 4.11.

The Monte-Carlo simulation has shown that the B-Spline algorithm is efficient that it ends up with much fewer knots than that of pixels and the reconstruction does not degrade much. Generally, the algorithm can lead to a 8-10 times decreasing in the complexity of the problem.

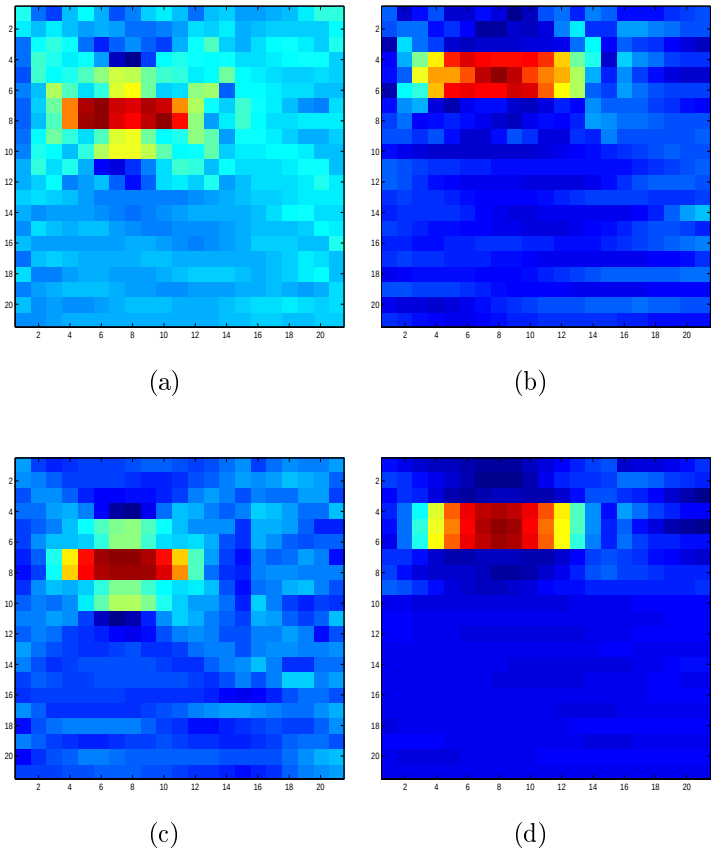


Figure 4.6: Example: Reconstruction obtained by Tikhonov regularization and edge-preserving regularization method. (a)'+' shaped object, Tikhonov regularization (MSE = 5.7393) (b)rectangular object, Tikhonov regularization (MSE = 5.2943) (c)'+' shaped object, edge-preserving regularization (MSE = 3.3930) (d)rectangular object, edge-preserving regularization (MSE = 2.9372)

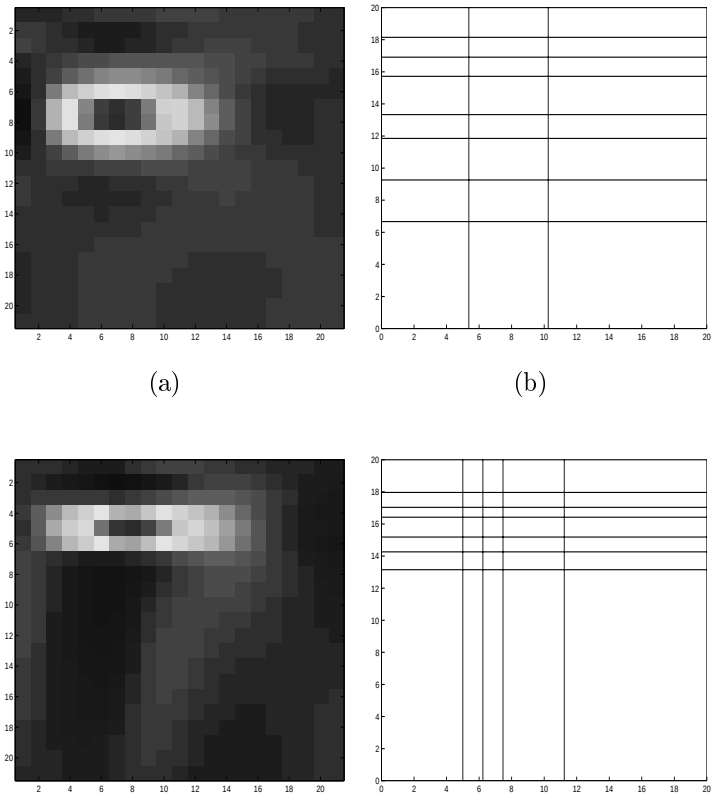


Figure 4.7: Example: Reconstruction and knots distribution using B-Spline based edge-preserving regularizer. (a),(b) '+' shaped object; (MSE = 7.8461) (c),(d) rectangular object. (MSE = 5.7867)

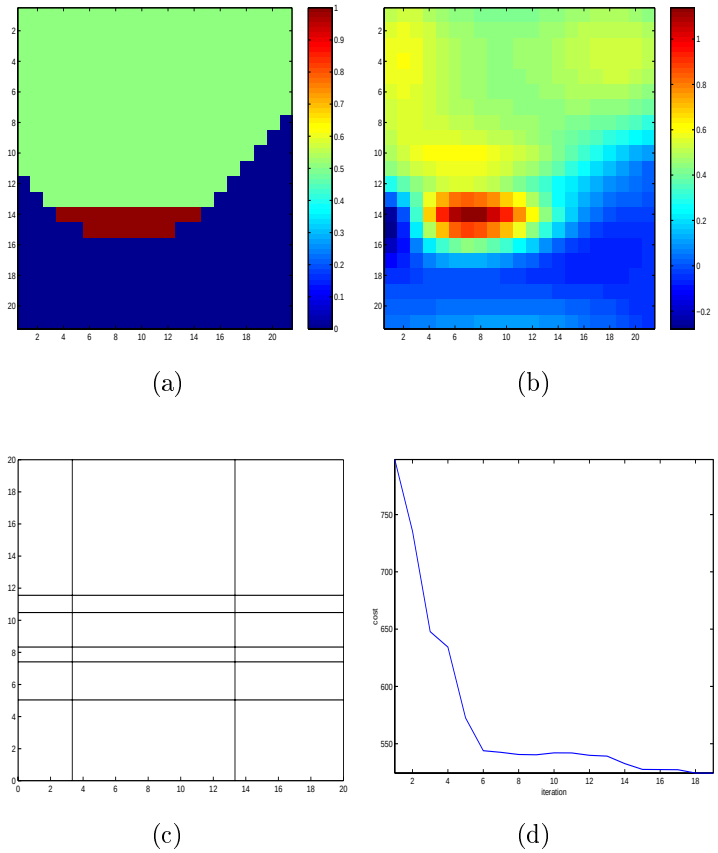


Figure 4.8: Example: B-Spline reconstruction for 3-layer problem (a)Object; (b)Reconstruction; (c)Knots distribution; (d)Convergence of cost (Each iteration consists of knot inserting and knot pruning procedure.)

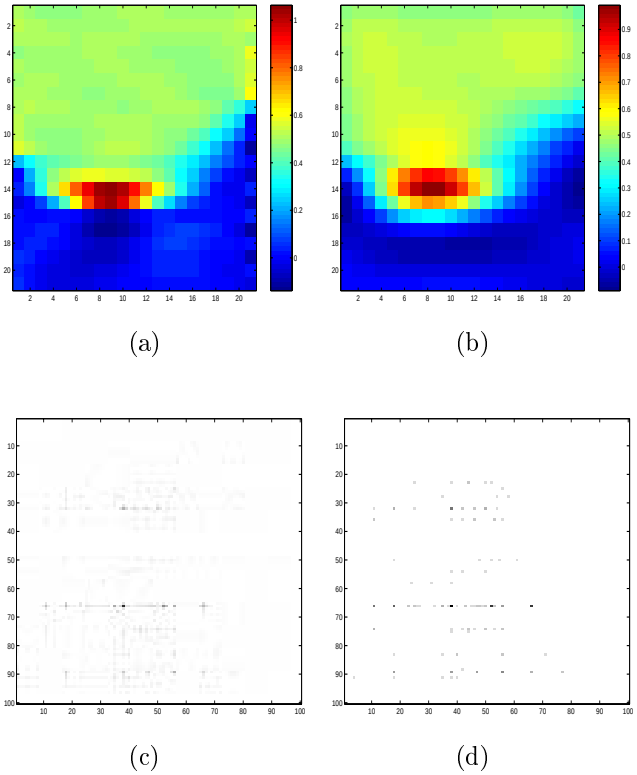


Figure 4.9: Monte-Carlo simulation for the 3-layer problem. (SNR = 20dB) (a)Average pixel-based reconstruction (MSE = 2.0874); (b)Average B-Spline reconstruction (MSE = 5.8922); (c)Total knots distribution; (d)Scaled knots distribution.

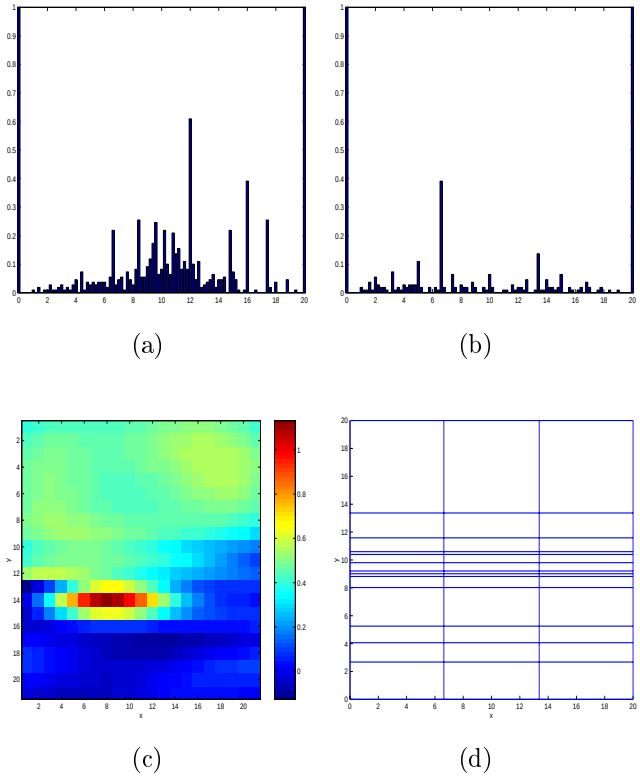
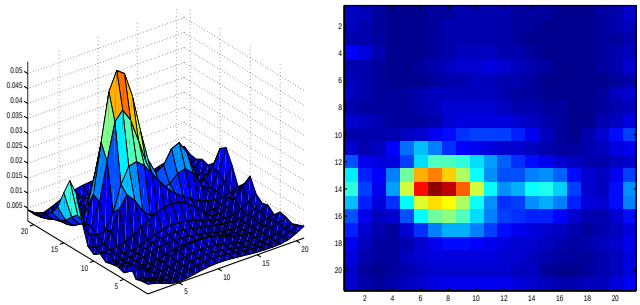


Figure 4.10: Knot selection by using separate 1D histograms. (a)Histogram of knot distribution along x direction. (b)Histogram of knot distribution along y direction. (c)Reconstruction based on histogram analysis. (d)Corresponding knot distribution.



(a) (b)

Figure 4.11: Variance of the Monte-Carlo simulation (2 different viewpoints)

Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis developed an adaptive B-Spline method for solving low order image reconstruction problems. Also an edge-preserving regularizer was introduced and integrated to make fine reconstructions. The edge-preserving regularizer has been shown to be very efficient in countering the smoothing effects of Tikhonov regularization and preserving the edges. But for multi-dimensional problems, the edge-preserving regularization leads to significant large complexity. To overcome this deficiency, B-Spline basis was chosen to make simple but efficient representation of the objects which decreased the dimension of the problem. Furthermore, an adaptive knot refining and knot pruning method was developed to automatically find a suitable knot distribution which is critical in solving the problem. Some examples were shown to illustrate the adaptive algorithm and the results were satisfying.

In Chapter 2, some background knowledge was introduced. First, the linear inverse problem model was given. The difficulty in solving the inverse problems lies in the ill-posed nature of such problems. As a result, the least squares solution always tends to have large high-frequency artifacts, and small observation errors would bring

large errors in solutions. After analyzing the problem from the SVD(Singular Value Decomposition) point of view, the regularization methods were introduced to filter out the abruptness brought by small singular values σ_i. The smoothing effects of Tikhonov regularization were investigated and led to the thought of edge-preserving regularization was introduced. Second, we introduced the basic knowledge of B-Splines. After giving the definition, some interesting properties, such as local support, minimal support and linear independent of B-Splines were introduced. A further study in the effects of knot distribution to the smoothness of the approximation showed that we can change the approximation curve/surface by simply manipulating the knots. An example was shown that proper placement of knots can lead to efficient representation of the object. Third, the geometries of B-Splines were introduced to compute the B-Spline surface curvatures, which is the guide of our manipulation of knots distribution. Tensor-product B-Splines was used in 2D case, and Gaussian and absolute curvatures were introduced.

In Chapter 3, a thorough discussion of edge-preserving regularization was given. The deficiency of Tikhonov regularization is that it blurs the edges while filtering out high-frequency artifacts. A study on the regularization term showed that a non-quadratic regularization functions can be used to preserve edges. Some common properties of such regularization functions were given and a non-linear least squares solution was proposed to solve the problem. Then a simple 1D example showed the edge-preserving regularizer did work. While the CRB analysis showed an interesting fact that with better preserving the edges, the error of estimate of edge-preserving regularizer actually had larger variance. A major problem of edge-preserving regularizer was its huge complexity, since the Gauss-Newton method need to evaluate all pixels in each iteration.

In Chapter 4, a B-Spline based adaptive algorithm was developed to decrease the large complexity of edge-preserving regularizer. The motivation was from that

the object can be efficiently represented using B-Splines. Thus after remodeling the inverse problem we had, we found that the dimension of the problem can be significantly decreased if we could find a suitable set of B-Splines. The process of finding such basis was actually to find a good knot distribution. Based on the discussions in Chapter 3, a knot refining and pruning algorithm was developed to adaptively finding a suboptimal solution. The refining part was to efficiently insert new knots according to curvature information. The pruning part was to find and delete those redundant knots. A measurement of the weight of each knot was given to assist the pruning process. The whole algorithm works to find a suitable set of knots with the reconstruction is within certain error. The 2D cross-well problem was revisited by using the algorithm, and the result appeared to be satisfying. It was able to decrease the dimension of the problem to 1/8 compared with the pixel-based approach, and the iteration process converges in about 10 iterations .

5.2 Future Work

Overall, this thesis has shown that the adaptive B-Spline incorporated with edge-preserving regularizer is useful and efficient in solving image reconstruction problems. The algorithm can adaptive locate the edges of the object as well as significantly decrease the complexity. However, there are still many aspects need to be explored and refined.

- The edge-preserving regularizer has been shown to be able to reconstruct the edges well. However, the CRB analysis in Chapter 3 also showed that the estimation variance at the edges increased compared with Tikhonov solutions. It is an interesting phenomenon. It seems the edge-preserving regularizer can locate the edges more accurately, but at the same time the estimation variance has been sacrificed. Further analysis in the framework of detection and estimation

theory may lead to some in-depth understanding of the problem.

- The knot refining process used curvature information as a guide, thus it depended much on the selection of threshold K_0. We were just using heuristic values in the examples. One could use a set of different thresholds to provide multiple results, each focusing on curvatures of different ranges. It'd better if there is a way to adaptively set the threshold.

- The knot pruning process is currently a greedy search process. An estimation of the weight of each knot was made first, then we tried to delete them one by one until the error of estimate exceeded a certain value. Though we have considered the vicinity effects and will not delete adjacent knots in the same iteration, the weight estimation may not be good enough to guide the deletion process. Actually, we also tried to organize the knots in quad-tree structure and use Branch-and-Bound method to search it [51]. Unfortunately, it was too difficult to determine the bounding function and we did not get good results. Further study of the pruning method may be helpful.

- In 2D case, we used tensor-product B-Splines. The advantage is that it is easily to be manipulated and the knots on different dimensions are separable. The major disadvantage is that we are bound to use rectangular patches to make the estimation thus lose some flexibility in manipulating the knots. A triangular patch [49] [50] may be interesting, though the computations are more difficult, the flexibility it brings will be worthy of trying.

Bibliography

[1] M. Bertero, "Linear Inverse and Ill-posed Problems", *Advances in Electronics and Electron Physics*, Vol. 75. Academic Press, 1989.

[2] A. S. Willsky, P. W. Fieguth and W. C. Karl, "Multiresolution Stochastic Imaging of Satellite Oceanographic Altimeter Data", *IEEE International Conference on Image Processing*, Vol. 2, 1994.

[3] E. L. Miller and A. S. Willsky, "A Multiscale Approach to Sensor Fusion and The Solution to Linear Inverse Problems", *Applied and Computational Harmonic Analysis*, Vol. 2, pp. 127-147, 1995

[4] J. Zhang, G. Wang and G. W. Pan, "Solution of Inverse Problems in Image Processing by Wavelet Expansion", *IEEE Trans. Image Processing*, Vol. 4, No. 5, pp. 579-593, 1995.

[5] R. F. Harrington, *Field Computations by Moment Methods*, Macmillan, 1968.

[6] K. A. Dines and R. J. Lytle, "Computerized Geophysical Tomography", *Proc. IEEE*, Vol. 67, pp. 471-480, 1979.

[7] E. L. Miller and A. S. Willsky, "A Multiscale, Statistically Based Inversion Scheme for Linearized Inverse Scattering Problems", *IEEE Trans. Geoscience and Remote Sensing*, Vol. 34, No. 2, 1996.

[8] K. I. Hopcraft and P. R. Smith, *An Introduction to Electromagnetic Inverse Scattering*, Kluwer Academic, 1992.

[9] R. G. Newton, *Scattering Theory of Waves and Particles*, Springer-Verlag, 1982.

[10] J. Lekner, *Theory of Reflection of Electromagnetic and Partical Waves*, Kluwer Academic, 1987.

[11] S. Coen, K. K. Mei, and D. J. Angelakos, "Inverse scattering technique applied to remote sensing of layered media", *IEEE Trans. Antennas Propagation*, Vol. AP-29, No. 2, pp. 298-306, 1981.

[12] D. K. Ghodgonkar, O. P. Gandhi, and M. J. Hagmann, "Estimation of Complex Permittivities of Three-dimensional Inhomogeneous Biological Bodies", *IEEE Trans. Microwave Theory Tech*, Vol. MTT-31, pp. 442-446, 1983.

[13] E. L. Miller, L. Nicolaides, and A. Mandelis, "Nonlinear Inverse Scattering Methods for Thermal Wave Slice Tomography: A Wavelet Domain Approach", *J. Optical Society of America (A)*, Vol. 15, No. 6, pp. 1545-1556, 1998.

[14] G. F. Roach, *Green's Functions*, Cambridge University Press, 2nd Edition, 1982.

[15] E. Wolf, "Three-dimensional Structure Determination of Semi-transparent Objects from Holographic Data", *Optical Communications*, Vol. 1, pp. 153-156, 1969.

[16] W. H. Carter, and P. C. Ho, "Reconstruction of Inhomogeneous Scattering Objects from Holograms", *Applied Optics*, Vol. 13, pp. 162-172, 1974.

[17] D. L. Colton, *Integral Equation Methods in Scattering Theory*, New York: Wiley, 1983.

[18] A. J. Devaney, "Geophysical Diffraction Tomography", *IEEE Trans. Geoscience and Remote Sensing*, Vol. GE-22, No. 1, pp. 3-13, 1984.

[19] G. Demoment, "Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems", *IEEE Trans. Acoustics, Speech And Signal Processing.* Vol. 37. No. 12, pp. 2024-2035, 1989.

[20] G. Archer and D. M. Titterington, "On Some Bayesian/Regularization Methods for Image Restoration", *IEEE Trans. Image Processing.* Vol. 4, No. 7, pp. 989-995, 1995.

[21] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, Washington, DC. Winston. 1977.

[22] Belge, Murat, "Multiscale and Curvature Methods for the Regularization of Linear Inverse Problems", PhD Thesis, Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, August, 1999

[23] H. W. Engl and W. Grever, "Using the l-curve for determining the optimal regularization parameters", *Numer. Math.* Vol. 69, No. 1, pp. 25-31, 1994.

[24] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic Edge-Preserving Regularization in Computed Imaging", *IEEE Trans. Image Processing.* Vol. 6, No. 2. pp, 298-311, 1997.

[25] D. Geman and C. Yang, "Nonlinear Image Recovery with Half-Quadratic Regularization", *IEEE Trans. Image Processing.* Vol. 4, No. 7, pp. 932-945, 1995.

[26] C. Bouman and K. Sauer, "A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation", *IEEE Trans. Image Processing.* Vol. 2, No. 3, pp, 296-310, 1993.

[27] M. Belge, M. Kilmer and E. L. Miller, "Wavelet Domain Bayesian Image Restoration Using Edge Preserving Prior Models", 1998 International Conf. on Image Processing, Chicago Il., Oct. 1998.

[28] P. E. Gill, W. Murry and M. H. Wright, *Practical Optimization*. New York: Academic, 1981.

[29] J. A. Schnabel, "Multi-Scale Active Shape Description in Medical Imaging", Ph.D. thesis (Dissertation), Department of Computer Science, University College London, Gower Street, London. WC1E6BT, 1997.

[30] R. C. Beach, *An Introduction to The Curves and Surfaces of Computer-Aided Design*, New York: Van Nostrand Reinhold, 1991.

[31] G. Farin, *Curves and Surfaces for Computer Aided Geometric Design, A practical Guide*, 2nd edition. Academic 1990.

[32] R. H. Bartels, J. C. Beatty, and B. A. Barsky, *An Introduction to Splines for Use in Computer Graphics & Geometric Modeling*, Morgan Kaufmann, 1987.

[33] T. Cataldi, and T. Rotunno, "Kalman Filter Optimized Cubic Spline Functions for Digital Smoothing", *Analytical Methods and Instrumentation*, Vol. 2, No. 1, pp. 27-34, 1995.

[34] Voskoboinikov, and Y. Ye, "Descriptive splines in Inverse Heat Conduction Problems: Method and Algorithms", *J. Engineering Physics and Thermophysics*, Vol. 65, No. 6, pp. 1230-1253, 1993.

[35] P. Dierckx, *Curve and Surface Fitting with Splines*, New York: Oxford University Press, 1993.

[36] C. de Boor, "On Calculation of B-Splines", *J. Approx. Theory*, Vol. 6, pp. 50-62, 1972.

[37] W. Boehm, and H. Prautzsch, "The Insertion Algorithm", *Computer-Aided Design*, Vol. 17, No. 2, 1985.

[38] M. Eck, and J. Hadenfeld, "Knot removal for B-Spline Curves", *Computer Aided Geometric Design*, Vol. 12, No. 3, pp. 259-272, 1995.

[39] R. N. Goldman, and T. Lyche, "Knot Insertion and Deletion Algorithms for B-Spline Curves and Surfaces", *SIAM*, 1993.

[40] T. Cham, and R. Cipolla, "Automated B-Spline Curve Representation Incorporating MDL and Error-Minimizing Control Point Insertion Strategies", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, No. 1, pp. 49-53, 1999.

[41] G. Farin, and N. Sapidis, "Curvaure and The Fairness of Curves and Surfaces", *IEEE Computer Graphics and Applications*, Vo. 9, pp. 53-57, 1989.

[42] D. Nira, and Y. Itai, "Optimal Distribution of Knots for Tensor-Product Spline Approximation", *Quarterly of Applied Mathematics*, Vol. 49, No. 1, pp. 19-36, 1991.

[43] N. Sapidis, and G. Farin, "Automatic Fairing Algorithm for B-Spline Curves", *Computer Aided Design*, Vol. 22, pp. 121-129, 1990

[44] A. Davies, and P. Samuels, *An Introduction to Computational Geometry for Curves and Surfaces*. New York: Oxford University Press. 1996.

[45] H. Flanders, *Differential Forms with Applications to the Physical Sciences*, New York: Dover Publications, 1989.

[46] H. Radha, R. Leonardi, M. Vetterli, and B. Naylor, "Binary Space Partitioning Tree Representation of Images", *J. Visual Communication and Image Representation.* Vol. 2, No. 3, pp. 202-221, 1991.

[47] G. Hunter and K. Steiglitz, "Operations on Images Using Quad Trees", *IEEE Trans. Pattern Analasys and Machine Intelligence*, Vol. PAMI-1, No. 2, pp. 145-153, 1979.

[48] W. Zhu, Y. Wang, Y. Deng, Y. Yao and R. L. Barbour, "A Wavelet-Based Multiresolution Regularized Least Squares Reconstruction Approach for Optical Tomography", *IEEE Trans. Medical Imaging.* Vol. 16, No. 2, pp. 210-217, 1997.

[49] M. Eck, T. deRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetzle, "Multiresolution Analysis of Arbitrary Meshes", *Computer Graphics(SIGGRAPH'95 Proceedings)*, pp. 173-182, 1995.

[50] W. Welch, and A. Witkin, "Free-form Shape Design Using Triangulated Surfaces", *Computer Graphics(SIGGRAPH'94 Proceedings)*, pp. 247-256, 1994.

[51] E. Horowitz, S. Sahni, and S. Rajasekaran, *Computer Algorithm*, New York: Computer Science Press, 1998.

Appendix A

Derivation of CRB

In our cross-well tomography problem, we set up the following methods:

1. Tikhonov method:

$$G(f) = \frac{1}{\sigma^2}\|y - Af\|_2^2 + \lambda_x^2\|D_x f\|_2^2 + \lambda_y^2\|D_y f\|_2^2 \tag{A.1}$$

2. Edge preserving method:

$$G(f) = \frac{1}{\sigma^2}\|y - Af\|_2^2 + \lambda_x^2\Phi_x^T(f)\Phi_x(f) + \lambda_y^2\Phi_y^T(f)\Phi_y(f) \tag{A.2}$$

It will be helpful to learn the nature of the problem if we can find the Cramer-Rao Bound(CRB) evaluated for the given object f, which represent the estimation variance of data for different cross-well regions.

For this problem we can find the CRB by the following way,

Write G(f) as:

$$G(f) = e^T(f)e(f) \tag{A.3}$$

and Let

$$J = \frac{\partial e(f)}{\partial f} \tag{A.4}$$

Note here f is a column vector, and the partial derivative is defined as:

$$[J]_{i,j} = \frac{\partial e_j}{\partial f_i} \tag{A.5}$$

Thus the Fisher's Information is:

$$F = J^T * J \tag{A.6}$$

and CRB can be found by:

$$CRB = F^{-1} = (J^T * J)^{-1} \tag{A.7}$$

A.1 Tikhonov Method

In this case, it's easily to see that:

$$e(f) = [\frac{1}{\sigma}(y - Af),\ \lambda_x D_x f,\ \lambda_y D_y f]^T \tag{A.8}$$

Where D_x and D_y are derivative matrix such that:

$$Dx(f) = D_x * f = f_{i+1,j} - f_{i,j} \tag{A.9}$$

$$Dy(f) = D_y * f = f_{i,j+1} - f_{i,j} \tag{A.10}$$

Here, simply use the definition of derivative of matrices and vectors:

$$\frac{d}{dx}Ax = A^T \tag{A.11}$$

$$\frac{d}{dx}x^T A = A \tag{A.12}$$

where, A is matrix and x is vector.

We can get:

$$J = \frac{\partial e}{\partial f} = \begin{bmatrix} \frac{1}{\sigma}A \\ \lambda_x D_x \\ \lambda_y D_y \end{bmatrix} \tag{A.13}$$

and,

$$CRB = (J^T * J)^{-1} \tag{A.14}$$

A.2 Edge Preserving Method

In this, we have,

$$e(f) = [\frac{1}{\sigma}(y - Af),\ \lambda_x \Phi(D_x f),\ \lambda_y \Phi(D_y f)]^T \tag{A.15}$$

where,

$$\Phi^T(g)\Phi(g) = \frac{g^2}{1+g^2} \tag{A.16}$$

$$\Phi(g) = \frac{g}{\sqrt{1+g^2}} \tag{A.17}$$

Thus, we can compute the derivative of e(f) by using the following equation:

$$[\frac{\partial \Phi}{\partial f}]_{i,j} = \frac{\partial \Phi_j}{\partial f_i} = \sum_k \frac{\partial \Phi_j}{\partial g_k}\frac{\partial g_k}{\partial f_i} \tag{A.18}$$

First, compute $\frac{\partial \Phi}{\partial g}$, where $g = D_x f$ or $g = D_y f$:

$$\frac{\partial \Phi_j}{\partial g_k} = \frac{\partial}{\partial g_k}(\frac{g_j}{\sqrt{1+g_j^2}}) = (1+g_j^2)^{-\frac{3}{2}}\delta(j-k) \tag{A.19}$$

Second,

$$\frac{\partial D_x f}{\partial f} = D_x^T \tag{A.20}$$

$$\frac{\partial D_y f}{\partial f} = D_y^T \tag{A.21}$$

Therefore, we get,

$$[\frac{\partial \Phi}{\partial f}]_{i,j} = \frac{\partial \Phi_j}{\partial f_i} = \sum_k \frac{1}{(1+g_j^2)^{\frac{3}{2}}} \delta(j-k)[D_x]_{k,i} = \frac{1}{(1+g_j^2)^{\frac{3}{2}}}[D_x]_{j,i} \tag{A.22}$$

Write it in matrix form, we get:

$$\frac{\partial \Phi_x}{\partial f} = G_x * D_x,\ \frac{\partial \Phi_y}{\partial f} = G_y * D_y \tag{A.23}$$

where

$$G_x = diag([1+(D_x f)_i^2]^{-\frac{3}{2}}),\ G_y = diag([1+(D_x f)_i^2]^{-\frac{3}{2}}) \tag{A.24}$$

Finally, we get:

$$J = \frac{\partial e}{\partial f} = \begin{bmatrix} \frac{1}{\sigma}A, \\ \lambda_x G_x D_x, \\ \lambda_y G_y D_y \end{bmatrix} \tag{A.25}$$

and

$$CRB = (J^T * J)^{-1} \tag{A.26}$$

A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets

Oliver Dorn* Eric L. Miller* Carey M. Rappaport *

March 20, 2000

Abstract

A two-step shape reconstruction method for electromagnetic (EM) tomography is presented which uses adjoint fields and level sets. The inhomogeneous background permittivity distribution and the values of the permittivities in some penetrable obstacles are assumed to be known, and the number, sizes, shapes, and locations of these obstacles have to be reconstructed given noisy limited-view EM data. The main application we address in the paper is the imaging and monitoring of pollutant plumes in environmental cleanup sites based on cross-borehole EM data. The first step of the reconstruction scheme makes use of an inverse scattering solver which first recovers equivalent scattering sources for a number of experiments, and then calculates from these an approximation for the permittivity distribution in the medium. The second step uses this result as an initial guess for solving the shape reconstruction problem. A key point in this second step is the fusion of the 'level set technique' for representing the shapes of the reconstructed obstacles, and an 'adjoint field technique' for solving the nonlinear inverse problem. In each step, a forward and an adjoint Helmholtz problem are solved based on the permittivity distribution which corresponds to the latest best guess for the representing level set function. A correction for this level set function is then calculated *directly* by combining the results of these two runs. Numerical experiments are presented which show that the derived method is able to recover one or more objects with nontrivial shapes given noisy cross-borehole EM data.

*Center for Electromagnetics Research, Northeastern University, Boston, MA, E-mail: dorn@cer.neu.edu, elmiller@ece.neu.edu, rappaport@neu.edu

Contents

1 Introduction **3**

2 The physical experiment **6**

2.1 The Helmholtz Equation 6

2.2 Formulation of the inverse problem 7

3 The shape reconstruction problem **8**

3.1 Shape reconstruction and inverse scattering 8

3.2 The domains $\Omega^{(n)}$ 9

3.3 Level set representation of the domains $\Omega^{(n)}$ 10

4 Step 2: Solving the shape reconstruction problem **11**

4.1 Function spaces 12

4.2 Operators 12

4.3 Linearized operators 14

4.4 A nonlinear Kaczmarz-type approach 15

4.5 The minimization problem 16

4.6 The adjoint linearized operators 17

4.7 The operators $T'_{jk}T'^{*}_{jk}$ 18

4.8 Regularization of $T'_{jk}T'^{*}_{jk}$ 18

4.9 Updating the level set function 20

4.10 Implementation: The levelART algorithm 22

5 Step 1: A Source-Type Adjoint Field method **23**

5.1 Source-type methods 24

5.2 Solving the inverse scattering problem 25

5.3 Looking for a scattering source 26

5.4 Recovery of the scattering potential 28

5.5 Combining the results from single experiments 29

5.6 The initial level set function 30

5.7 Implementation: The STAF algorithm 31

6 Numerical Experiments **31**

6.1 Discretization of the computational domain. 31

6.2 A full-view example 33

6.3 A cross-borehole situation with multiple objects 34

6.4 A cross-borehole situation with a single inclusion and variable conductivity 36

7 Summary and future directions **37**

8 Acknowledgments **39**

9 Appendix **39**

9.1 Proof of theorem 5.1 39

9.2 Proof of theorem 5.2 40

10 Figures **43**

1 Introduction

In this paper, we investigate the retrieval of an unknown number of penetrable objects (inclusions) imbedded in an inhomogeneous background medium based on observations of electromagnetic (EM) fields. The electromagnetic characteristics (permittivity and conductivity) of the background medium as well as of the material forming the inclusions are assumed to be known, but the main topological information concerning the number, sizes, shapes, and locations of the inclusions is missing and has to be reconstructed from the EM data.

One possible technique for using EM fields in cross-borehole tomography is ElectroMagnetic Induction Tomography (EMIT) [3, 18, 44, 49, 50, 51] which typically operates at frequencies between 1 to 20 kHz. In this frequency band, electromagnetic fields tend to diffuse rather than propagate as waves through the Earth. Penetration depths of 100 m or more are possible at these low frequencies, but the diffusional behavior of the fields makes the inverse problem severely ill-posed.

However, if the typical distances in the area of interest are not much larger than 10-20 m, we can use EM fields in the higher frequency band of 5 to 30 MHz instead. The wavelengths of these fields are typically between 2-15 m in moist soil, where the relative dielectric constant is typically around 20 [48]. Therefore, we can make use of wave propagation phenomena in the inversion process. In the present paper, we address this situation. The main application we have in mind is the imaging and monitoring of pollutant plumes at environmental cleanup sites given cross-borehole EM data, where the distances of the boreholes are not much larger than 10-20 m.

We assume that the known conductivity distribution is positive but small everywhere, and that the permittivity distribution in the medium has to be recovered. Inside the pollutant plumes, the permittivity is assumed to be constant with a known value, and the background permittivity is arbitrary but also known. Therefore, the task is to find the *number, shapes, sizes and locations of the pollutant plumes* from cross-borehole data gathered for a small number (less than 10) of frequencies. No

topological constraints are made on the shapes of these plumes. For example, they are allowed to be multiply connected, and to enclose 'cavities' or 'holes' filled with background material.

The main difficulties which arise in this situation are 1.) We want to allow for an (arbitrary) inhomogeneous background permittivity distribution in the inversion; 2.) The inverse problem is usually strongly nonlinear because of the high contrast of the permittivity values inside the plumes to the background medium; 3.) The data in our application are typically noisy and have only limited view; and 4.) The number of the plumes is typically unknown, and their shapes can have a complicated geometry.

In this paper, we propose a new shape reconstruction method which works in a two-step fashion in order to overcome these difficulties.

The first step of this combined inversion scheme plays the role of an initializing procedure for the second step, and employs a 'source-type' inversion method (which is described in more details in section 5) to deal with the high nonlinearity in the problem due to the presence of strong scatterers.

Then, the second step directly starts with the outcome of this initializing procedure, and continues by using a combination of the 'adjoint field technique' and a level set representation of the shapes until the inversion task is completed. Using a level set representation in this second step enables us to easily describe and keep track of complicated geometries which arise during the inversion process.

Both steps use an 'adjoint field technique' for the inversion which has the very useful property that the inverse problem can be solved approximately by making two uses of the same forward modelling code. Using a somewhat oversimplified description of our technique, the updates to the level set function are obtained by first making one pass through the code using the permittivity distribution corresponding to the latest best guess of the level set function, and then another pass with the adjoint operator applied to the differences in computed and measured data. Then the results of these two calculations are combined to determine updates to the level set function. The resulting procedure is iterative, and can be applied successively to parts of the data, e.g., data associated with one transmitter location and one frequency can be used to update the model before other transmitter locations and other frequencies are considered. This general procedure has several of the same advantages as wave equation migration in reflection seismology [10] and is also related to recent methods in EM migration introduced in Zhdanov *et al.* [51]. A similar technique has been successfully applied recently as part of an iterative nonlinear inversion scheme in [17, 18, 33].

The *level set method* was originally developed by Osher and Sethian for describ-

ing the motion of curves and surfaces [35, 41]. Since then, it has found applications in a variety of quite different situations. Examples are image enhancement, computer vision, interface problems, crystal growth, or etching and deposition in the microchip fabrication. For an overview we refer to [42].

The idea of using a level set representation as part of a solution scheme for inverse problems involving obstacles was first suggested by Santosa in [40]. More recently, a similar method was applied to a nonlinear inverse scattering problem by Litman *et al.* in [28]. In that work, an inverse transmission problem in free space is solved by a controlled evolution of a level set function. This evolution is governed by a Hamilton-Jacobi type equation, whose velocity function has to be determined properly in order to minimize a given cost functional.

The approach developed here does not lead to a Hamilton-Jacobi type equation. We follow an optimization approach, and employ a very specific inversion routine (an adjoint field technique) for solving it. This has the advantage that we do not have to propagate the level set function explicitly by computing a numerical Hamiltonian. Instead, our inversion routine provides us in each step with an update that has to be applied *directly* to the most recent level set function. Doing so, we automatically 'propagate' the level set function until the method converges.

This gain in simplicity, however, has its price. In order to arrive at an efficient scheme which is practically useful as well as easy to implement, we will apply some suitable approximations when deriving the algorithm. We will point out and discuss these approximations in those sections of the paper where they are applied.

For interesting approaches to solving shape recovery problems in various applications we refer to [23, 24, 26] and to the references therein. For alternative approaches to the shape reconstruction method in geophysical applications see [30, 39, 43] and the references therein. The treatment of more general inverse scattering problems is for example addressed in [4, 8, 11, 12, 15, 22, 31, 33, 38].

The paper is organized as follows. In section 2 we will present the basic equations of 2D EMs in a form convenient for development of the shape reconstruction technique. In section 3 we formulate the shape reconstruction problem and introduce the level set formulation of this problem. In section 4, we derive the basic shape reconstruction algorithm using level sets and adjoint fields. Section 5 describes how to calculate a suitable initialization for the shape reconstruction algorithm. In Section 6 numerical experiments are presented which demonstrate the performance of the algorithm in different situations. The final section summarizes the results of this paper and indicates some directions for future research.

2 The physical experiment

2.1 The Helmholtz Equation

We consider the 2D Helmholtz Equation

$$\Delta u + k^2(x)u \;=\; q(x) \quad \text{ in } \mathbb{R}^2, \tag{1}$$

with complex wavenumber

$$k^2(x) \;=\; \omega^2 \mu_0 \epsilon_0 \left[\epsilon(x) + i\frac{\sigma(x)}{\omega\epsilon_0}\right]. \tag{2}$$

Here, $i^2 = -1$, ω denotes the angular frequency $\omega = 2\pi f$, μ_0 is the magnetic permeability in free space $\mu_0 = 4\pi \times 10^{-7}$ Henrys per meter, ϵ_0 is the dielectric permittivity in free space $\epsilon_0 = 8.854 \times 10^{-12}$ Farads per meter, ϵ is the relative dielectric permittivity (dimensionless), and σ is the electric conductivity in Siemens per meter. The form of (2) corresponds to time-harmonic line sources $\tilde{q}(x,t)$ which have a time-dependence $\tilde{q}(x,t) = q(x)e^{-i\omega t}$. For these sources we require that there exists a radius $r_0 > 0$ such that $\text{supp}(q) \subset\subset B_{r_0}(0)$, where $B_r(x) = \{y \in \mathbb{R},\ |x-y| < r\}$ denotes the open ball centered in x with radius $r > 0$. For simplicity we assume throughout the paper that we can find a ball $B_R(0)$ with $R > r$ such that the complex wavenumber $k^2(x)$ is constant with value k_0^2 in $\mathbb{R}^2 \backslash B_R(0)$, and that for this k_0 the field u generated by (1) satisfies the Sommerfeld radiation condition

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial u}{\partial r} \;-\; ik_0 u\right) \;=\; 0 \tag{3}$$

with $r = |x|$ where the limit is assumed to hold uniformly in all directions $x/|x|$. With this assumption, the problem (1)-(3) possesses a uniquely determined solution u in $\mathbb{R}^2$ [12].

Furthermore we will consider in this paper only the case that the conductivity is positive everywhere, $\sigma > 0$ in $\mathbb{R}^2$, and that it is small in some sense which will be specified later. Typical values in our geophysical examples will be $\sigma \approx 10^{-3} - 10^{-4}$ Siemens per meter or less [48].

We want to introduce some notation here which will be useful in the following. We denote the wavenumber $k^2(x)$ in short form by

$$k^2(x) \;=\; \kappa(x) \;=\; a\epsilon(x) + ib\sigma(x), \quad a = \omega^2\mu_0\epsilon_0, \quad b = \omega\mu_0. \tag{4}$$

We only consider positive frequencies $\omega > 0$ such that $a, b > 0$.

2.2 Formulation of the inverse problem

We assume that we are given p different source distributions q_j, $j = 1, \ldots, p$. For each of these sources, data are gathered at the detector positions x_d, $d = 1, \ldots, D_j$, for various frequencies f_k, $k = 1, \ldots, K$. The total number of receivers D_j, as well as their positions x_d, might vary with the source q_j. We assume, for simplicity in the notation, that these positions do not depend on the frequency f_k. This restriction is, however, not necessary for the derivation of the inversion method. We require that there exists a radius $r_1 > 0$ such that all receiver positions are inside the ball of radius r_1, i.e. $x_{jd} \in B_{r_1}(0)$ for all $d = 1, \ldots, D_j$, $j = 1, \ldots, p$.

For a given source q_j and a given frequency f_k we collect a set of data $\tilde{G}_{jk}$ which is described by

$$\tilde{G}_{jk} \;=\; \Big(\tilde{u}_{jk}(x_{j1}), \ldots, \tilde{u}_{jk}(x_{jd}), \ldots, \tilde{u}_{jk}(x_{jD_j})\Big)^T \quad \in\; Z_j \tag{5}$$

with $Z_j = \mathbb{C}^{D_j}$ being the data space corresponding to a single experiment using one source and one frequency only. In (5), the fields $\tilde{u}_{jk}$ solve (1)-(3) with the correct permittivity distribution $\tilde{\epsilon}(x)$, i.e.

$$\Delta \tilde{u}_{jk} \;+\; [a_k \tilde{\epsilon}(x) + i b_k \sigma(x)]\, \tilde{u}_{jk} \;=\; q_j(x) \qquad \text{in } \mathbb{R}^2 \tag{6}$$

with

$$a_k = \omega_k^2 \mu_0 \epsilon_0, \quad b_k = \omega_k \mu_0, \quad \omega_k = 2\pi f_k. \tag{7}$$

In a slightly more formal way, we define for a given source q_j the measurement operator M_j acting on solutions u of (1) by

$$M_j u \;=\; \left(\int_{\mathbb{R}^2} u(x) \delta(x - x_{jd}) dx \right)^T_{d=1,\ldots,D_j}. \tag{8}$$

With this notation, (5) is written as

$$\tilde{G}_{jk} \;=\; M_j \tilde{u}_{jk}, \qquad j = 1, \ldots, p,\; k = 1, \ldots, K. \tag{9}$$

We gather these data sets $\tilde{G}_{j,k}$ for all sources q_j, $j = 1, \ldots, p$, and all frequencies f_k, $k = 1, \ldots, K$, and the aim is to recover from this collection of data sets

$$\tilde{G} = (\tilde{G}_{1,1}, \ldots, \tilde{G}_{p,K})^T \tag{10}$$

the unknown parameter distribution $\tilde{\epsilon}(x)$ in the domain of interest.

In the application of *EM cross-borehole tomography*, the sources and receivers are typically situated in some boreholes, and the permittivity distribution ϵ (and/or the conductivity distribution σ) between these boreholes has to be recovered from

the gathered data. In the 2D geometry considered here, typical sources are time-harmonic line sources which can be modelled in (1) by

$$q_j(x) \;=\; J_j\delta(x - x_j), \quad j = 1, \ldots, p, \tag{11}$$

where x_j denotes the 2D coordinates of the j-th line source, $j = 1, \ldots, p$, and the complex number J_j is the strength of the source. We will use these sources in our numerical experiments in section 6.

3 The shape reconstruction problem

In this section we formulate the shape reconstruction problem which we want to solve, and cast it in a form which makes use of the level set representation of the domains.

3.1 Shape reconstruction and inverse scattering

To start with we introduce some terminology which we will use throughout the paper.

Definition 3.1 *Let us assume that we are given a constant $\hat{\epsilon} > 0$, an open ball $B_r(0) \subset \mathbb{R}^2$ with $r > \max(r_0, r_1) > 0$, and a bounded function $\epsilon_b : \mathbb{R}^2 \to \mathbb{R}$. We call a pair (Ω, ϵ), which consists of a compact domain $\Omega \subset\subset B_r(0)$ and a bounded function $\epsilon : \mathbb{R}^2 \to \mathbb{R}$,* admissible *if we have*

$$\epsilon|_\Omega = \hat{\epsilon}, \quad \epsilon|_{\mathbf{R}^2\setminus\Omega} = \epsilon_{\mathrm{b}}|_{\mathbf{R}^2\setminus\Omega}. \tag{12}$$

In other words, a pair (Ω, ϵ) is *admissible* if ϵ is equal to a preassigned constant value $\hat{\epsilon}$ inside of Ω, and equal to the preassigned *background permittivity* ϵ_{b} outside of Ω. The domain Ω is called the *scattering domain*.

Remark 3.1 *For an admissible pair (Ω, ϵ), and for given $\hat{\epsilon}$, ϵ_{b}, the permittivity ϵ is uniquely determined by Ω.*

With this definition, we can now formulate the *shape reconstruction problem*.

Shape reconstruction problem. Let us assume that we are given a constant $\hat{\epsilon} > 0$, a bounded function $\epsilon_{\mathrm{b}} : \mathbb{R}^2 \to \mathbb{R}$, and some data $\tilde{G}$ as in (10). Find a domain $\tilde{\Omega}$ such that the admissible pair $(\tilde{\Omega}, \tilde{\epsilon})$ reproduces the data, i.e. (9) holds with $\tilde{u}_{jk}$ given by (6) for $j = 1, \ldots, p,\ k = 1, \ldots, K$.

Using the same notation and assumptions as in definition 3.1, we want to formulate another inverse problem which we will call the *inverse scattering problem* and which will play an important part when solving the shape reconstruction problem.

Inverse Scattering Problem. Let us assume that we are given a bounded function $\epsilon_{\mathrm{b}} : \mathbb{R}^2 \to \mathbb{R}$, and some data $\tilde{G}$ as in (10). Find a bounded function $\tilde{\epsilon}_{\mathrm{s}} : \mathbb{R}^2 \to \mathbb{R}$ with $\mathrm{supp}(\epsilon_{\mathrm{s}}) \subset\subset B_r(0)$ such that $\tilde{\epsilon} = \epsilon_{\mathrm{b}} + \tilde{\epsilon}_{\mathrm{s}}$ reproduces the data, i.e. (9) holds with $\tilde{u}_{jk}$ given by (6) for $j = 1, \ldots, p,\ k = 1, \ldots, K$.

The inverse scattering problem gives rise to the following decomposition of ϵ in $\mathbb{R}^2$.

Decomposition of $\epsilon(x)$:

$$\text{(i)} \quad \epsilon = \epsilon_{\mathrm{b}} + \epsilon_{\mathrm{s}} \quad \text{in } \mathbb{R}^2 \tag{13}$$

$$\text{(ii)} \quad \mathrm{supp}(\epsilon_{\mathrm{s}}) \subset\subset B_r(0). \tag{14}$$

In other words, the permittivity distribution ϵ is decomposed into the background distribution ϵ_{b} and the perturbation ϵ_{s} which is assumed to have compact support and which we will refer to as the *scattering potential* in the following.

Solving the shape reconstruction problem requires only to find the *shape* of the domain $\tilde{\Omega}$, since the function $\tilde{\epsilon}$ is then uniquely determined by (12). Solving the inverse scattering problem, on the other hand, amounts to finding the entire function $\tilde{\epsilon}_{\mathrm{s}}$ from the given data, which is much harder to do. However, it will turn out that finding a good *approximate* solution of the inverse scattering problem is much easier to achieve and will provide us with an excellent initial guess for starting our shape reconstruction routine.

Definition 3.1 allows us to formulate a first version of the strategy which we want to use for solving the shape reconstruction problem.

Strategy for solving the shape reconstruction problem. Construct a series of admissible pairs $(\Omega^{(n)}, \epsilon^{(n)})$, $n = 0, 1, 2, \ldots$, such that the misfit between the data (10) and the calculated data corresponding to $(\Omega^{(n)}, \epsilon^{(n)})$ decreases with increasing n, and ideally, i.e. in absence of noise, tends to zero in the limit $n \to \infty$. Use the *approximate* solution of the inverse scattering problem (i),(ii) to initialize this series by determining a good starting element $(\Omega^{(0)}, \epsilon^{(0)})$.

3.2 The domains $\Omega^{(n)}$

In our numerical examples, each of the domains $\Omega^{(n)}$ which we are looking for can be given as a collection of a finite number L_n of disjoint, compact subdomains $\Omega_l^{(n)}$,

$l = 1, \ldots, L_n$, with

$$\Omega^{(n)} = \bigcup_{l=1}^{L_n} \Omega_l^{(n)}, \qquad \Omega_l^{(n)} \cap \Omega_{l'}^{(n)} = \emptyset \quad \text{for } l \neq l'. \tag{15}$$

The shapes of these subdomains $\Omega_l^{(n)}$ can in principle be arbitrary. In particular, they are allowed to be multiply connected, and to enclose some 'cavities' or 'holes' filled with background material. Moreover, the number L_n of these subdomains might (and usually does) vary with the iteration number n. For the derivation of the inversion method, we assume that the boundaries $\partial\Omega_l^{(n)}$ of these domains are sufficiently smooth (e.g. C^1).

It is essential for the success and the efficiency of the reconstruction scheme to have a good and flexible way of keeping track of the shape evolution during the reconstruction process. The method we have chosen in our reconstruction algorithm is a *level set representation* of the shapes as it was suggested by Santosa [40]. This representation has the advantage that the level set functions, which are in principle only used for *representing* the shapes, can in a natural way be made part of the reconstruction scheme itself. Doing so, it is not necessary anymore to refer to the shapes of the domains until the reconstruction process is completed. The *final* shape is then recovered from the representing level set function easily. In the following we will discuss in a more formal way how this can be achieved.

3.3 Level set representation of the domains $\Omega^{(n)}$

Assume that we are given a domain $\Omega \subset\subset B_r(0)$. The characteristic function $\chi_\Omega : \mathbb{R}^2 \to \{0,1\}$ is defined in the usual way as

$$\chi_\Omega(x) = \begin{cases} 1 & , \quad x \in \Omega \\ 0 & , \quad x \in \mathbb{R}^2 \backslash \Omega. \end{cases} \tag{16}$$

Definition 3.2 *We call a function $\phi : \mathbb{R}^2 \to \mathbb{R}$ a* level set representation of Ω *if*

$$\chi_\Omega(x) = \Psi_\phi(x) \qquad \textit{on } \mathbb{R}^2 \tag{17}$$

where $\Psi_\phi : \mathbb{R}^2 \to \{0,1\}$ is defined as

$$\Psi_\phi(x) = \begin{cases} 1 & , \quad \phi(x) \leq 0 \\ 0 & , \quad \phi(x) > 0. \end{cases} \tag{18}$$

For each function $\phi : \mathbb{R}^2 \to \mathbb{R}$ there is a domain Ω associated with ϕ by (17),(18) which we call $\Omega[\phi]$. It is clear that different functions ϕ_1, ϕ_2, $\phi_1 \neq \phi_2$, can be associated with the same domain $\Omega[\phi_1] = \Omega[\phi_2]$, but that different domains cannot have

the same level set representation. Therefore, we can use the level set representation for unambiguously specifying a domain Ω by any one of its associated level set functions.

The boundary $\Gamma = \partial\Omega[\phi]$ of a domain $\Omega[\phi]$, represented by the level set function ϕ, is defined as

$$\begin{aligned} \Gamma \quad = \quad \{x \in \mathbb{R}^2 : \quad & \text{for all } \epsilon > 0 \text{ we can find } x_1, x_2 \in B_\epsilon(x) \\ & \text{with } \phi(x_1) > 0 \text{ and } \phi(x_2) < 0 \,\} \end{aligned} \tag{19}$$

Definition 3.3 *We call a triple (Ω, ϵ, ϕ), which consists of a domain $\Omega \subset\subset B_r(0)$ and bounded functions $\epsilon, \phi : \mathbb{R}^2 \to \mathbb{R}$,* admissible *if the pair (Ω, ϵ) is admissible in the sense of definition 3.1, and ϕ is a valid level set representation of Ω.*

Remark 3.2 *For an admissible triple (Ω, ϵ, ϕ), and for given $\hat{\epsilon}$, ϵ_{b}, the pair (Ω, ϵ) is uniquely determined by ϕ.*

We use these definitions to reformulate our shape reconstruction problem.

Level set formulation of the shape reconstruction problem. Given a constant $\hat{\epsilon} > 0$, a background distribution ϵ_{b}, and some data $\tilde{G}$ as in (10). Find a level set function $\tilde{\phi}$ such that the corresponding admissible triple $(\tilde{\Omega}, \tilde{\epsilon}, \tilde{\phi})$ reproduces the data, i.e. (9) holds with $\tilde{u}_{jk}$ given by (6) for $j = 1, \ldots, p,\ k = 1, \ldots, K$.

The strategy for solving this shape reconstruction problem has to be reformulated, too. It reads now as follows.

Strategy for solving the reformulated shape reconstruction problem. Construct a series of admissible triples $(\Omega^{(n)}, \epsilon^{(n)}, \phi^{(n)})$, $n = 0, 1, 2, \ldots$, such that the misfit between the data (10) and the calculated data corresponding to $(\Omega^{(n)}, \epsilon^{(n)}, \phi^{(n)})$ decreases with increasing n, and ideally, i.e. in absence of noise, tends to zero in the limit $n \to \infty$. *For finding this series we only have to keep track of $\epsilon^{(n)}$ and $\phi^{(n)}$, but not of $\Omega^{(n)}$.* The function $\epsilon^{(n)}$ is needed in each step for solving a forward problem (1), and a corresponding adjoint problem. The knowledge of $\phi^{(n)}$ is used in each step to determine $\epsilon^{(n)}$. The final level set function $\phi^{(N)}$, which satisfies some stopping criterion, is used to recover the final shape $\Omega^{(N)}$ via (17).

4 Step 2: Solving the shape reconstruction problem

In this section we derive the basic shape reconstruction method which uses adjoint fields and the level set representation introduced above. The initializing procedure ('Step 1') for this reconstruction routine will be discussed in section 5.

4.1 Function spaces

We want to specify now the function spaces which we will be working with. The main objective of this section is to introduce the inner products on these function spaces, which will become important when defining the adjoint linearized operators in sections 4.6 and 5.3.

The space of sources and scattering sources Y is defined as

$$Y = \Big\{ q : \mathbb{R}^2 \to \mathbb{C}, \quad q = 0 \text{ on } \mathbb{R}^2 \backslash B_r(0), \quad \int_{B_r(0)} |q|^2 dx < \infty \Big\}, \tag{20}$$

$$\langle q_1 \, , \, q_2 \rangle_Y = \int_{B_r(0)} q_1(x) \overline{q_2(x)} \, dx, \tag{21}$$

where the bar means 'complex conjugate'. The space F of scattering potentials is defined as

$$F = \Big\{ \epsilon_{\mathrm{s}} : \mathbb{R}^2 \to \mathbb{R}, \quad \epsilon_{\mathrm{s}} = 0 \text{ on } \mathbb{R}^2 \backslash B_r(0), \quad \int_{B_r(0)} |\epsilon_{\mathrm{s}}|^2 dx < \infty \Big\}, \tag{22}$$

$$\langle \epsilon_{s,1} \, , \, \epsilon_{s,2} \rangle_F = \int_{B_r(0)} \gamma \, \epsilon_{s,1}(x) \epsilon_{s,2}(x) \, dx, \tag{23}$$

with some positive weighting factor $\gamma > 0$ which is introduced here for convenience. Analogously, the space of level set functions Φ is defined as

$$\Phi = \Big\{ \phi : \mathbb{R}^2 \to \mathbb{R}, \quad \phi = 0 \text{ on } \mathbb{R}^2 \backslash B_r(0), \quad \int_{B_r(0)} |\phi|^2 dx < \infty \Big\}, \tag{24}$$

$$\langle \phi_1 \, , \, \phi_2 \rangle_\Phi = \int_{B_r(0)} \gamma \, \phi_1(x) \phi_2(x) \, dx. \tag{25}$$

This space for the level set functions is mainly chosen in order to have an inner product available which is convenient for the derivation of the shape reconstruction algorithm. We mention that the regularity of an arbitrary function in Φ is, strictly speaking, not sufficient for our purposes, such that we will apply further regularity constraints on those level set functions $\phi \in \Phi$ which we choose for representing the boundaries $\partial\Omega[\phi]$.

The data space Z_j corresponding to source q_j, $j = 1, \ldots, p$, was already introduced earlier, and is given by $Z_j = \mathbb{C}^{D_j}$, where D_j is the total number of receivers corresponding to source q_j.

4.2 Operators

In the following, we will introduce some operators which will enable us to formulate the shape reconstruction problem in a way suitable for deriving the inversion algorithm.

Given a constant $\hat{\epsilon}$ and a bounded function $\epsilon_{\rm b} : \mathbb{R}^2 \to \mathbb{R}$. Then, with each level set function $\phi \in \Phi$ a uniquely determined scattering potential $\Lambda(\phi)$ is associated by putting

$$\Lambda(\phi)(x) \;=\; \begin{cases} \hat{\epsilon} - \epsilon_{\rm b}(x) & , \quad \phi(x) \le 0 \\ 0 & , \quad \phi(x) > 0. \end{cases} \tag{26}$$

With (18) we can write this also as

$$\Lambda(\phi)(x) \;=\; \Psi_\phi(x)(\hat{\epsilon} - \epsilon_{\rm b}(x)) \quad , \quad x \in \mathbb{R}^2. \tag{27}$$

Notice that the operator Λ is chosen such that the triple (Ω, ϵ, ϕ) with $\epsilon = \epsilon_{\rm b} + \Lambda(\phi)$ and domain $\Omega[\phi]$ forms an admissible triple (Ω, ϵ, ϕ) in the sense of definition 3.3. Moreover, for (Ω, ϵ, ϕ) an admissible triple, we see that $\Lambda(\phi)$ is just the scattering potential $\epsilon_{\rm s}$ as defined in (13),(14)

$$\Lambda(\phi)(x) \;=\; \epsilon_{\rm s}(x) \;=\; \chi_\Omega(x)(\hat{\epsilon} - \epsilon_{\rm b}(x)) \quad , \quad x \in \mathbb{R}^2. \tag{28}$$

Let us assume now that we are given a background permittivity $\epsilon_{\rm b}$ and that we have collected some data $\tilde{G}_{jk}$ which correspond to the 'true' permittivity distribution

$$\tilde{\epsilon} \;=\; \epsilon_{\rm b} + \tilde{\epsilon}_s, \tag{29}$$

where $\tilde{\epsilon}_s$ is the 'true' scattering potential. The *residual operators* R_{jk} map for a source position q_j and a frequency f_k a given scattering potential ϵ_s to the corresponding mismatch in the data

$$R_{jk} : F \longrightarrow Z_j \quad , \quad R_{jk}(\epsilon_{\rm s}) \;=\; M_j u_{jk} - \tilde{G}_{jk} \tag{30}$$

where u_{jk} solves

$$\Delta u_{jk} \;+\; [a_k(\epsilon_{\rm b} + \epsilon_{\rm s})(x) + ib\sigma(x)]\, u_{jk} \;=\; q_j \tag{31}$$

and M_j is the measurement operator defined in (8). From (9) we see that for the 'true' scattering potential the residuals vanish,

$$R_{jk}(\tilde{\epsilon}_{\rm s}) \;=\; 0 \qquad \text{for } j = 1, \ldots, p,\; k = 1, \ldots, K, \tag{32}$$

if the data are noise-free.

The *forward operators* T_{jk} which map a given level set function $\phi \in \Phi$ into the corresponding mismatch in the data are defined by

$$T_{jk} : \Phi \longrightarrow Z_j \quad , \quad T_{jk}(\phi) \;=\; R_{jk}(\Lambda(\phi)) \tag{33}$$

for $j = 1, \ldots, p$, $k = 1, \ldots, K$. The goal is to find a level set function $\tilde{\phi} \in \Phi$ such that

$$T_{jk}(\tilde{\phi}) \;=\; 0 \qquad \text{for } j = 1, \ldots, p,\; k = 1, \ldots, K. \tag{34}$$

We mention that all three operators Λ, R_{jk} and T_{jk} are nonlinear.

4.3 Linearized operators

For the derivation of the shape reconstruction algorithm, we will need expressions for the linearized operators corresponding to the nonlinear operators introduced above, and for their adjoints with respect to the given inner products. In this section, we define the linearized operators, and expressions for their adjoints are derived in section 4.6.

In Santosa [40] it is shown that, for a homogeneous background $\epsilon_{\rm b}$, the infinitesimal response $\delta\epsilon_{\rm s}(x)$ in the scattering potential $\epsilon_{\rm s}(x)$ to an infinitesimal change $\delta\phi(x)$ of the level set function $\phi(x)$ has the form

$$\delta\epsilon_{\rm s}(x) \;=\; -\,[\hat{\epsilon}-\epsilon_{\rm b}]\,\frac{\delta\phi(x)}{|\nabla\phi(x)|}\bigg|_{x\in\partial\Omega[\phi]}. \tag{35}$$

The function $\delta\epsilon_{\rm s}$ in (35) can be interpreted as a 'surface measure' on the boundary $\Gamma=\partial\Omega[\phi]$. Similar to (35), we would like to define the linearized operator $\tilde{\Lambda}'[\phi]$ by

$$\left(\tilde{\Lambda}'[\phi]\delta\phi\right)(x) \;=\; -\,[\hat{\epsilon}-\epsilon_{\rm b}(x)]\,\frac{\delta\phi(x)}{|\nabla\phi(x)|}\hat{\delta}_\Gamma(x) \tag{36}$$

where $\hat{\delta}_\Gamma(x)$ denotes the Dirac delta distribution concentrated on $\Gamma=\partial\Omega[\phi]$. In this interpretation, (36) describes an infinitesimal 'surface load' of permittivity on Γ which has to be recovered from the mismatch in the data.

However, the expression on the right hand side of (36) is not an element of F which causes problems when we want to calculate the inner products defined in section 4.1. Therefore, we will introduce an approximation to the operator (36) which maps from Φ into F and which will be more convenient for the derivation of the reconstruction method.

For a given level set function $\phi\in\Phi$, let $\Gamma=\partial\Omega[\phi]$ and $B_\rho(\Gamma)=\cup_{y\in\Gamma}B_\rho(y)$ a small neighborhood of Γ with some given constant $0<\rho\ll 1$. The (approximated) linearized operator $\Lambda'[\phi]$ is defined as

$$\Lambda'[\phi]\;:\Phi\longrightarrow F,\quad \left(\Lambda'[\phi]\delta\phi\right)(x) \;=\; -\,[\hat{\epsilon}-\epsilon_{\rm b}(x)]\,\frac{\delta\phi(x)}{|\nabla\phi(x)|}C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x). \tag{37}$$

Here, $C_\rho(\Gamma)=L(\Gamma)/{\rm Vol}(B_\rho(\Gamma))$ where $L(\Gamma)=\int_{B_r(0)}\hat{\delta}_\Gamma(x)dx$ is the length of the boundary Γ, and ${\rm Vol}(B_\rho(\Gamma))=\int_{B_r(0)}\chi_{B_\rho(\Gamma)}(x)dx$ is the volume of $B_\rho(\Gamma)$. For a very small ρ we will get a very large weight $C_\rho(\Gamma)$, whereas for increasing ρ this weight $C_\rho(\Gamma)$ decreases accordingly. The operator defined in (37) maps now from Φ into F such that we can make use of the inner products defined on these spaces.

We mention that the term $|\nabla\phi(x)|$ in (35), (36), as well as the derivation of these expressions, implies some regularity constraint on ϕ. For example, $\phi\in C^1$ would be possible. Another possibility would be to use a 'signed distance function'

as a standard representation of the boundary [42]. We do not want to specify the regularity of ϕ at this point, but assume instead that it is 'sufficiently smooth' for our purposes.

The linearized residual operator $R'_{jk}[\epsilon_{\rm s}]$ is defined by

$$R'_{jk}[\epsilon_{\rm s}] : F \longrightarrow Z_j \quad , \quad R'_{jk}[\epsilon_{\rm s}]\delta\epsilon_{\rm s} \;=\; M_j v_{jk} \tag{38}$$

where v_{jk} solves the linearized equation

$$\Delta v_{jk} \;+\; [a_k(\epsilon_{\rm b} + \epsilon_{\rm s})(x) + ib\sigma(x)]\, v_{jk} \;=\; -\, a_k \delta\epsilon_{\rm s}(x) u_{jk}(x) \tag{39}$$

with u_{jk} a solution of (31). This representation can be derived by perturbing

$$\epsilon_{\rm s} \;\to\; \epsilon_{\rm s} + \delta\epsilon_{\rm s} \quad , \quad u_{jk} \;\to\; u_{jk} + v_{jk}, \tag{40}$$

plugging this into (31) and neglecting terms which are of higher than linear order in the perturbations $\delta\epsilon_{\rm s}$, v_{jk}.

Notice that the right hand side of (39) can be interpreted as a 'scattering source'. We will use this concept later for solving our inverse scattering problem approximately in order to find a starting guess for the shape reconstruction scheme. But we want to mention here already that the linearization assumption built into (38), (39), namely that v_{jk} is small compared to u_{jk}, will not be necessary when solving the inverse scattering problem. That will allow us to circumvent some difficulties which often arise in high contrast inverse problems due to the occurence of strong nonlinearities.

As our third linearized operator, we introduce the linearized forward operator $T'_{jk}[\phi]$ by putting

$$T'_{jk}[\phi] : \Phi \longrightarrow Z_j \quad , \quad T'_{jk}[\phi]\delta\phi \;=\; R'_{jk}\,[\Lambda(\phi)]\,\Lambda'[\phi]\delta\phi. \tag{41}$$

All three operators $\Lambda'[\phi]$, $R'_{jk}[\epsilon_{\rm s}]$, and $T'_{jk}[\phi]$ are linear.

4.4 A nonlinear Kaczmarz-type approach

The algorithm works in a 'single-step fashion' as follows. Instead of using the data (10) for all sources and all frequencies simultaneously, we only use the data for one source and frequency at a time while updating the linearized residual operator after each determination of the corresponding incremental correction $\delta\phi$. So, in each step we will look for a solution of the equation

$$T'_{jk}[\phi]\delta\phi_{jk} \;=\; -T_{jk}(\phi) \tag{42}$$

for a given source index $j = 1, \ldots, p$ and a given frequency index $k = 1, \ldots, K$. After correcting ϕ by

$$\phi \longrightarrow \phi + \delta\phi_{jk}, \tag{43}$$

we use the updated residual equation (42) to compute the next correction $\delta\phi_{j'k'}$. Doing this for one equation after the other, until each of the sources q_j and each of the frequencies f_k has been considered exactly once, will yield one complete sweep of the algorithm. This procedure is similar to the Kaczmarz method for solving linear systems, or the algebraic reconstruction technique (ART) in x-ray tomography [32] and the simultaneous iterative reconstruction technique (SIRT) as presented in [16]. Related approaches have also been employed in ultrasound tomography by Natterer and Wübbeling [33], in more general bilinear inverse problems by Natterer [34], in optical tomography by Dorn [17], and in 3D-electromagnetic induction tomography (EMIT) by Dorn *et alii* [18].

4.5 The minimization problem

Let us assume now that we are given a level set function $\phi^{(n)}(x)$ and a scattering potential $\epsilon_{\mathrm{s}}^{(n)}(x)$ such that $(\Omega^{(n)}, \epsilon_{\mathrm{b}} + \epsilon_{\mathrm{s}}^{(n)}, \phi^{(n)})$ forms an admissible triple in the sense of definition 3.3. Using a data set $\tilde{G}_{jk}$ corresponding to the fixed source position q_j and the frequency f_k, we want to find an update $\delta\phi^{(n)}$ to $\phi^{(n)}$ such that for the admissible triple

$$\begin{aligned} &\left(\Omega^{(n+1)}, \epsilon_{\mathrm{b}} + \epsilon_{\mathrm{s}}^{(n+1)}, \phi^{(n+1)}\right) := \\ &\left(\Omega[\phi^{(n)} + \delta\phi^{(n)}], \epsilon_{\mathrm{b}} + \Lambda(\phi^{(n)} + \delta\phi^{(n)}), \phi^{(n)} + \delta\phi^{(n)}\right) \end{aligned} \tag{44}$$

the residuals in the data corresponding to this source and this frequency vanish

$$T_{jk}(\phi^{(n+1)}) \;=\; T_{jk}(\phi^{(n)} + \delta\phi^{(n)}) \;=\; 0. \tag{45}$$

Applying a Newton-type approach, we get from (45) a correction $\delta\phi^{(n)}$ for $\phi^{(n)}$ by solving

$$T'_{jk}[\phi^{(n)}]\delta\phi^{(n)} \;=\; -\,T_{jk}(\phi^{(n)}) \;=\; -\,\left(M_j u_{jk} - \tilde{G}_{jk}\right) \tag{46}$$

where u_{jk} satisfies (31) with $\epsilon_{\mathrm{s}} = \Lambda(\phi^{(n)})$

$$\Delta u_{jk} \;+\; \left[a_k(\epsilon_{\mathrm{b}} + \Lambda(\phi^{(n)}))(x) + ib\sigma(x)\right] u_{jk} \;=\; q_j(x) \tag{47}$$

and

$$\epsilon_{\mathrm{b}}(x) + \Lambda(\phi^{(n)})(x) \;=\; \begin{cases} \hat{\epsilon} & , \quad x \in \Omega[\phi^{(n)}] \\ \epsilon_{\mathrm{b}}(x) & , \quad x \in \mathbb{R}^2 \setminus \Omega[\phi^{(n)}]. \end{cases} \tag{48}$$

Since we have only few data given for one source and one frequency, equation (46) usually will have many solutions (in the absence of noise), such that we have to pick

one according to some criterion. We choose to take that solution which minimizes the energy norm of $\delta\phi^{(n)}$

$$\text{Min } \|\delta\phi^{(n)}\|_2 \quad \text{subject to} \quad T'_{jk}(\phi^{(n)})\delta\phi^{(n)} \;=\; -\left(M_j u_{jk} - \tilde{G}_{jk}\right). \tag{49}$$

This solution can be formulated explicitly. It is

$$\delta\phi^{(n)}_{\text{MN}} \;=\; -\,T'_{jk}[\phi^{(n)}]^* \left(T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^*\right)^{-1} \left(M_j u_{jk} - \tilde{G}_{jk}\right), \tag{50}$$

where $T'_{jk}[\phi^{(n)}]^*$ denotes the adjoint operator to $T'_{jk}[\phi^{(n)}]$.

4.6 The adjoint linearized operators

In order to calculate the minimal norm solution (50), we will need practically useful expressions for the adjoints of the linearized operators of section (4.3). We will present such expressions in this section. The calculation of the actions of these operators will typically require to solve an adjoint Helmholtz problem. This explains the name 'adjoint field method' of the inversion method employed here.

To start with, a simple calculation gives us the following theorem.

Theorem 4.1 *The adjoint operator $\Lambda'[\phi]^*$ which corresponds to the linearized operator $\Lambda'[\phi]$ is given by*

$$\Lambda'[\phi]^* : F \longrightarrow \Phi \quad , \quad \left(\Lambda'[\phi]^*\delta\epsilon_{\text{s}}\right)(x) \;=\; -\,[\hat{\epsilon} - \epsilon_{\text{b}}(x)]\,\frac{\delta\epsilon_{\text{s}}(x)}{|\nabla\phi(x)|}C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x). \tag{51}$$

The next theorem describes the adjoint operator $R'_{jk}[\epsilon_{\text{s}}]^*$ which corresponds to $R'_{jk}[\epsilon_{\text{s}}]$. Its proof is analogous to the proof of Theorem 4.1 in the appendix, or to the proof given in a similar situation in Dorn *et al.* [18], and is therefore omitted here.

Theorem 4.2 *Let $\zeta = (\zeta_1, \ldots, \zeta_{D_j})^T \in Z_j$ and let x_{jd}, $d = 1, \ldots, D_j$ be the detector positions corresponding to q_j. Then the action of the adjoint operator $R'_{jk}[\epsilon_{\text{s}}]^*$ on ζ is given by*

$$R'_{jk}[\epsilon_{\text{s}}]^*\zeta \;=\; -\,\frac{1}{a_k}\mathcal{R}e\left(\overline{u_{jk}z_{jk}}\right)\,\chi_{B_r(0)} \tag{52}$$

where u_{jk} solves

$$\Delta u_{jk} \;+\; \kappa_k(x)u_{jk} \;=\; q_j(x), \tag{53}$$

and z_{jk} solves the 'adjoint equation'

$$\Delta z_{jk} \;+\; \kappa_k(x)z_{jk} \;=\; \sum_{d=1}^{D_j}\bar{\zeta}_d\delta(x - x_{jd}) \tag{54}$$

with

$$\kappa_k(x) \;=\; a_k[\epsilon_{\text{b}}(x) + \epsilon_{\text{s}}(x)] \;+\; ib_k\sigma(x) \tag{55}$$

and a_k, b_k defined as in (7).

Finally, by combining theorems 4.1 and 4.2, we get an expression for the adjoint operator $T'_{jk}[\phi]^*$ which corresponds to the linearized forward operator $T'_{jk}[\phi]$. It is described in the following theorem.

Theorem 4.3 *Let $\zeta = (\zeta_1, \ldots, \zeta_{D_j})^T \in Z_j$ and let x_{jd}, $d = 1, \ldots, D_j$ be the detector positions corresponding to q_j. Then the adjoint operator $T'_{jk}[\epsilon_{\mathrm{s}}]^*$ acts on ζ in the following way*

$$\begin{aligned} T'_{jk}[\phi]^*\zeta &= \Lambda'[\phi]^* R'_{jk}[\Lambda(\phi)]^*\zeta \\ &= \frac{[\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)]}{a_k|\nabla\phi(x)|}\, \mathcal{R}e\,(\overline{u_{jk}z_{jk}})\; C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x), \end{aligned} \tag{56}$$

where u_{jk} solves (53) and z_{jk} solves (54) with ϵ_{s} replaced by $\Lambda(\phi)$.

4.7 The operators $T'_{jk}T'^*_{jk}$

Let us consider the operator

$$C^{(n)}_{jk} := T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^* \tag{57}$$

in (50) more closely. Using (41) it gets the form

$$C^{(n)}_{jk} = R'_{jk}[\Lambda(\phi^{(n)})]\,\Lambda'[\phi^{(n)}]\,\Lambda'[\phi^{(n)}]^*\,R'_{jk}[\Lambda(\phi^{(n)})]^*. \tag{58}$$

With (37), (51) we see that, due to the operator $\Lambda'[\phi^{(n)}]\,\Lambda'[\phi^{(n)}]^*$ in (58), $C^{(n)}_{jk}$ maps first from the data space to functions in F or Φ which are supported on $B_\rho(\Gamma)$, and then back to the data space. In a discretized setting, it might happen that for a coarse mesh (and a small ρ) the number of pixels representing $B_\rho(\Gamma)$ becomes close to (or even smaller than) the number of data points. This observation lets us expect that the inversion of $C^{(n)}_{jk}$ in (50) will be highly ill-conditioned and unstable. This is confirmed by our numerical experiments so far.

Therefore, we will regularize the inversion of $C^{(n)}_{jk}$ and the action of its inverse on the right hand side of (46).

4.8 Regularization of $T'_{jk}T'^*_{jk}$

A standard way of regularization is the Tychonov-Phillips regularization scheme which amounts to replacing the operator $T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^*$ in (50) by the operator

$$T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^* + \lambda I \tag{59}$$

with some suitably chosen regularization parameter $\lambda > 0$. A small λ means little regularization, whereas in the case of very noisy data we might wish to use a very

large λ such that (59) is dominated by the term λI and we can approximate it further by a simple multiplication with the regularization parameter λ.

However, motivated by the above mentioned observations, we choose a different form of regularization. First, we add a Tychonov-Phillips term to $\Lambda'[\phi^{(n)}]\,\Lambda'[\phi^{(n)}]^*$ such that the right hand side of (58) becomes

$$C_{jk}^{(n)} \;\approx\; R'_{jk}[\Lambda(\phi^{(n)})]\,\left(\Lambda'[\phi^{(n)}]\,\Lambda'[\phi^{(n)}]^* + \lambda I\right)\,R'_{jk}[\Lambda(\phi^{(n)})]^*. \tag{60}$$

Now, using a very large regularization parameter λ, we approximate (60) further by

$$C_{jk}^{(n)} \;\approx\; \lambda\, R'_{jk}[\Lambda(\phi^{(n)})]\,R'_{jk}[\Lambda(\phi^{(n)})]^*. \tag{61}$$

Since calculating the operator (61) in each step of the inversion routine is still very time-consuming, we approximate this operator further by replacing the argument $\Lambda(\phi^{(n)})$ by the background scattering potential which is zero. Therefore, we end up with the following approximation for $C_{jk}^{(n)}$

$$C_{jk}^{(n)} \;\approx\; \hat{C}_{jk} \;:=\; R'_{jk}[0]\,R'_{jk}[0]^* \quad \text{for all} \quad n = 1, 2, \ldots \tag{62}$$

The multiplicator λ is neglected in (62) since it becomes part of the relaxation parameter in our inversion scheme. We see that we have replaced in (62) the highly ill-conditioned and difficult to calculate operator $C_{jk}^{(n)}$ by a much better conditioned operator $\hat{C}_{jk}$ which has to be computed only once and which can be precalculated before starting the inversion routine.

The next theorem tells us how to practically compute the operator $\hat{C}_{jk}$ for a given background permittivity ϵ_{b}.

Theorem 4.4 *Let us assume that we are given a background permittivity distribution ϵ_{b}, a finite set of sources q_j, $j = 1, \ldots, p$, and for each of these sources a finite set of receiver positions x_{jd}, $d = 1, \ldots, D_j$. We apply each of the sources with K different frequencies f_k, $k = 1, \ldots, K$. The operators $\hat{C}_{jk}$, $j = 1, \ldots, p$, $k = 1, \ldots, K$, are then described by $D_j \times D_j$ matrices of the form*

$$\hat{C}_{jk} \;=\; \left\{ \int_{B_r(0)} \overline{u_{jk}(x)\varphi_{dk}(x)}\, u_{jk}(x)\varphi_{lk}(x) \right\}_{\substack{l = 1, \ldots, D_j \\ d = 1, \ldots, D_j}}, \tag{63}$$

where $\varphi_{\nu k}$ solves

$$\Delta\varphi_{\nu k} \;+\; (a_k\epsilon_{\mathrm{b}} + ib_k\sigma)\,\varphi_{\nu k} \;=\; \delta(x - x_{j\nu}) \tag{64}$$

and u_{jk} solves

$$\Delta u_{jk} \;+\; (a_k\epsilon_{\mathrm{b}} + ib_k\sigma)\,u_{jk} \;=\; q_j. \tag{65}$$

The index ν in (64) can stand for a receiver index d or l. In (63), the receiver index l is the row index, and the receiver index d is the column index of $\hat{C}_{jk}$.

The proof of this theorem is similar to the proof of Theorem 5.2 given in the appendix such that we omit it here.

4.9 Updating the level set function

In order to calculate a correction $\delta\phi_{\mathrm{MN}}^{(n)}$ by (50) we have to apply the operator $T'_{jk}[\phi^{(n)}]^*$ to the vector

$$\zeta := \hat{C}_{jk}^{-1}\left(M_j u_{jk} - \tilde{G}_{jk}\right). \tag{66}$$

An explicit formula for $T'_{jk}[\phi^{(n)}]$ was already given in (56)

$$T'_{jk}[\phi^{(n)}]^*\zeta = \frac{[\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)]}{a_k|\nabla\phi^{(n)}(x)|}\,\mathcal{R}e\left(\overline{u_{jk}(x)z_{jk}(x)}\right)\,C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x), \tag{67}$$

where u_{jk} and z_{jk} solve (53)- (55) with ϵ_{s} replaced by $\Lambda(\phi^{(n)})$.

To stabilize the reconstruction scheme, we replace the term $|\nabla\phi(x)|$ in (67) by some constant c_1. Doing so we avoid dividing by numerical derivatives which might cause instabilities due to numerical noise and roundoff errors. This is justified as long as $|\nabla\phi(x)|$ does not vary too much along the boundary. It turns out that the updates we apply in our numerical examples to the level set functions usually justify this assumption. In cases with limited view and very noisy data, however, we will apply an additional 'smoothing procedure' (which is described in section 4.10) to the level set functions near the boundary after each update in order to guarantee the necessary regularity for the succeeding steps.

With these modifications, (50) yields the following *update formula for the level set function*

$$\delta\hat{\phi}^{(n)}(x) = -\frac{\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)}{c_1 a_k}\mathcal{R}e\left(\overline{u_{jk}(x)z_{jk}(x)}\right)\,C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x) \tag{68}$$

where u_{jk} and z_{jk} solve (53)- (55) with ζ given by (66) and ϵ_{s} replaced by $\Lambda(\phi^{(n)})$.

Notice that, although we did not explicitly impose any regularity constraints on the updates (68), they are in the range of $R'_{jk}[\Lambda(\phi^{(n)})]^*$ (up to the factor $\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)$) which implicitly gives us some information about the regularity we can expect. Our numerical experiments so far indicate that the degree of regularity which is achieved by applying (68) is typically sufficient 'for practical purposes' in those situations where the data are not too noisy and where we have suitably arranged receiver positions all around the obstacles. (This is the 'full view' situation.)

However, in cases of noisy and limited-view data, the resulting boundaries look rough and fuzzy, in particular when high-frequency data are used for the reconstruction. In these situations, we can improve the results by applying some additional regularization on ϕ. A possible way of doing so is to filter the level set function after

each update in order to smooth it locally. An example for such a procedure is 'curve shortening by diffusion', which is briefly described at the end of section 4.10.

We mention that an interesting (and from a mathematical point of view more satisfactory) alternative to this procedure would be to apply some additional regularity constraints already in the derivation of (68), such that we would not have to worry at all about the smoothness of the resulting level set functions. We will investigate possible ways of doing so in our future research.

4.10 Implementation: The levelART algorithm

In brief algorithmic form, the nonlinear Kaczmarz-type method for shape reconstruction using level sets (which we call for short 'levelART' because of its above mentioned similarity to the 'ART' algorithm in x-ray tomography) can be written in the following way.

Preparation step.

- Calculate $\hat{C}_{jk}$ and
$$D_{jk} = \hat{C}_{jk}^{-1} \tag{69}$$
according to (63) for each source q_j, $j = 1, \ldots, p$, and each frequency f_k, $k = 1, \ldots, K$, and store in memory for later use.

- Build groups of frequencies $G_m = \{f_1, \ldots, f_{K_m}\}$, $m = 1, \ldots, M$.

Initialization.

$n = 0$;
$(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$ given from STAF.

Reconstruction loop.

FOR $m = 1 : M$ march over frequency groups G_m
  FOR $i = 1 : I_m$ perform I_m sweeps for frequency group G_m
    FOR $k = 1 : K_m$ march over frequencies in G_m
      FOR $j = 1 : p$ march over sources q_j for each frequency
      $\zeta_{jk} = D_{jk}(M_j u_{jk} - \tilde{G}_{jk})$; u_{jk} solves (53) with $\epsilon^{(n)}$
      $\delta\phi_{jk} = -\frac{\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)}{a_k} \mathcal{R}e(\overline{u_{jk} z_{jk}}) \chi_{B_\rho(\Gamma)}$; z_{jk} solves (54) with $\epsilon^{(n)}$ and ζ_{jk}
      END
    $\delta\phi^{(n)}(x) = \sum_{j=1}^{p} \delta\phi_{jk}(x)$;
    $\phi^{(n+1)} = C_{\mathrm{LS}}^{(n)}(\phi^{(n)} + \eta \frac{C_\rho(\Gamma)}{c_1} \delta\phi^{(n)})$; update level set function
      Optional step: 'curve shortening' by diffusion. See separate chart.
    $\epsilon^{(n+1)} = \epsilon_{\mathrm{b}} + \Lambda(\phi^{(n+1)})$; $n = n + 1$; Reinitialization $n \to n + 1$
    END
  END alternatively, some stopping criteria can be used here
END
$(\Omega^{(N)}, \epsilon^{(N)}, \phi^{(N)}) = (\Omega[\phi^{(n)}], \epsilon_{\mathrm{b}} + \Lambda(\phi^{(n)}), \phi^{(n)})$; Final reconstruction.

Here, η is a relaxation parameter for the update of the level set function which is determined empirically. The constant $C_\rho(\Gamma)$ could be calculated explicitly for the actual curve $\Gamma^{(n)}$, or it could be approximated by some value corresponding to a simple geometrical object (to give an example, in case of a single circle it would be $C_\rho(\Gamma) = (2\rho)^{-1}$). In our numerical experiments so far, however, it is simply considered as part of η. The same holds true for c_1. The constant ρ is in our numerical experiments chosen between 30-40 cm, which corresponds to 2-3 grid cells. The scaling factor $C_{\mathrm{LS}}^{(n)}$ is determined after each update to keep the global minimum (or maximum) of the level set function at a constant value.

The following smoothing filter on the level set function is optional. We usually apply it when we use noisy high-frequency data for the reconstruction. Especially in the limited-view examples presented in sections 6.3 and 6.4, the application of this filter improves the reconstructions significantly. Smoothing the level set function with this filter has the effect of local curve shortening. Roughness and small scale oscillations in the reconstructed boundaries are smoothed out such that the 'energy' of the reconstructed boundaries is reduced. The filtering step can be described as follows.

Optional step: 'Curve shortening' by diffusion.

Introduce artificial time $\tau \in \mathbb{R}$. $\frac{\partial}{\partial\tau}$ = time derivative, Δ = Laplace operator.
Solve initial value problem (with absorbing boundary conditions) on $B_r(0)$:

$$\begin{aligned}
&\tilde{\phi}(x,0) = \phi^{(n+1)}(x);\\
&\qquad \frac{\partial}{\partial\tau}\tilde{\phi}(x,\tau) = \Delta\tilde{\phi}(x,\tau), \quad \tau \in [0,T_\lambda], \quad x \in B_r(0).\\
&\phi^{(n+1)}(x) = \tilde{\phi}(x,T_\lambda),
\end{aligned}$$

with regularization parameter T_λ.

5 Step 1: A Source-Type Adjoint Field method

For starting our shape reconstruction method using level sets we will need an initial guess $(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$.

Although it is possible just to create an arbitrary initial guess without using any data at all, we believe that it is important for the efficiency and the robustness of such a method to start it with a *good* initial guess. There are several reasons for this. First, when deriving the shape reconstruction method (which we will call for short 'levelART' in the following) we used a perturbation approach which is strictly

justified only when we already have a good first guess available. Moreover, we see in our numerical experiments that starting with a good initial guess stabilizes the shape reconstruction routine, in particular in cases where the data are incomplete and noisy. In addition, finding a good first guess reduces the amount of work which has to be done by the levelART routine itself, such that in the end the combined code will converge much more rapidly than levelART alone would do without a good initialization.

In this section, we will present the second key point of our combined inversion scheme, namely a *fast*, *inexpensive* and *stable* method for finding a very good first guess $(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$ for levelART. This method is designed to share basic features with the levelART algorithm, such that it can be implemented in addition to levelART with almost no extra cost.

5.1 Source-type methods

In the framework of inverse scattering problems, the method we propose here can be considered as a 'source-type inversion method'. Roughly speaking, the general idea of source-type reconstruction methods in inverse scattering is to split a given nonlinear inverse scattering problem into two subproblems. The first one is linear, and tries to recover a virtual 'equivalent source' in the medium which would be able to fit the data if applied with the known background distribution. This equivalent source is related to the unknown scattering potential by a nonlinear 'constitutive' relation. Therefore, in the second part of the algorithm, a nonlinear inverse problem has to be solved to derive the scattering potential from the recovered equivalent source distribution.

This idea is not at all new. It has been applied for example in the Source-Type Integral Equation (STIE) method of Habashy *et al* [21], or in the method presented by Chew *et al.* in [9]. More recently, similar ideas have been applied by Abdullah *et al.* [1], Caorsi *et al.* [6], and van den Berg *et al.* [46, 47].

All of these approaches have in common that they use the source-type method as a stand-alone inversion scheme. Such a method has the advantage that it is not as sensitive to strong nonlinearities in the inverse problem as for example perturbation methods or the Born or Rytov approximation are [15, 20, 27].

On the other hand, interpreting the inverse scattering problem as an inverse source problem is not without drawbacks. For example, the existence of so-called 'non-radiating sources' or 'invisible sources' gives rise to a nonuniqueness in the inverse source problem, which is difficult to deal with when solving the nonlinear part [1, 14, 21]. Moreover, it is not clear at all how to combine properly the information corresponding to different experiments, since each experiment creates its own 'equiv-

alent sources' and its own 'invisible sources'. For more information about possible applications, advantages and drawbacks of the source-type scheme as a stand-alone inversion tool we refer to [1, 5, 6, 9, 14, 21, 46, 47].

Our approach is different from those mentioned above. We only want to find a good *approximation* to the scattering potential, and a corresponding *initial level set function* suitable to start the shape reconstruction routine. Moreover, we can make use of our prior information about the permittivity distribution. This will allow us to circumvent most of the problems of source-type schemes which have been mentioned above.

We will now describe this method, which we will call the Source-Type Adjoint Field (STAF) method, in more details.

5.2 Solving the inverse scattering problem

Consider the inverse scattering problem formulated in section 3.1. The decomposition (13), (14) reads

$$\tilde{\epsilon}(x) \;=\; \epsilon_{\mathrm{b}}(x) + \tilde{\epsilon}_{\mathrm{s}}(x) \tag{70}$$

with some (known) background distribution ϵ_{b} and the (unknown) scattering potential $\tilde{\epsilon}_{\mathrm{s}}$ having compact support, $\mathrm{supp}(\tilde{\epsilon}_{\mathrm{s}}) \subset\subset B_r(0)$. The goal is to find $\tilde{\epsilon}_{\mathrm{s}}(x)$ from the data (9).

We already mentioned above that we actually will *not* recover the entire function $\tilde{\epsilon}_{\mathrm{s}}(x)$ from the data $\tilde{G}_{jk}$ in this preprocessing step. All we will find is 1.) A very good first guess for the scattering potential $\epsilon_{\mathrm{s}}^{(0)}$ which will be part of the initializing triple $(\Omega^{(0)}, \epsilon_{\mathrm{b}} + \epsilon_{\mathrm{s}}^{(0)}, \phi^{(0)})$, and 2.) A corresponding level set function $\phi^{(0)}$. We will freely make use of the prior information resulting from the knowledge of $\hat{\epsilon}$ inside the estimated scatterer $\Omega^{(0)}$. However, our numerical results so far indicate that the method proposed here -if suitably adapted- can actually be used to find, in addition to $\Omega^{(0)}$ and $\phi^{(0)}$, also a good first estimate $\hat{\epsilon}^{(0)}$ for the contrast $\hat{\epsilon}$. This will be used in our future work to start a reconstruction method which tries to recover $\tilde{\Omega}$ and $\hat{\epsilon}$ *simultaneously* from the given data.

For a fixed frequency f_k and a source q_j, let $\tilde{u}_{jk}$ be the solution of

$$\Delta \tilde{u}_{jk} \;+\; (a_k(\epsilon_{\mathrm{b}} + \tilde{\epsilon}_{\mathrm{s}})(x) \;+\; ib\sigma(x))\, \tilde{u}_{jk} \;=\; q_j(x), \tag{71}$$

and let u_{jk} be the solution of the 'unperturbed' equation

$$\Delta u_{jk} \;+\; (a_k \epsilon_{\mathrm{b}}(x) \;+\; ib\sigma(x))\, u_{jk} \;=\; q_j(x). \tag{72}$$

Define

$$\tilde{v}_{jk} \;:=\; u_{jk} - \tilde{u}_{jk}. \tag{73}$$

Subtraction of (72) from (71) shows that $\tilde{v}_{jk}$ solves

$$\Delta\tilde{v}_{jk} \,+\, (a_k\epsilon_{\rm b}(x) \,+\, ib\sigma(x))\,\tilde{v}_{jk} \,=\, \tilde{Q}^s_{jk}(x), \tag{74}$$

where the 'scattering source' $\tilde{Q}^s_{jk}(x)$ is defined as

$$\tilde{Q}^s_{jk}(x) \,=\, a_k\tilde{\epsilon}_{\rm s}(x)\tilde{u}_{jk}(x). \tag{75}$$

We introduce a 'source type' forward operator A_{jk} by putting

$$A_{jk} : Y \longrightarrow Z_j \quad , \quad A_{jk}Q^s_{jk} \,=\, M_j v_{jk} \tag{76}$$

where M_j is the measurement operator defined in (8), and v_{jk} solves

$$\Delta v_{jk} \,+\, (a_k\epsilon_{\rm b}(x) \,+\, ib\sigma(x))\,v_{jk} \,=\, Q^s_{jk}(x). \tag{77}$$

The operator A_{jk} is linear.

Let us assume now that we apply the 'correct' scattering source $\tilde{Q}^s_{jk}(x)$ defined by (75) as argument of A_{jk}. Then we know from (9), (29), (30) that

$$A_{jk}\tilde{Q}_{jk} \,=\, M_j\tilde{v}_{jk} \,=\, M_j(u_{jk} - \tilde{u}_{jk}) \,=\, M_j u_{jk} - \tilde{G}_{jk} \,=\, R_{jk}(0). \tag{78}$$

The vectors $R_{jk}(0)$ are easily computed by solving a forward problem on the background distribution (72). Therefore, all we have to do to get back the scattering source $\tilde{Q}^s_{jk}$ from the data $\tilde{G}_{jk}$ is to solve (78) for $\tilde{Q}^s_{jk}$. Doing so amounts to solving an ill-posed but *linear* inverse problem.

Once we have recovered $\tilde{Q}^s_{jk}(x)$, we want to get back $\tilde{\epsilon}_{\rm s}(x)$ out of it by using the constitutive relation (75). This second part of the inversion scheme can be interpreted as solving a *nonlinear* inverse problem since $\tilde{u}_{jk}(x)$ depends on $\tilde{\epsilon}_{\rm s}(x)$.

Notice that $\tilde{Q}^s_{jk}(x)$ varies with different sources and frequencies, but that $\tilde{\epsilon}_{\rm s}(x)$ is the same for all sources and all frequencies (if we neglect dispersion). We will make use of this observation when we try to solve the nonlinear part (75). In the following, we describe the method which we will use to recover the scattering source $\tilde{Q}^s_{jk}(x)$ from a given data set $\tilde{G}_{jk}$ for a fixed source q_j and a fixed frequency f_k.

5.3 Looking for a scattering source

Since for a fixed (primary) source position and a fixed frequency we have only few data given to recover $\tilde{Q}^s_{jk}$, and since we have to take into account that also 'non-radiating' and 'invisible' sources have been generated in the experiment, we assume that there will be many solutions (in absence of noise) of (78). To pick one we are looking for the solution with minimal norm

$$\text{Min}\ \|Q^s_{jk}\|_Y \quad \text{subject to} \quad A_{jk}Q^s_{jk} \,=\, R_{jk}(0). \tag{79}$$

It is given by

$$Q^s_{jk,\mathrm{MN}} \;=\; A^*_{jk}\left(A_{jk}A^*_{jk}\right)^{-1} R_{jk}(0), \tag{80}$$

where A^*_{jk} denotes the adjoint operator to A_{jk}.

The following theorem, which is proven in the appendix, tells us how to calculate the action of A^*_{jk} on a vector $\zeta \in Z_j$ in an efficient way.

Theorem 5.1 *Let $\zeta = (\zeta_1, \ldots, \zeta_{D_j})^T \in Z_j$ and let x_{jd}, $d = 1, \ldots, D_j$ be the detector positions corresponding to the source q_j. Then, $A^*_{jk}\zeta$ is given by*

$$A^*_{jk}\zeta \;=\; \overline{z_{jk}}\chi_{B_r(0)}, \tag{81}$$

where z_{jk} solves

$$\Delta z_{jk} \;+\; (a_k\epsilon_{\mathrm{b}} + ib_k\sigma)\, z_{jk} \;=\; \sum_{d=1}^{D_j} \overline{\zeta_d}\delta(x - x_{jd}). \tag{82}$$

Corollary 5.1 *Let $\varphi_{\nu k}$ solve*

$$\Delta\varphi_{\nu k} \;+\; (a_k\epsilon_{\mathrm{b}} + ib_k\sigma)\,\varphi_{\nu k} \;=\; \delta(x - x_{j\nu}). \tag{83}$$

Then, we can write (81) in the alternative form

$$\left(A^*_{jk}\zeta\right)(x) \;=\; \overline{\sum_{d=1}^{D_j} \overline{\zeta_d}\varphi_{dk}(x)}\chi_{B_r(0)}(x) \;=\; \sum_{d=1}^{D_j} \zeta_d\overline{\varphi_{dk}(x)}\chi_{B_r(0)}(x). \tag{84}$$

The next theorem, which is proven in the appendix, gives an explicit expression for the operators $A_{jk}A^*_{jk}$.

Theorem 5.2 *Let us assume that we are given a background permittivity distribution ϵ_{b} and a finite set of receiver positions x_{jd}, $d = 1, \ldots, D_j$. The operators $A_{jk}A^*_{jk}$, $j = 1, \ldots, p$, $k = 1, \ldots, K$, are then described by $D_j \times D_j$ matrices of the form*

$$A_{jk}A^*_{jk} \;=\; \left\{\int_{B_r(0)} \overline{\varphi_{dk}(x)}\varphi_{lk}(x)\right\}_{\substack{l\,=\,1,\ldots,D_j \\ d\,=\,1,\ldots,D_j}}, \tag{85}$$

where $\varphi_{\nu k}$ solves (83) and the index ν can stand for a receiver index d or l. In (85), the receiver index l is the row index, and the receiver index d is the column index of $\hat{C}_{jk}$.

Remark. Notice that (85) does not depend on the sources q_j, but only on the arrangement of the detectors and on the background permittivity distribution ϵ_{b}. The operators $A_{jk}A^*_{jk}$ can be precomputed before starting the inversion routine,

and this has to be done only once for each frequency as long as we use the same arrangement of detectors for all sources q_j, $j = 1, \ldots, p$.

In the case of noisy data we will invert $A_{jk}A_{jk}^* + \lambda I$ instead of $A_{jk}A_{jk}^*$ in (80) with a suitably chosen regularization parameter $\lambda > 0$. This amounts to applying Tychonov-Phillips regularization.

5.4 Recovery of the scattering potential

After we have found a scattering source Q_{jk}^s which satisfies (79), we want to use the constitutive relation

$$Q_{jk}^s(x) \;=\; a_k \tilde{\epsilon}_{\mathrm{s}}(x) \tilde{u}_{jk}(x), \tag{86}$$

which holds for the 'correct' scattering source $\tilde{Q}_{jk}^s$ according to (75), to find an approximation for $\tilde{\epsilon}_{\mathrm{s}}(x)$.

Let $\tilde{u}_{jk}$ be a solution of (71) and u_{jk} a solution of (72). We decompose Q_{jk}^s, $\tilde{u}_{jk}$, u_{jk} and $\tilde{\epsilon}_{\mathrm{s}}$ into amplitude and phase

$$Q_{jk}^s(x) \;=\; |Q_{jk}^s(x)|\, e^{ir(x)}, \qquad \tilde{u}_{jk} \;=\; |\tilde{u}_{jk}|\, e^{i\tilde{s}(x)}, \tag{87}$$

$$u_{jk} \;=\; |u_{jk}|\, e^{is(x)}, \qquad \epsilon_{\mathrm{s}}(x) \;=\; |\epsilon_{\mathrm{s}}(x)|\, e^{it(x)}, \tag{88}$$

where we have omitted the subscripts jk in the argument functions r, $\tilde{s}$, s, and t for simplicity in the notation. Making use of the fact that $\tilde{\epsilon}_{\mathrm{s}}(x) \in \mathbb{R}$ we see that

$$t(x) \;\in\; \{0, \pi\} \quad \text{for all } x \in \mathbb{R}^2. \tag{89}$$

With (87),(88) equation (79) decomposes into two equations, one for the amplitude and one for the phase. They are

$$|Q_{jk}^s(x)| \;=\; a_k\, |\tilde{\epsilon}_{\mathrm{s}}(x)|\, |\tilde{u}_{jk}|, \tag{90}$$

$$r(x) \;=\; \tilde{s}(x) \;+\; t(x). \tag{91}$$

The observation in our numerical experiments is that, although $s(x)$ and $\tilde{s}(x)$ might be quite different from each other for large perturbations $\tilde{\epsilon}_{\mathrm{s}}(x)$, the amplitudes $|u_{jk}(x)|$ and $|\tilde{u}_{jk}(x)|$ most often do not differ too much from each other in the scattering region. Therefore, in our applications it is a reasonable approximation to assume that

$$|\tilde{u}_{jk}(x)| \;\approx\; |u_{jk}(x)| \quad \text{in } B_r(0). \tag{92}$$

With this approximation, (90) yields the following estimate for $|\tilde{\epsilon}_{\mathrm{s}}(x)|$

$$|\tilde{\epsilon}_{\mathrm{s}}^{(jk)}(x)| \;\approx\; \frac{|Q_{jk}^s(x)|}{a_k |u_{jk}(x)|} \quad \text{in } B_r(0). \tag{93}$$

We have added the indices j and k on the left hand side of (93) to indicate that we have used only the data $\tilde{G}_{jk}$ corresponding to source q_j and frequency f_k for its determination.

Notice that the step (92), (93) is *nonlinear* since taking the amplitude of a complex number is a nonlinear operation. Therefore, the approach presented here is quite different from the usual Born approximation which approximates $\tilde{u}_{jk}$ by u_{jk}.

For the purposes of the present paper, the determination of $|\tilde{\epsilon}_{\rm s}(x)|$ is already sufficient in order to get a good first guess for the scattering potential $\epsilon_{\rm s}^{(0)}(x)$ and for the level set function $\phi^{(0)}(x)$, since we can now make use of our prior information about the correct value of $\hat{\epsilon}$ in (12).

We mention, however, that the recovery of the phase $t(x)$ is also possible from (86). (This will be necessary for example when we try to recover $\tilde{\Omega}$ and $\hat{\epsilon}$ simultaneously from the given set of data $\tilde{G}_{jk}$.) We can do this by using equation (91). We already mentioned that the assumption $\tilde{s}(x) \approx s(x)$ might be quite wrong for situations with large scattering potentials $\tilde{\epsilon}_{\rm s}(x)$. However, in our situation we only have to decide whether $t(x)$ is zero or π, which means that we have to determine whether in (91) $r(x) = \tilde{s}(x)$ or $r(x) = \tilde{s}(x) \pm \pi$ is satisfied. Therefore, a reasonable estimate for $t(x)$ is to put

$$t^{(jk)}(x) \;=\; \begin{cases} 0 & , \quad |s(x) - r(x)| < \pi/2, \\ \pi & , \quad \text{elsewhere.} \end{cases} \tag{94}$$

Our numerical experiments so far show that a suitable combination of these estimates resulting from many source positions gives a very good reconstruction of the phase $t(x)$ in $B_r(0)$ even in the situation of limited view and noisy data. We will not need this estimate in the present paper.

5.5 Combining the results from single experiments

We can combine now the estimates $|\tilde{\epsilon}_{\rm s}^{(jk)}(x)|$ from many source positions q_j, $j = 1, \ldots, p$, by putting

$$|\tilde{\epsilon}_{\rm s}^{(k)}(x)| \;=\; \frac{1}{p} \sum_{j=1}^{p} |\tilde{\epsilon}_{\rm s}^{(jk)}(x)|. \tag{95}$$

If we want to take into account also the information corresponding to many frequencies we can do so by putting

$$|\tilde{\epsilon}_{\rm s}(x)| \;\approx\; \frac{1}{pK} \sum_{k=1}^{K} \sum_{j=1}^{p} |\tilde{\epsilon}_{\rm s}^{(jk)}(x)| \tag{96}$$

where the sum is over all frequencies f_k, $k = 1, \ldots, K$, and all source positions q_j, $j = 1, \ldots, p$. In (96) we have neglected dispersion.

A similar strategy can be employed to improve the estimates for the phase $t^{(jk)}(x)$.

5.6 The initial level set function

We are now ready to define the initial triple $(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$.

We assume that we are working in a high contrast situation, such that exactly one of the following conditions is satisfied

$$\hat{\epsilon} - \epsilon_{\mathrm{b}}(x) \gg 0 \quad \text{for all } x \in \tilde{\Omega} \tag{97}$$

$$\hat{\epsilon} - \epsilon_{\mathrm{b}}(x) \ll 0 \quad \text{for all } x \in \tilde{\Omega}. \tag{98}$$

Since we know $\hat{\epsilon}$ and $\epsilon_{\mathrm{b}}(x)$, we know the constant

$$\mathrm{sign}(\tilde{\Omega}) := \begin{cases} 1 & , \text{ if (97) holds} \\ -1 & , \text{ if (98) holds.} \end{cases} \tag{99}$$

Let us assume that we want to use $|\tilde{\epsilon}_{\mathrm{s}}^{(k)}(x)|$ as defined in (95) for a fixed frequency f_k to specify the level set function $\phi^{(0)}$. Choose a threshold value $0 < \gamma_{\mathrm{LS}} < 1$ (in our numerical examples presented in section 6 we use $\gamma_{\mathrm{LS}} = 0.7$) and define

$$\epsilon_{\mathrm{LS}} := \gamma_{\mathrm{LS}} \max_{x \in B_r(0)} |\tilde{\epsilon}_{\mathrm{s}}^{(k)}(x)|. \tag{100}$$

For the level set zero $L_0^{(0)}$ of $\phi^{(0)}$ we require that

$$L_0^{(0)} = \left\{ x \in B_r(0) : \quad |\tilde{\epsilon}_{\mathrm{s}}^{(k)}(x)| = \epsilon_{\mathrm{LS}} \right\}. \tag{101}$$

This means that we want all points of $B_r(0)$ where the reconstruction $|\tilde{\epsilon}_{\mathrm{s}}^{(k)}(x)|$ has exactly the value ϵ_{LS} to be mapped to zero by the level set function $\phi^{(0)}$

$$\phi^{(0)}(x) = 0 \quad \text{for all } x \in L_0^{(0)}. \tag{102}$$

The level set function is now defined as

$$\phi^{(0)}(x) = C_{\mathrm{LS}}^{(0)} \mathrm{sign}(\tilde{\Omega}) \left(\epsilon_{\mathrm{LS}} - |\tilde{\epsilon}_{\mathrm{s}}^{(k)}(x)| \right), \tag{103}$$

where $C_{\mathrm{LS}}^{(0)}$ is some suitably chosen scaling factor. Notice that (102) and (103) are consistent.

The initial scattering domain $\Omega^{(0)}$ and the permittivity $\epsilon^{(0)}$ are defined as

$$\Omega^{(0)} = \Omega[\phi^{(0)}], \qquad \epsilon^{(0)} = \epsilon_{\mathrm{b}} + \Lambda(\phi^{(0)}). \tag{104}$$

Together with $\phi^{(0)}$ they form an admissible triple $(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$.

5.7 Implementation: The STAF algorithm

In brief algorithmic form, the Source Type Adjoint Field (STAF) scheme can be written in the following way.

Preparation step.

- Select a group of frequencies $G_s = \{f_1, \ldots, f_{K_s}\}$ which are used for the STAF reconstruction. Typically, this is just one frequency.
- Calculate $A_{jk}A_{jk}^*$ according to (85) for each frequency $f \in G_s$. The operator $A_{jk}A_{jk}^*$ does not depend on the index j if we use the same detector positions for all sources.
- Calculate

$$B_{jk} = \left(A_{jk}A_{jk}^*\right)^{-1} \quad \text{or} \quad B_{jk} = \left(A_{jk}A_{jk}^* + \lambda I\right)^{-1} \tag{105}$$

  for all $f \in G_s$ and store in memory for later use.

Reconstruction step.

FOR $k = 1 : K_s$

  FOR $j = 1 : p$

    $R_{jk}(0) = M_j u_{jk} - \tilde{G}_{jk}, \qquad u_{jk}$ solves (72)

    $\zeta_{jk} = B_{jk}R_{jk}(0), \qquad B_{jk}$ from (105)

    $Q_{jk}^s = A_{jk}^*\zeta_{jk} = \overline{z_{jk}}\chi_{B_r(0)}, \qquad z_{jk}$ solves (82)

    $|\tilde{\epsilon}_{\mathrm{s}}^{(jk)}(x)| = \frac{|Q_{jk}^s(x)|}{a_k|u_{jk}(x)|}$

  END

END

$|\tilde{\epsilon}_{\mathrm{s}}(x)| = \frac{1}{pK_s}\sum_{k=1}^{K_s}\sum_{j=1}^{p}|\tilde{\epsilon}_{\mathrm{s}}^{(jk)}(x)|, \qquad x \in B_r(0)$

$\phi^{(0)}(x) = C_{\mathrm{LS}}^{(0)}\mathrm{sign}(\tilde{\Omega})(\epsilon_{\mathrm{LS}} - |\tilde{\epsilon}_{\mathrm{s}}(x)|) \qquad$ as in (103)

$(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)}) = (\Omega[\phi^{(0)}], \epsilon_{\mathrm{b}} + \Lambda(\phi^{(0)}), \phi^{(0)})$.

6 Numerical Experiments

6.1 Discretization of the computational domain.

In our numerical experiments, we use a Finite-Differences Frequency Domain (FDFD) code written in MATLAB for solving (1)-(3). The code uses appropriately designed

perfectly matched layers (PML) to avoid reflections at the artificial computational boundaries [36, 37].

The system which results after discretization is solved by a simple Gauss elimination scheme, which is reflected in the implementation shown in sections 4.10 and 5.7. The LR-factorization corresponding to the most recent best guess is used there to calculate the fields for all transmitters and all receivers simultaneously. Therefore, the computational cost for solving all necessary forward and adjoint problems is just one LR-factorization for STAF, and one LR-factorization per update for levelART. If a different solver is used (e.g. GMRES or QMR), then we might find more efficient strategies than those presented in sections 4.10 and 5.7. We mention also that an iterative solver has been developed recently in [25] which is optimized to work on several source distributions simultaneously.

The physical domain is partitioned into 100×100 elementary cells (pixels) in the first numerical example, and into 180×110 elementary cells in the second and third example. Each of these grid cells has a physical size of about 0.14×0.14 m^2, such that the total computational domain in the first example covers an area of 14×14 m^2, and in the other two examples of 15×25 m^2. The eight layers which are closest to the boundaries of the computational domain are used as a PML.

We will refer to the first numerical example as the 'full-view' situation, and to the other two numerical examples as the 'limited-view', 'cross-borehole' or 'geophysical' situations. This terminology is motivated by the source and receiver geometries used, which are as follows.

In the full-view example, we have 64 sources and receivers given which surround the domain of interest. Each source position is at the same time a receiver position and vice versa. The distance of two adjacent sources or receivers from each other is four pixels or about 55 cm. The area enclosed by these sources and receivers has a size of 10×10 m^2.

In the two limited-view examples, 74 sources and receivers are positioned equally spaced in two boreholes. The distance of two adjacent sources or receivers from each other is again 4 pixels or 55 cm, and the distance of the two boreholes from each other is about 10 m.

We mention that, in all of our numerical examples, the regions beyond the source and receiver positions are part of the inversion problem, too. This means, the area which has to be recovered from the data is the whole area situated between the PML boundaries. In some of our numerical experiments, artifacts can be observed developping in the outer areas during the early stages of the reconstruction process.

We apply time-harmonic dipole sources of the form (11) with frequencies of $f = 5$, 10, 15, 20, 25, or 30 MHz. In our examples, this corresponds to wavelengths

in the background medium between 2 meters for $f = 30$ MHz and 13 meters for $f = 5$ MHz. The size of an individual grid cell is chosen such that each of these wavelengths is sampled by at least 16 pixels in order to avoid numerical artifacts due to undersampling.

The data in our numerical examples are generated by running the FDFD forward modelling code on the correct permittivity and conductivity distributions. Using the same forward code for creating the data and for doing the reconstruction is usually called 'inverse crime'. Therefore, to make sure that the situations we model in our experiments are as realistic as possible, we have tested the forward modelling code thoroughly, and add Gaussian noise with signal-to-noise ratios between 10 and 5 dB to the real and imaginary parts of the generated data.

6.2 A full-view example

Our first numerical example tests whether the derived algorithm is able to reconstruct a relatively complicated shape in the ideal situation where sources and receivers completely surround the area of interest. The geometry of this example is shown in Figure 1. The positions of the sources and receivers are indicated by dots in the Figure. The background medium in this example consists of a homogeneous conductivity distribution $\sigma_b = 3.0 \times 10^{-4}$ Siemens/m, and a homogeneous permittivity distribution $\epsilon_{\mathrm{b}} = 20$. Inside the object, the permittivity is $\hat{\epsilon} = 15$, having a moderate contrast to the background distribution.

Notice that an interesting feature of this geometry is the 'hole' in the body of the object which is difficult to reconstruct. We will see that, during the evolution of the permittivity in levelART, the boundaries of the reconstructed domain will split and merge in the attempt to recover this geometry correctly.

First, we test the STAF algorithm in Figure 2 by reconstructing the permittivity $\epsilon^{(0)}$ using only the data $G_s = \{30 \text{ MHz}\}$. These data are noisy with a signal-to-noise ratio (SNR) of 10 dB in the real and the imaginary parts. Compare the result with the upper left image of Figure 3 where we used (noise-free) data with frequency $G_s = \{5 \text{ MHz}\}$.

A comparison of reconstructed permittivities for different frequencies between 5 and 30 MHz shows that -in the ideal situation of sources and receivers completely surrounding the area of interest- the STAF algorithm usually yields already a decent approximation to the shape of the inclusion when applied to the data with the highest frequencies, whereas it yields a reconstruction with decreasing resolution when applied to data corresponding to lower frequencies.

Therefore, it seems to be most efficient to apply the STAF algorithm directly to the highest frequency data, such that we do not need at all any low-frequency

information for the reconstruction. We will see in the following two geophysical examples that this is certainly not true in applications where we have only data with limited view available. In these cases, the use of lower frequency data stabilizes the reconstruction process, and is necessary for preparing the final reconstruction step using the higher frequency data.

We also want to demonstrate the performance of the levelART reconstruction scheme when applied to this geometry. We start the algorithm by using as initial permittivity $\epsilon^{(0)}$ the low resolution STAF reconstruction which is shown in the upper left image of Figure 3, and the corresponding initial level set function $\phi^{(0)}$. No noise is added to the data. Figure 3 shows different stages in the reconstruction process. We first apply levelART with a frequency of 10 MHz to this initial guess and run it for 30 sweeps. Then, we run levelART with 20 MHz for 30 more sweeps, and finally for another 30 sweeps with 30 MHz. The final reconstruction $\epsilon^{(N)}$ is shown in the lower right image of Figure 3. The final level set function $\phi^{(N)}$ corresponding to this reconstruction is displayed in Figure 14.

We see from this example that the shape reconstruction algorithm using level sets is able to split and merge boundaries easily in order to build up relatively complicated geometries. In the present situation, splitting and merging of boundaries was necessary for building the 'hole' in the reconstructed domain.

6.3 A cross-borehole situation with multiple objects

In our second numerical example, we consider a situation which is typical for geophysical applications. Comparable situations occur for example when we wish to monitor pollutant plumes at environmental cleanup sites from cross-borehole EM data.

We assume that we have 74 sources and receivers equidistantly distributed over two boreholes. The distance of the boreholes from each other is 10 meters, and the distance of two adjacent sources or receivers is 55 cm. The area between the two boreholes has to be monitored given the gathered data. The geometry is shown in Figure 4.

The background permittivity distribution in this example consists of four tilted layers with values of $\epsilon_{\mathrm{b}} = 21$ in the top layer, and then continuing downwards with 20 , 19, and again 21 for the deepest layer. The conductivity distribution σ_b is homogeneous with a value of $\sigma_b = 3.0 \times 10^{-4}$ S/m everywhere.

Embedded in this background are three compact inclusions as shown in Figure 4. The permittivity inside these inclusions is $\hat{\epsilon} = 5$, having a high contrast to the background values. The three inclusions are oriented such that there are two 'channels' of background material between them, one of them in the vertical and

one in the horizontal direction. The difficulty in this example is to separate the three inclusions from each other from the limited-view data. In particular, the reconstruction of the vertical channel is critical since we expect that the resolution in the horizontal direction will suffer from the missing data.

Again, we first apply the STAF reconstruction scheme to the data to get a first guess of the permittivity distribution $\epsilon^{(0)}$. Figure 5 shows the result for $G_s = \{30$ MHz$\}$. Comparing this result with the reconstruction for the same frequency in our first numerical example, we conclude that the performance of STAF for high-frequency data is in the limited-view case not as good as in the case where we can use data with full view. We observe that the vertical resolution of the reconstruction is still good, whereas in the horizontal direction severe artifacts build up which reduce the quality of the high-frequency STAF reconstructions as an initial guess for levelART.

Figure 6 shows the corresponding STAF reconstruction using $G_s = \{5$ MHz$\}$. We do not observe any artifacts in this reconstruction which might be caused by the limited view in the data. Therefore, we see that the decreased resolution of STAF using low-frequency data is in this situation compensated by a much higher robustness with respect to missing data. Keeping this in mind, we will typically start our reconstructions in the limited-view geometry by using the STAF result which correspond to (one or more of) the lowest available frequencies.

Starting out from the permittivity $\epsilon^{(0)}$ as shown in Figure 6, and the corresponding level set function $\phi^{(0)}$, we want to use the levelART algorithm in order to calculate a series of shape deformations which transforms the initial shape into the correct permittivity distribution. Figure 6 shows a reconstruction which uses data where the real part and the imaginary part have been contaminated by additive Gaussian noise with a signal-to-noise ratio (SNR) of 10 dB before starting the reconstruction process. Figure 7 shows the same reconstruction scheme, but with an even lower SNR of now 5 dB in the data.

Different strategies are possible for levelART. Which one works best, depends on the specific situation, for example the number and arrangement of sources and detectors, and on the noise level of the data. The strategy we use here (for both SNR values) is the following: First, we apply 20 steps of levelART with a frequency of 15 MHz to the initial guess. The result is shown in the lower left images of Figures 6 and 7. At this stage, the task of splitting the initial object into three subsets is almost completed.

Then, we apply levelART with a combination of three frequencies, namely 20, 25, and 30 MHz. This means that in a given sweep each of these three frequencies is used exactly once in the prescribed order, before starting again with the lowest frequency

(20 MHz) for the new sweep. This final step completes the reconstruction in just a few (about 10) sweeps. The succeeding sweeps do not improve the reconstruction significantly. Moreover, the norms of the residuals approach a constant value as the Figures 8 to 11 show.

Figures 8 to 11 show the evolution of the norms of the residuals during the reconstruction process for different signal-to-noise ratios in the data. The graphs with the symbol '$\circ$' correspond to noiseless data, the graphs for a SNR of 10 dB are indicated by '$*$', and those for a SNR of 5 dB by '$+$'.

Figure 8 shows the norms of the residuals for $f = 15$ MHz during the 20 steps using the data with this frequency. The norms of the residuals decrease in all cases continually during the reconstruction process. Figures 9 to 11 show the evolution of the norms during the final 30 sweeps with the frequencies 20, 25 and 30 MHz. We see that after 10 sweeps the residuals approach some constant value which depends on the noise level of the data. In the lower right image of Figure 6, the reconstruction for a signal-to-noise ratio of 10 dB after completion of these 10 sweeps is shown. The corresponding level set function is shown in Figure 15.

Our experience is that marching over the higher frequencies in the described fashion stabilizes the inversion procedure especially in the limited-view situation considered here. However, so far we do not have any theoretical analysis which supports this observation.

Notice the artifacts which appear in the case of an extremely low SNR of 5 dB in Figure 7. These artifacts remain more or less stable when applying levelART with a constant frequency of 15 MHz, and disappear when marching to the higher frequencies in the succeeding reconstruction step. Notice that we also apply a diffusion ('curve shortening') filter for these higher frequencies, see section 4.10.

We observe again that levelART has no problems in propagating and tracking these multiple artifacts, even when they finally shrink and disappear. Notice also that, similar to the first numerical example, the algorithm splits the original boundary in order to arrive at the three separated inclusions forming the final reconstruction.

6.4 A cross-borehole situation with a single inclusion and variable conductivity

In our third numerical example, we want to test a situation where the conductivity inside the inclusions is different from the given background values. Since we did not take these conductivity changes into account when deriving the reconstruction algorithm, the question arises how much the performance of the reconstruction method will be effected by such changes in the conductivity distribution.

We assume therefore that we know the two values $\hat{\epsilon}$ and $\hat{\sigma}$ of the permittivity and the conductivity inside the inclusions, but both of them are different from the background values. We run the STAF routine as usual, which amounts to treating the conductivity changes simply as an additional form of noise. When applying levelART, we calculate the updates $\delta\phi^{(n)}$ for the level set function $\phi^{(n)}$ in the same way as derived above, but when determining the corresponding updated permittivity distribution $\epsilon^{(n)}$, we update at the same time the conductivity distribution $\sigma^{(n)}$ such that $\sigma(x) = \hat{\sigma}$ where the level set function $\phi^{(n)}$ has negative values. Strictly speaking, we still invert only for the permittivity ϵ, but we make use of the fact that inside the obstacles the conductivity and the permittivity are closely related to each other.

Figure 12 shows the geometry of this example, and Figure 13 shows the results for two different conductivity values $\hat{\sigma} = 1.0 \times 10^{-6}$ S/m and $\hat{\sigma} = 1.0 \times 10^{-2}$ S/m. Notice that these two values differ from each other and from the background value $\sigma_b = 3.0 \times 10^{-4}$ S/m by orders of magnitude! In both cases, Gaussian noise has been added to the real and imaginary parts of the data with a signal-to-noise ratio of 10 dB before starting the reconstruction routine.

As before, we start the reconstruction with the STAF guess corresponding to a frequency of 5 MHz. After only six sweeps of levelART, using the frequencies 15, 20, 25, and 30 MHz one after the other in each sweep, we arrive in both cases at a very good reconstruction of the permittivity distribution and of the conductivity distribution. We conclude that the performance of the reconstruction method (STAF and levelART) is not significantly effected by the changes in the conductivity distribution.

This robustness with respect to changes in the conductivity certainly has its limits. However, the example presented here makes us confident that in practical situations, when the conductivity value $\hat{\sigma}$ inside the obstacles is approximately known and not too large ($< 1.0 \times 10^{-2}$ S/m), the shape reconstruction method will perform stably and reliably and will give us a good reconstruction of the actual permittivity distribution.

7 Summary and future directions

We have presented a stable and efficient two-step shape reconstruction algorithm for EM cross-borehole tomography which uses adjoint fields and level sets. We have shown that this method is able to recover one ore more objects with nontrivial shapes given noisy cross-borehole EM data.

The first step of this combined inversion scheme plays the role of an initializing procedure for the second step, and employs a 'source-type' inversion scheme to deal

with the high nonlinearity in the problem due to the presence of strong scatterers.

Although we believe that the preprocessing routine we propose here will work well in most situations, it can be replaced by any other preprocessing tool which shares the main features of the derived algorithm.

The second step of the inversion routine starts directly with the outcome of this initializing procedure, and continues by using a combination of an 'adjoint field technique' and a level set representation of the shapes until the inversion task is completed. We have shown that using a level set representation in this second step enables us to easily describe and keep track of complicated geometries which arise during the inversion process.

We mention that the FDFD routine, which has been employed in both steps to solve the forward and the adjoint Helmholtz problems, can be replaced by any other more efficient Helmholtz solver which has been tested to work reliably in the given situation.

We have shown in our numerical experiments that the proposed reconstruction scheme performs stably with respect to changes in the conductivity distribution, although these conductivity changes have not been taken into account for the derivation of the scheme. It would be desirable, however, to extend the reconstruction scheme to work simultaneously on the permittivity and the conductivity distribution. This seems to be possible, and we will address this problem in our future work.

We also assume that we know the permittivity values inside the obstacles, and that we only have to recover their shapes and their locations. In our future research, we will investigate the situation where both, the shapes and the permittivity values inside the obstacles, have to be recovered from the given data.

Throughout the paper, we have not clearly specified the degree of regularity which we require for the level set functions ϕ representing the domains Ω. A possible choice would be $\phi \in C_0^1(B_r(0))$ (i.e. continuously differentiable on $B_r(0)$ and zero on $\partial B_r(0)$), which would require some additional regularization in our numerical experiments. We also have introduced in (72) an approximated linearized operator $\Lambda'[\phi]$ motivated by our wish to use convenient inner products. Are there any function spaces which are more useful for our purposes? Will a practically useful reconstruction scheme result if we use different inner products instead of introducing $\Lambda'[\phi]$? To answer both questions, a thorough theoretical analysis has to be done to investigate the implications of using different function spaces for the level set representation.

The main ideas of the reconstruction method presented here are not restricted to a 2D geometry. Therefore, we believe that it is possible to extend the method to a more realistic 3D situation. All what is needed for this is an efficient forward

solver for the 3D system of Maxwell's equations. A forward solver which has been tested for such situations has been presented in [7, 18]. Moreover, applications to situations in medical imaging [2, 11], or in the nondestructive testing of materials [45], seem interesting and possible.

8 Acknowledgments

This work has been supported by a grant from the U. S. Dept. of Energy DE-FG07-97ID3566, by the Center for Subsurface Sensing and Imaging Systems at Northeastern University, by A CAREER Grant from the National Science Foundation MIP-9623721, and by The Army Research Office, Multidisciplinary University Research Initiative Grant No. DAAG55-97-0013.

9 Appendix

9.1 Proof of theorem 5.1

Green's formula for an infinite domain (without boundary terms since $\sigma > 0$) reads for general v, z

$$\int_{\mathbf{R}^2} [\Delta v + \kappa_k v]\, z\, dx \;+\; \int_{\mathbf{R}^2} v(x)(\sum_{d=1}^{D_j} \overline{\zeta_d}\delta(x - x_{jd}))\, dx \tag{106}$$

$$= \int_{\mathbf{R}^2} v\,[\Delta z + \kappa_k z]\; dx \;+\; \sum_{d=1}^{D_j} (\int_{\mathbf{R}^2} v(x)\delta(x - x_{jd})dx)\overline{\zeta_d}$$

where we have used the notation $\kappa_k = a_k\epsilon_{\mathrm{b}} + ib_k\sigma$. Let now v_{jk} be a solution of (77), and z_{jk} a solution of (82). Then the first term on the left hand side of (106) reads

$$\int_{\mathbf{R}^2} Q^s_{jk}(x) z_{jk}(x)\, dx \;=\; \left\langle Q^s_{jk}\,,\, \overline{z_{jk}}\chi_{B_r(0)} \right\rangle_Y , \tag{107}$$

whereas the second term on the right hand side is

$$\langle M_j v_{jk}\,,\, \zeta\rangle_{Z_j} \;=\; \left\langle A_{jk} Q^s_{jk}\,,\, \zeta \right\rangle_{Z_j} . \tag{108}$$

The remaining two terms cancel each other because of (82). Therefore, (106) gets the form

$$\left\langle Q^s_{jk}\,,\, \overline{z_{jk}}\chi_{B_r(0)} \right\rangle_Y \;=\; \left\langle A_{jk} Q^s_{jk}\,,\, \zeta \right\rangle_{Z_j} , \tag{109}$$

which proves the theorem.

9.2 Proof of theorem 5.2

Since $A_{jk}A_{jk}^*$ is a linear operator acting on the finite-dimensional data space, it is sufficient to find the action of $A_{jk}A_{jk}^*$ on each of the basis vectors $e_d = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, $d = 1, \ldots, D_j$, where the '1' is at the d-th position. From (83) we see that

$$\left(A_{jk}^* e_d\right)(x) \;=\; \overline{\varphi_{dk}(x)}\,\chi_{B_r(0)}(x). \tag{110}$$

Application of A_{jk} yields for the l-th component ($l = 1, \ldots, D_j$)

$$\left(A_{jk}A_{jk}^* e_d\right)_l \;=\; v_{jk}(x_{jl}), \tag{111}$$

where v_{jk} is given by

$$v_{jk}(y) \;=\; \int_{B_r(0)} G_k(y,x)\overline{\varphi_{dk}(x)}\,dx \tag{112}$$

and $G_k(y,x)$ is Green's function satisfying

$$\Delta G_k(y,x) \;+\; (a_k\epsilon_{\mathrm{b}} + ib_k\sigma)\,G_k(y,x) \;=\; \delta(y-x). \tag{113}$$

Reciprocity yields $G_k(x_{jl}, x) = \varphi_{lk}(x)$. Therefore, we get from (111), (112)

$$\left(A_{jk}A_{jk}^* e_d\right)_l \;=\; \int_{B_r(0)} \varphi_{lk}(x)\overline{\varphi_{dk}(x)}\,dx, \tag{114}$$

which proves the theorem.

References

[1] Abdullah H and Louis A K 1999 The approximate inverse for solving an inverse scattering problem for acoustic waves in an inhomogeneous medium *Inverse Problems* **15** 1213-1229

[2] Albanese R A, Medina R L and Penn J W 1994 Mathematics, medicine and microwaves *Inverse Problems* **10** 995-1007

[3] Alumbaugh D L and Morrison H F 1995 Theoretical and practical considerations for crosswell electromagnetic tomography assuming a cylindrical geometry *Geophysics* **60** (3) 846-870

[4] Angell T S, Kleinman R E and Roach F G 1987 An inverse transmission problem for the Helmholtz equation *Inverse Problems* **3** 149-80

[5] Angell T S, Jiang X and Kleinman R E 1997 A distributed source method for inverse acoustic scattering *Inverse Problems* **13** 531-46

[6] Caorsi S and Gragnani G L 1999 Inverse-scattering method for dielectric objects based on the reconstruction of the nonmeasurable equivalent current density *Radio Science* **34** (1) 1-8

[7] Champagne N J, II, Berryman J G, Buettner H M, Grant J B and Sharpe R M 1999 A finite-difference frequency-domain code for electromagnetic induction tomography, in the *Proceedings of SAGEEP*, Oakland, CA, March 14–18, 931–940

[8] Chen Y (1997) Inverse scattering via Heisenberg's uncertainty principle *Inverse Problems* **13** 253-282

[9] Chew W C, Wang Y M, Otto G, Lesselier D and Bolomey J Ch 1994 On the inverse source method of solving inverse scattering problems *Inverse Problems* **10** 547-553

[10] Claerbout J F 1976 *Fundamentals of Geophysical Data Processing: With Applications to Petroleum Prospecting* McGraw-Hill, New York

[11] Colton D and Monk P 1994 The detection and monitoring of leukemia using electromagnetic waves: mathematical theory *Inverse Problems* **10** 1235-1251

[12] Colton D L and Kress R 1998 Inverse Acoustic and Electromagnetic Scattering Theory 2nd edn (Berlin: Springer)

[13] Davis J L and Annan A P 1989 Ground penetrating radar for high resolution mapping of soil and rock stratigraphy *Geophys. Prospect.* **37** 531-51

[14] Devaney A J and Sherman G C 1984 Nonuniqueness in inverse source and scattering problems *IEEE Trans. Antennas Propag.* **30** 1034-1037

[15] Devaney A J 1986 Reconstructive tomography with diffracting wavefields *Inverse Problems* **2** 161-83

[16] Dines K A and Lytle R J 1979 Computerized geophysical tomography *Proc. IEEE* **67** 1065–1073

[17] Dorn O 1998 A transport-backtransport method for optical tomography *Inverse Problems* **14** 1107–1130

[18] Dorn O, Bertete-Aguirre H, Berryman J G and Papanicolaou G C 1999 A nonlinear inversion method for 3D electromagnetic imaging using adjoint fields *Inverse Problems* **15** 1523-1558

[19] Fisher E, McMechan G A and Annan A P 1992 Acquisition and processing of wide-aperture ground penetrating radar data *Geophysics* **57** 495-504

[20] Habashy T M, Groom R W, Spies B R 1993 Beyond the Born and Rytov Approximations: A Nonlinear Approach to Electromagnetic Scattering *Journal of Geophysical Research* **98** (B2) 1759-1775

[21] Habashy T M, Oristaglio M L and de Hoop A T 1994 Simultaneous nonlinear reconstruction of two-dimensional permittivity and conductivity *Radio Science* **29** (4) 1101-1118

[22] Hettlich F 1995 Frechet derivatives in inverse obstacle scattering *Inverse Problems* **11** 371-382

[23] Hettlich F and Rundell W 1997 Recovery of the support of a source term in an elliptic differential equation *Inverse Problems* **13** 959-976

[24] Kaup P G, Santosa F and Vogelius M 1996 Method of imaging corrosion damage in thin plates from electrostatic data *Inverse Problems* **12** 279-93

[25] Kilmer M, Miller E L and Rappaport C 1999 QMR-Based Projection Techniques for the Solution of Non-Hermitian Systems with Multiple Right Hand Sides, submitted to *SIAM Journal on Scientific and Statistical Computing*

[26] Kolehmainen V, Arridge S R, Lionheart W R B, Vauhkonen M and Kaipio J P 1999 Recovery of region boundaries of piecewise constant coefficients of an elliptic PDE from boundary data *Inverse Problems* **15** 1375-1391

[27] Ladas K T and Devaney A J 1992 Iterative methods in geophysical diffraction tomography *Inverse Problems* **8** 119-132

[28] Litman A, Lesselier D and Santosa F 1998 Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set *Inverse Problems* **14** 685-706

[29] MATLAB: High-performance numeric computation and visualization software - Reference Guide 1992 *MathWorks* Natick MA

[30] Miller E, Kilmer M and Rappaport C 1999 A New Shape-Based Method for Object Localization and Characterization from Scattered Field Data, submitted to *IEEE Trans. Geoscience and Remote Sensing*

[31] Nachman A 1996 Global uniqueness for a two-dimensional inverse boundary value problem *Ann. Math.* **143** 71-96

[32] Natterer F 1986 *The Mathematics of Computerized Tomography* (Stuttgart: Teubner)

[33] Natterer F and Wübbeling F 1995 A propagation-backpropagation method for ultrasound tomography *Inverse Problems* **11** 1225–1232

[34] Natterer F 1996 Numerical Solution of Bilinear Inverse Problems *Preprints* "Angewandte Mathematik und Informatik" 19/96-N Münster

[35] Osher S and Sethian J 1988 Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations *Journal of Computational Physics* **56** 12-49

[36] Rappaport C M 1996 Interpreting and Improving the PML Absorbing Boundary Condition Using Anisotropic Lossy Mapping of Space *IEEE Transactions on Magnetics* **32** (3) 968-974

[37] Rappaport C M, Kilmer M and Miller E 1999 Accuracy Considerations in Using the PML ABC with FDFD Helmholtz Equation Computation *Journal of Numerical Modelling* in press

[38] Ramm A G 1986 *Scattering by obstacles* (Dordrecht: Reidel)

[39] Rozier C, Lesselier D, Angell D and Kleinman R E 1996 Shape retrieval of an obstacle immersed in shallow water from single frequency fields using a complete family method *Inverse Problems* **13** 487-508

[40] Santosa F 1996 A Level-Set Approach for Inverse Problems Involving Obstacles *ESAIM: Control, Optimization and Calculus of Variations* **1** 17-33

[41] Sethian J A 1990 Numerical algorithms for propagating interfaces: Hamilton-Jacobi equations and conservation laws *J. Diff. Geom.* **31** 131-61

[42] Sethian J A 1999 Level Set Methods and Fast Marching Methods (2nd ed) Cambridge University Press

[43] Souriau L, B Duchene, Lesselier D and Kleinman R E 1996 Modified gradient approach to inverse scattering for binary objects in stratified media *Inverse Problems* **12** 463-481

[44] Spies B P and Habashy T M 1995 Sensitivity analysis of crosswell electromagnetics *Geophysics* **60** 834–845

[45] Special section on the conference on 'inverse problems, control and shape optimization' held 1998 in Carthage (Tunisia) 1999 *Inverse Problems* **15** (1) 1-134

[46] Van den Berg P M and Kleinman R E 1997 A contrast source inversion method *Inverse Problems* **13** 1607-20

[47] Van den Berg P M, van Broekhoven A L and Abubakar A 1999 Extended contrast source inversion *Inverse Problems* **15** 1325-1344

[48] von Hippel A 1953 *Dielectric Materials and Applications* (Wiley: New York) 3-4

[49] Ward S H and Hohmann G W 1987 Electromagnetic theory for geophysical applications, in *Electromagnetic Methods in Applied Geophysics – Theory*, Vol 1, Nabighian M N (ed.), SEG, Tulsa, Oklahoma 131-311

[50] Wilt M J, Morrison H F, Becker A, Tseng H W, Lee K H, Torres-Verdin C, and Alumbaugh D 1995 Crosshole electromagnetic tomography: A new technology for oil field characterization, *The Leading Edge* **14** 173-177

[51] Zhdanov M S, Traynin P and Booker J R 1996 Underground imaging by frequency-domain electromagnetic migration *Geophysics* **61**, 666–682

10 Figures

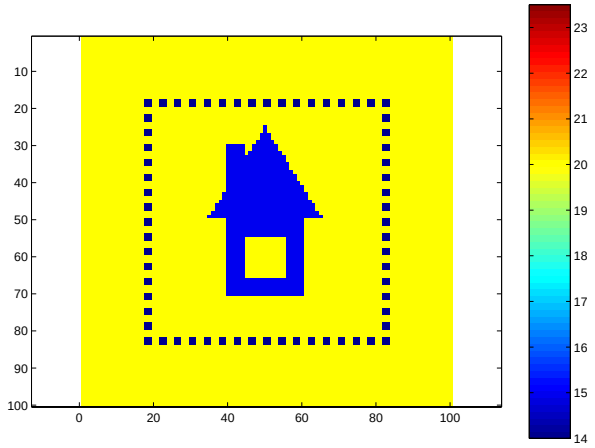


Figure 1: Original object for the example with full view. The dots in the figure indicate the source and receiver positions. The permittivity in the background is $\epsilon_{\mathrm{b}} = 20$, and in the object $\hat{\epsilon} = 15$.

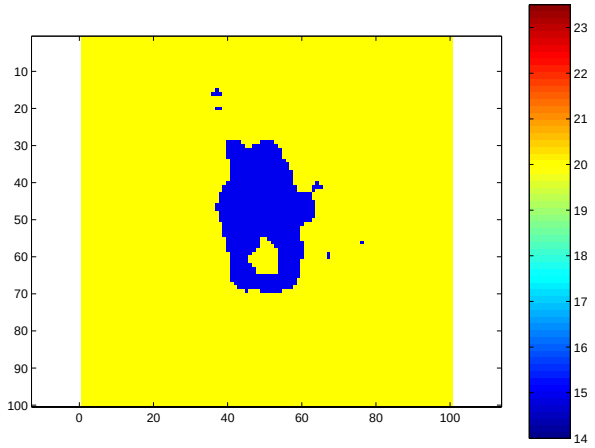


Figure 2: STAF reconstruction of permittivity distribution for the example with full view using $f = 30$ MHz and noisy data with 10 dB SNR.

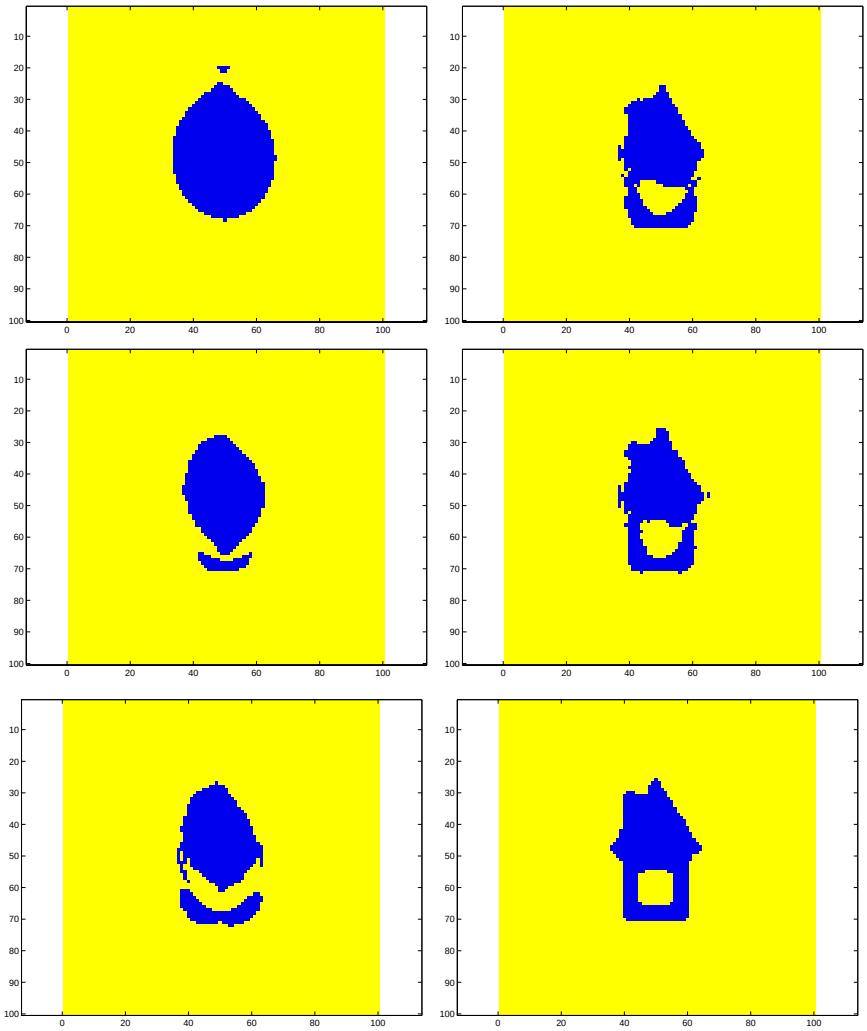


Figure 3: Evolution of permittivity $\epsilon^{(n)}$. Left column from top to bottom: STAF reconstructions of $\epsilon^{(0)}$ for 5 MHz (top left); This is the starting guess for the following reconstruction using levelART. After 10 steps of levelART with 10 MHz; After 30 steps with 10 MHz; Right column from top to bottom: After 10 steps with 20 MHz; After 30 steps with 20 MHz; Final reconstruction after 30 steps of levelART with 30 MHz (bottom right). The algorithm used noise-free data.

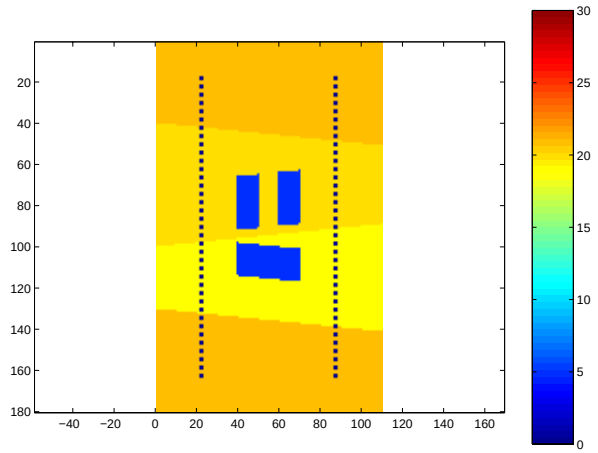


Figure 4: True permittivity distribution in the cross-borehole example. The dots in the figure indicate the source and receiver positions. The permittivity in the background layers is (from top to bottom) $\epsilon_{\rm b} = 21$, 20, 19, and 21. Inside the object it is $\hat{\epsilon} = 5$.

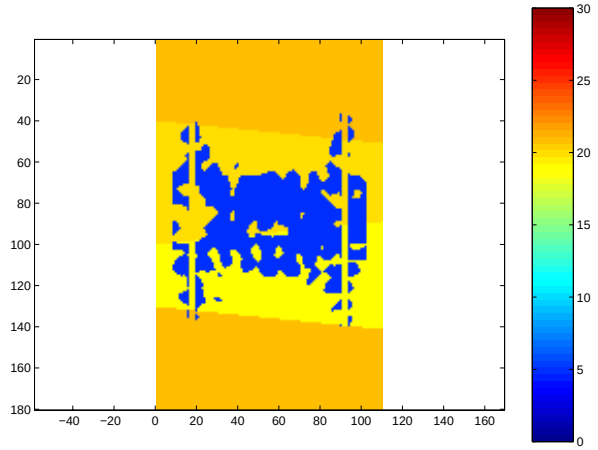


Figure 5: STAF reconstruction of permittivity distribution for cross-borehole example using noisy data with 30 MHz and 10 dB SNR.

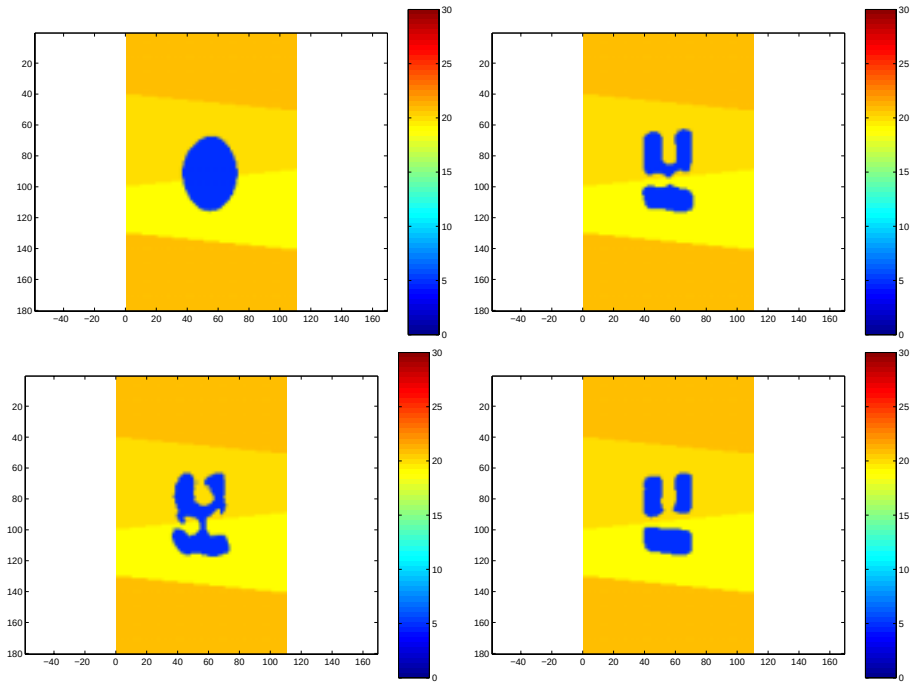


Figure 6: Evolution of the permittivity distribution $\epsilon^{(n)}$ in the cross-borehole example using noisy data with 10 dB SNR and limited view. Top left: STAF reconstruction of $\epsilon^{(0)}$ for 5 MHz. This is the starting guess for the following reconstruction using the levelART algorithm. Bottom left: After 20 steps of levelART with 15 MHz; Top right: After 2 sweeps with 20, 25, and 30 MHz; Bottom right: After 10 sweeps of levelART with 20, 25, and 30 MHz.

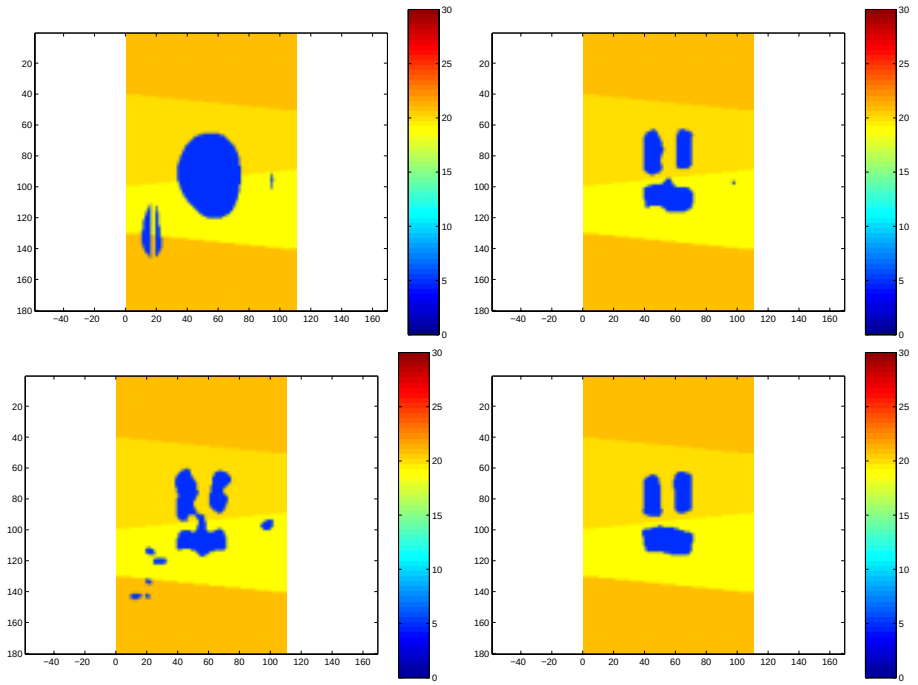


Figure 7: Evolution of the permittivity distribution $\epsilon^{(n)}$ in the cross-borehole example using noisy data with 5 dB SNR and limited view. Top left: STAF reconstruction of $\epsilon^{(0)}$ for 5 MHz. This is the starting guess for the following reconstruction using the levelART algorithm. Bottom left: After 20 steps of levelART with 15 MHz; Top right: After 2 sweeps with 20, 25, and 30 MHz; Bottom right: After 10 sweeps of levelART with 20, 25, and 30 MHz.

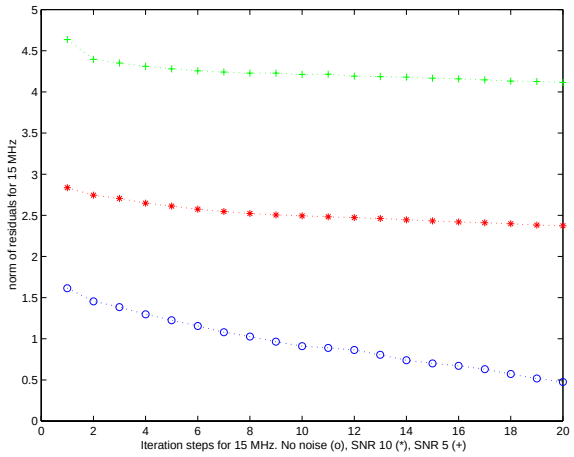


Figure 8: Norm of residuals for 15 MHz in cross-borehole example.

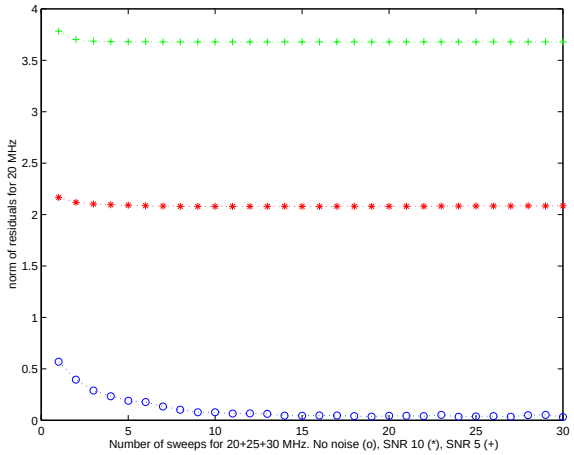


Figure 9: Norm of residuals for 20 MHz in cross-borehole example.

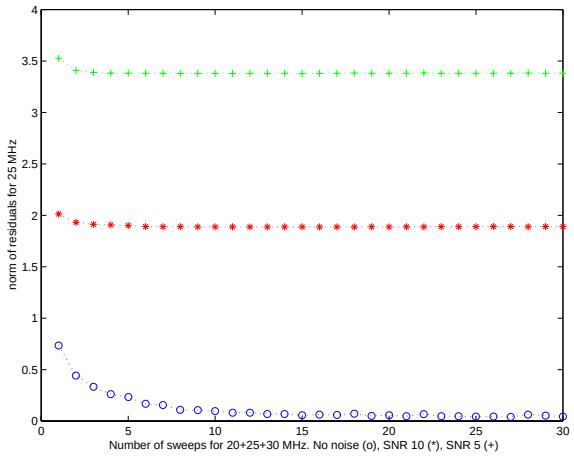


Figure 10: Norm of residuals for 25 MHz in cross-borehole example.

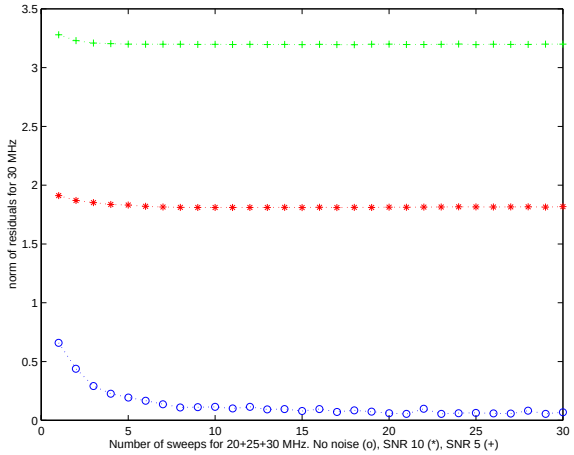


Figure 11: Norm of residuals for 30 MHz in cross-borehole example.

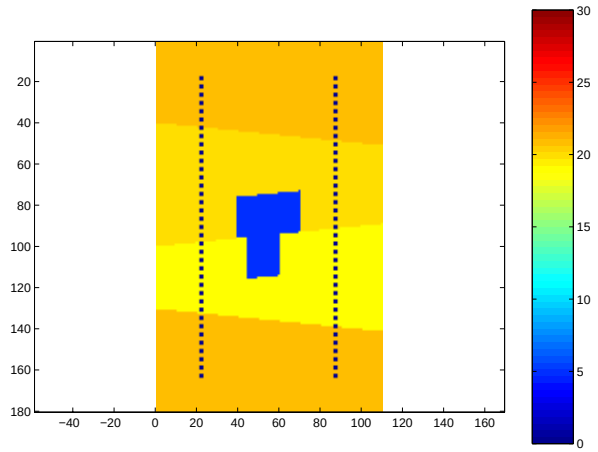


Figure 12: Original permittivity distribution $\epsilon^{(n)}$ in the cross-borehole example using noisy data with 10 dB SNR and limited view. The permittivity $\epsilon_{\rm b}$ in the background is the same as in figure 4, and in the inclusion it is $\hat{\epsilon} = 5$. The conductivity in the background is $\sigma_b = 3.0 \times 10^{-4}$ S/m, and in the inclusion it is either $\hat{\sigma} = 1.0 \times 10^{-6}$ S/m (first example), or $\hat{\sigma} = 1.0 \times 10^{-2}$ S/m (second example). The dots in the figure indicate the source and receiver positions.

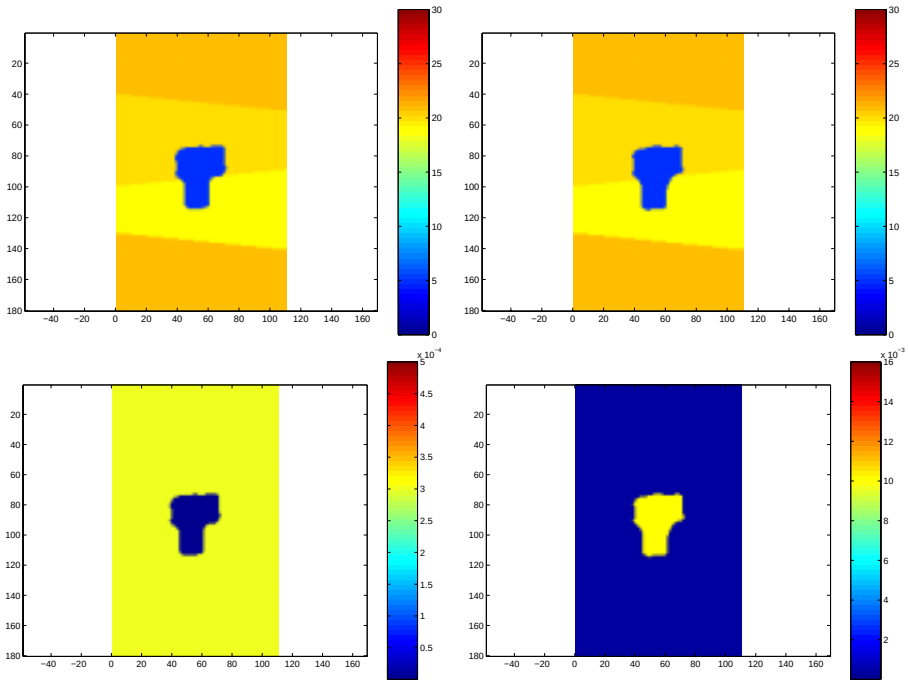


Figure 13: Reconstruction of permittivity and conductivity distributions after applying STAF with 5 MHz, and 6 sweeps of levelART with 15, 20, 25, and 30 MHz. Left column: Example with $\hat{\sigma} = 1.0 \times 10^{-6}$ S/m inside the obstacle. Shown is the reconstructed permittivity (top) and reconstructed conductivity (bottom). Right column: Example with $\hat{\sigma} = 1.0 \times 10^{-2}$ S/m inside the obstacle. Reconstructed permittivity (top) and reconstructed conductivity (bottom). The conductivity was treated as noise in STAF, and was considered linked to the current reconstruction of the permittivity distribution in levelART. All data were contaminated with white Gaussian noise of 10 dB SNR before starting the reconstruction process.

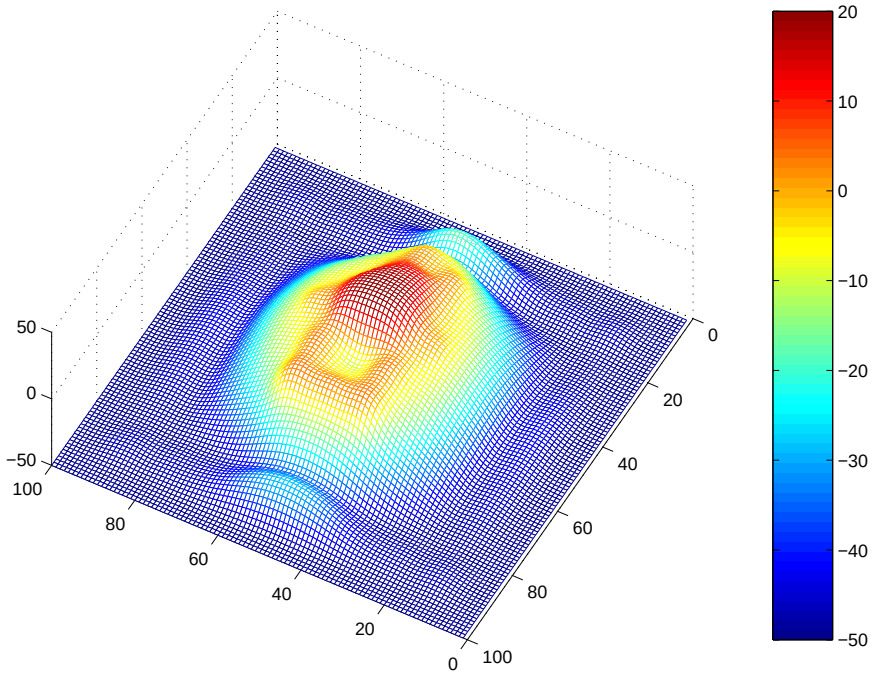


Figure 14: Final level set function $-\phi^{(N)}$ for the full-view example. Figure 3 shows the corresponding permittivity distribution $\epsilon^{(N)}$.

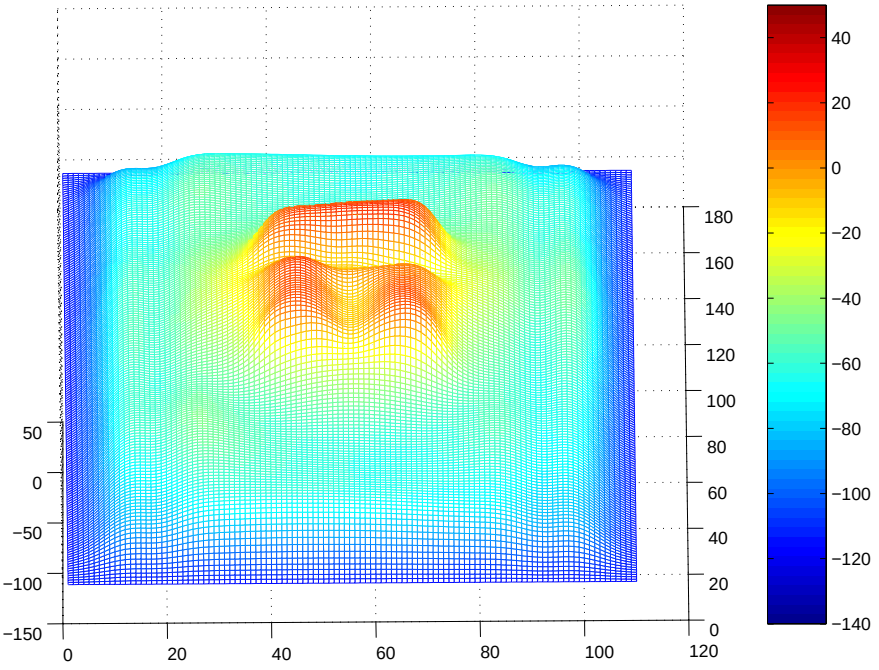


Figure 15: Final level set function $-\phi^{(N)}$ for the cross-borehole example using data with 10 dB SNR. Figure 6 shows the corresponding permittivity distribution $\epsilon^{(N)}$.

2000]Primary 35R30, 86A20;Secondary 78A46

Shape reconstruction in 2D from limited-view multifrequency electromagnetic data

Oliver Dorn, Eric L. Miller, and Carey M. Rappaport

Abstract. In geophysical applications it is often the case that electromagnetic data are available for a number of frequencies but only in a strongly limited-view geometry. In these applications, it might be expected that the highest frequency data alone already give the reconstructions with the highest possible resolution, and the question arises how the additional information given by the lower frequency data should be incorporated into the reconstruction process. In the present paper we consider a shape reconstruction problem in electromagnetic cross-borehole tomography and demonstrate a possible way of using the lower frequency data as a stabilizing and regularizing component in the reconstructions. We employ an iterative shape reconstruction routine which was introduced earlier by the authors and which makes use of an adjoint field inversion technique as well as a level set representation of the shapes. We present numerical experiments which show that stable reconstructions of non-trivial shapes can be achieved from noisy limited-view electromagnetic data by properly incorporating the available lower frequency information.

1. Introduction

In a recent paper [**12**] the authors have introduced a two-step shape reconstruction method for electromagnetic tomography which uses adjoint fields and level sets. The first step of the method employes a source-type reconstruction scheme and yields a good first guess for the shapes. This preprocessing step is designed to be fast and stable and, in particular, to be able to handle the strong nonlinearity in the inverse problem which is due to the high contrasts of the parameter perturbations. The second step is iterative and uses the outcome of the preprocessing step as an initialization. In this second step a level set representation for the shapes is employed in order to be able to keep track of topological changes of the shapes which often occur during the iteration process.

The aim is to retrieve an unknown number of penetrable objects (inclusions) imbedded in an inhomogeneous background medium from observations of electromagnetic (EM) fields. The EM fields in our application are in the frequency band of 5 to 30 MHz. The wavelengths of these fields are typically between 2-15 m in

1991 *Mathematics Subject Classification.* [.

Key words and phrases. Electromagnetic tomography, shape reconstruction, inverse problems.

The work of OD has been supported in part under a grant from the U. S. Dept. of Energy DE-FG07-97ID3566, and under the NSERC Research Grant 84306.

The work of ELM has been supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (award number EEC-9986821), by a grant from the U. S. Dept. of Energy DE-FG07-97ID3566, and by A CAREER Grant from the National Science Foundation MIP-9623721 .

The work of CMR has been supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (award number EEC-9986821), and by The Army Research Office, Multidisciplinary University Research Initiative Grant No. DAAG55-97-0013.

moist soil, where the relative dielectric constant is typically around 20 [**32**]. We are in particular interested in the imaging and monitoring of pollutant plumes at environmental cleanup sites given cross-borehole EM data, where the distances of the boreholes are not much larger than 10-20 m.

We assume that the known conductivity distribution is positive but small everywhere. The electromagnetic characteristics (permittivity and conductivity) of the background medium as well as of the material forming the inclusions are assumed to be known, but the main topological information concerning the *number, sizes, shapes,* and *locations* of the inclusions is missing and has to be reconstructed from the EM data. Inside the pollutant plumes, the permittivity is assumed to be constant with a known value which is typically much smaller (or larger) than the background permittivity, and the background permittivity itself is arbitrary but also known. No topological constraints are made on the shapes of the plumes. For example, they are allowed to be multiply connected, and to enclose 'cavities' or 'holes' filled with background material.

For alternative approaches to solving shape recovery problems in various applications we refer to [**16, 19, 25, 29**] and the references therein.

The paper is organized as follows. In section 2, we will formulate the shape reconstruction problem which we want to solve. Section 3 gives a short overview of the iterative shape reconstruction routine which uses a level set representation of the shapes and adjoint fields for the inversion task. The source-type preprocessing routine is outlined in section 4, and numerical experiments are presented and discussed in section 5. Finally, section 6 summarizes the results of the paper and indicates some interesting directions for future research.

2. The inverse problem

2.1. The Helmholtz Equation. We consider the 2D Helmholtz Equation

$$\Delta u + k^2(x)u \;=\; q(x) \quad \text{in } \mathbb{R}^2, \tag{2.1}$$

with complex wavenumber

$$k^2(x) \;=\; \omega^2\mu_0\epsilon_0\left[\epsilon(x) + i\frac{\sigma(x)}{\omega\epsilon_0}\right]. \tag{2.2}$$

Here, $i^2 = -1$, ω denotes the angular frequency $\omega = 2\pi f$, μ_0 is the magnetic permeability in free space $\mu_0 = 4\pi \times 10^{-7}$ Henrys per meter, ϵ_0 is the dielectric permittivity in free space $\epsilon_0 = 8.854 \times 10^{-12}$ Farads per meter, ϵ is the relative dielectric permittivity (dimensionless), and σ is the electric conductivity in Siemens per meter. The form of (2.2) corresponds to time-harmonic line sources $\tilde{q}(x,t)$ which have a time-dependence $\tilde{q}(x,t) = q(x)e^{-i\omega t}$. For these sources we require that there exists a radius $r_0 > 0$ such that $\text{supp}(q) \subset\subset B_{r_0}(0)$, where $B_r(x) = \{y \in \mathbb{R},\ |x-y| < r\}$ denotes the open ball centered in x with radius $r > 0$. For simplicity we assume throughout the paper that we can find a ball $B_R(0)$ with $R > r$ such that the complex wavenumber $k^2(x)$ is constant with value k_0^2 in $\mathbb{R}^2\backslash B_R(0)$, and that for this k_0 the field u generated by (2.1) satisfies the Sommerfeld radiation condition

$$\lim_{r\to\infty}\sqrt{r}\left(\frac{\partial u}{\partial r} - ik_0u\right) \;=\; 0 \tag{2.3}$$

with $r = |x|$ where the limit is assumed to hold uniformly in all directions $x/|x|$. With this assumption, the problem (2.1)-(2.3) possesses a uniquely determined solution u in $\mathbb{R}^2$ [**6**].

Furthermore we will consider in this paper only the case that the conductivity is positive everywhere, $\sigma > 0$ in $\mathbb{R}^2$, and that it is small. Typical values in our geophysical examples will be $\sigma \approx 10^{-3} - 10^{-4}$ Siemens per meter or less [**32**].

We want to introduce some notation here which will be useful in the following. We denote the wavenumber $k^2(x)$ in short form by

$$k^2(x) \;=\; \kappa(x) \;=\; a\epsilon(x) + ib\sigma(x), \quad a = \omega^2\mu_0\epsilon_0, \quad b = \omega\mu_0. \tag{2.4}$$

The frequencies are assumed to be positive, $\omega > 0$, such that $a, b > 0$.

2.2. Formulation of the shape reconstruction problem. We assume that we are given p different source distributions q_j, $j = 1, \ldots, p$. For each of these sources, data are gathered at the detector positions x_d, $d = 1, \ldots, D_j$, for various frequencies f_k, $k = 1, \ldots, K$. The total number of receivers D_j, as well as their positions x_d, might vary with the source q_j. We assume, for simplicity in the notation, that these positions do not depend on the frequency f_k. This restriction is, however, not necessary for the derivation of the inversion method. We require that there exists a radius $r_1 > 0$ such that all receiver positions are inside the ball of radius r_1, i.e. $x_{jd} \in B_{r_1}(0)$ for all $d = 1, \ldots, D_j$, $j = 1, \ldots, p$.

In the application of *EM cross-borehole tomography*, the sources and receivers are typically situated in some boreholes, and the permittivity distribution ϵ (and/or the conductivity distribution σ) between these boreholes has to be recovered from the gathered data. In the 2D geometry considered here, typical sources are time-harmonic line sources which can be modelled in (2.1) by

$$q_j(x) \;=\; J_j\delta(x - x_j), \quad j = 1, \ldots, p, \tag{2.5}$$

where x_j denotes the 2D coordinates of the j-th line source, $j = 1, \ldots, p$, and the complex number J_j is the strength of the source. We will use these sources in our numerical experiments in section 5.

For a given source q_j and a given frequency f_k we collect a set of data $\tilde{G}_{jk}$ which is described by

$$\tilde{G}_{jk} \;=\; \big(\tilde{u}_{jk}(x_{j1}), \ldots, \tilde{u}_{jk}(x_{jd}), \ldots, \tilde{u}_{jk}(x_{jD_j})\big)^T \quad \in Z_j \tag{2.6}$$

with $Z_j = \mathbb{C}^{D_j}$ being the data space corresponding to a single experiment using one source and one frequency only. In (2.6), the fields $\tilde{u}_{jk}$ solve (2.1)-(2.3) with the correct permittivity distribution $\tilde{\epsilon}(x)$, i.e.

$$\Delta\tilde{u}_{jk} \;+\; [a_k\tilde{\epsilon}(x) + ib_k\sigma(x)]\,\tilde{u}_{jk} \;=\; q_j(x) \qquad \text{in } \mathbb{R}^2 \tag{2.7}$$

with

$$a_k = \omega_k^2\mu_0\epsilon_0, \quad b_k = \omega_k\mu_0, \quad \omega_k = 2\pi f_k. \tag{2.8}$$

More generally, we define for a given source q_j the measurement operator M_j acting on solutions u_{jk} of (2.1) by

$$M_j u_{jk} \;=\; \big(u_{jk}(x_{j1}), \ldots, u_{jk}(x_{jd}), \ldots, u_{jk}(x_{jD_j})\big)^T \quad \in Z_j. \tag{2.9}$$

With this notation, (2.6) is written as

$$\tilde{G}_{jk} \;=\; M_j\tilde{u}_{jk}, \qquad j = 1, \ldots, p,\; k = 1, \ldots, K. \tag{2.10}$$

We gather these data sets $\tilde{G}_{j,k}$ for all sources q_j, $j = 1, \ldots, p$, and all frequencies f_k, $k = 1, \ldots, K$, and the aim is to recover from this collection of data sets

$$\tilde{G} = (\tilde{G}_{1,1}, \ldots, \tilde{G}_{p,K})^T \tag{2.11}$$

the unknown parameter distribution $\tilde{\epsilon}(x)$ in the domain of interest.

Definition 2.1. Let us assume that we are given a constant $\hat{\epsilon} > 0$, an open ball $B_r(0) \subset \mathbb{R}^2$ with $r > \max(r_0, r_1) > 0$, and a bounded function $\epsilon_b : \mathbb{R}^2 \to \mathbb{R}$. We call a pair (Ω, ϵ), which consists of a compact domain $\Omega \subset\subset B_r(0)$ and a bounded function $\epsilon : \mathbb{R}^2 \to \mathbb{R}$, *admissible* if we have

$$\epsilon|_\Omega = \hat{\epsilon}, \quad \epsilon|_{\mathbb{R}^2 \setminus \Omega} = \epsilon_{\mathrm{b}}|_{\mathbb{R}^2 \setminus \Omega}. \tag{2.12}$$

In other words, a pair (Ω, ϵ) is *admissible* if ϵ is equal to a preassigned constant value $\hat{\epsilon}$ inside of Ω, and equal to the preassigned *background permittivity* ϵ_{b} outside of Ω. The domain Ω is called the *scattering domain*.

Remark 2.2. For an admissible pair (Ω, ϵ), and for given $\hat{\epsilon}$, ϵ_{b}, the permittivity ϵ is uniquely determined by Ω.

Formulation of the shape reconstruction problem. Given a constant $\hat{\epsilon} > 0$, a background distribution ϵ_{b}, and some data $\tilde{G}$ as in (2.11). Find a shape $\tilde{\Omega}$ such that the corresponding admissible pair $(\tilde{\Omega}, \tilde{\epsilon})$ reproduces the data, i.e. (2.10) holds with $\tilde{u}_{jk}$ given by (2.7) for $j = 1, \ldots, p$, $k = 1, \ldots, K$.

3. A level set method for shape reconstruction

The *level set method* was originally developed by Osher and Sethian for describing the motion of curves and surfaces [**23, 27**]. Since then, it has found applications in a variety of quite different situations. Examples are image enhancement, computer vision, interface problems, crystal growth, or etching and deposition in the microchip fabrication. For an overview we refer to [**28**].

The idea of using a level set representation as part of a solution scheme for inverse problems involving obstacles was first suggested by Santosa in [**26**]. In that paper, two linear inverse problems, a deconvolution problem and the problem of reconstructing a diffraction screen, are solved by employing an optimization approach as well as a time evolution approach using the level set technique for describing the shapes. Santosa also outlines in that paper how the two presented methods can be generalized to nonlinear shape recovery problems.

More recently, a related method was applied to a nonlinear shape recovery problem by Litman *et al.* in [**18**]. In that work, an inverse transmission problem in free space is solved by a controled evolution of a level set function. This evolution is governed by a Hamilton-Jacobi type equation, whose velocity function has to be determined properly in order to minimize a given cost functional.

The approach developed here is not based on a Hamilton-Jacobi type equation. We follow an optimization approach, and employ a very specific inversion routine (an adjoint field technique) for solving it. This has the advantage that we do not have to propagate the level set function explicitly by computing a numerical Hamiltonian. Instead, our inversion routine provides us in each step with an update that has to be applied directly to the most recent level set function. Doing so, we automatically 'propagate' the level set function until the method converges.

3.1. Level set representation of the domains Ω. Assume that we are given a domain $\Omega \subset\subset B_r(0)$. The characteristic function $\chi_\Omega : \mathbb{R}^2 \to \{0,1\}$ is defined in the usual way as

$$\chi_\Omega(x) = \begin{cases} 1 & , \quad x \in \Omega \\ 0 & , \quad x \in \mathbb{R}^2 \backslash \Omega. \end{cases} \tag{3.1}$$

DEFINITION 3.1. We call a function $\phi : \mathbb{R}^2 \to \mathbb{R}$ a *level set representation of* Ω if

$$\chi_\Omega(x) = \Psi_\phi(x) \qquad \text{on } \mathbb{R}^2 \tag{3.2}$$

where $\Psi_\phi : \mathbb{R}^2 \to \{0,1\}$ is defined as

$$\Psi_\phi(x) = \begin{cases} 1 & , \quad \phi(x) \le 0 \\ 0 & , \quad \phi(x) > 0. \end{cases} \tag{3.3}$$

For each function $\phi : \mathbb{R}^2 \to \mathbb{R}$ there is a domain Ω associated with ϕ by (3.2),(3.3) which we call $\Omega[\phi]$. It is clear that different functions ϕ_1, ϕ_2, $\phi_1 \not\equiv \phi_2$, can be associated with the same domain $\Omega[\phi_1] = \Omega[\phi_2]$, but that different domains cannot have the same level set representation. Therefore, we can use the level set representation for unambiguously specifying a domain Ω by any one of its associated level set functions.

DEFINITION 3.2. We call a triple (Ω, ϵ, ϕ), which consists of a domain $\Omega \subset\subset B_r(0)$ and bounded functions $\epsilon, \phi : \mathbb{R}^2 \to \mathbb{R}$, *admissible* if the pair (Ω, ϵ) is admissible in the sense of definition 2.1, and ϕ is a valid level set representation of Ω.

REMARK 3.3. For an admissible triple (Ω, ϵ, ϕ), and for given $\hat{\epsilon}$, $\epsilon_{\rm b}$, the pair (Ω, ϵ) is uniquely determined by ϕ.

We use this definition to reformulate our shape reconstruction problem.

Level set formulation of the shape reconstruction problem. Given a constant $\hat{\epsilon} > 0$, a background distribution $\epsilon_{\rm b}$, and some data $\tilde{G}$ as in (2.11). Find a level set function $\tilde{\phi}$ such that the corresponding admissible triple $(\tilde{\Omega}, \tilde{\epsilon}, \tilde{\phi})$ reproduces the data, i.e. (2.10) holds with $\tilde{u}_{jk}$ given by (2.7) for $j = 1, \ldots, p,\ k = 1, \ldots, K$.

3.2. Operators. Throughout this paper we will denote the space of parameters $\epsilon_{\rm s}$ by F, and the space of level set functions ϕ describing the shapes Ω by Φ. For simplicity, both spaces are assumed to be L_2-Hilbert spaces equipped with suitable inner products. The space of data $\tilde{G}_{jk}$ corresponding to the source position q_j and frequency f_k is given by $Z_j = \mathbb{C}^{D_j}$, where D_j is the number of receivers correponding to the source q_j.

We consider the following *decomposition of* ϵ *in* $\mathbb{R}^2$

$$\text{(i)} \quad \epsilon = \epsilon_{\rm b} + \epsilon_{\rm s} \quad \text{in } \mathbb{R}^2 \tag{3.4}$$

$$\text{(ii)} \quad \operatorname{supp}(\epsilon_{\rm s}) \subset\subset B_r(0). \tag{3.5}$$

It says that the permittivity distribution ϵ is decomposed into the background distribution $\epsilon_{\rm b}$ and the perturbation $\epsilon_{\rm s}$ which is assumed to have compact support and which we will refer to as the *scattering permittivity* in the following.

Given a constant $\hat{\epsilon}$ and a bounded function $\epsilon_{\rm b} : \mathbb{R}^2 \to \mathbb{R}$, there is with each level set function $\phi \in \Phi$ a uniquely determined scattering permittivity $\Lambda(\phi)$ associated by putting

$$\Lambda(\phi)(x) \;=\; \begin{cases} \hat{\epsilon} - \epsilon_{\rm b}(x) & , \quad \phi(x) \le 0 \\ 0 & , \quad \phi(x) > 0. \end{cases} \tag{3.6}$$

With (3.3) we can write this also as

$$\Lambda(\phi)(x) \;=\; \Psi_\phi(x)(\hat{\epsilon} - \epsilon_{\rm b}(x)) \quad , \quad x \in \mathbb{R}^2. \tag{3.7}$$

Notice that the operator Λ is chosen such that the triple (Ω, ϵ, ϕ) with $\epsilon = \epsilon_{\rm b} + \Lambda(\phi)$ and domain $\Omega[\phi]$ forms an admissible triple (Ω, ϵ, ϕ) in the sense of definition 3.2.

We see that, for (Ω, ϵ, ϕ) an admissible triple, $\Lambda(\phi)$ is just the scattering permittivity $\epsilon_{\rm s}$ as defined in (3.4),(3.5)

$$\Lambda(\phi)(x) \;=\; \epsilon_{\rm s}(x) \;=\; \chi_\Omega(x)(\hat{\epsilon} - \epsilon_{\rm b}(x)) \quad , \quad x \in \mathbb{R}^2. \tag{3.8}$$

Let us assume now that we are given a background permittivity $\epsilon_{\rm b}$ and that we have collected some data $\tilde{G}_{jk}$ which correspond to the 'true' permittivity distribution

$$\tilde{\epsilon} \;=\; \epsilon_{\rm b} + \tilde{\epsilon}_s, \tag{3.9}$$

where $\tilde{\epsilon}_s$ is the 'true' scattering permittivity. The *residual operators* R_{jk} map for a source position q_j and a frequency f_k a given scattering permittivity ϵ_s to the corresponding mismatch in the data

$$R_{jk} : F \longrightarrow Z_j \quad , \quad R_{jk}(\epsilon_{\rm s}) \;=\; M_j u_{jk} - \tilde{G}_{jk} \tag{3.10}$$

where u_{jk} solves

$$\Delta u_{jk} \;+\; \left[a_k(\epsilon_{\rm b} + \epsilon_{\rm s})(x) + i b_k \sigma(x)\right] u_{jk} \;=\; q_j \tag{3.11}$$

and M_j is the measurement operator defined in (2.9). From (2.10) we see that for the 'true' scattering permittivity the residuals vanish,

$$R_{jk}(\tilde{\epsilon}_{\rm s}) \;=\; 0 \qquad \text{for } j = 1, \ldots, p, \; k = 1, \ldots, K, \tag{3.12}$$

if the data are noise-free.

The *forward operators* T_{jk} which map a given level set function $\phi \in \Phi$ into the corresponding mismatch in the data are defined by

$$T_{jk} : \Phi \longrightarrow Z_j \quad , \quad T_{jk}(\phi) \;=\; R_{jk}(\Lambda(\phi)) \tag{3.13}$$

for $j = 1, \ldots, p$, $k = 1, \ldots, K$. The goal is to find a level set function $\tilde{\phi} \in \Phi$ such that

$$T_{jk}(\tilde{\phi}) \;=\; 0 \qquad \text{for } j = 1, \ldots, p, \; k = 1, \ldots, K. \tag{3.14}$$

We mention that all three operators Λ, R_{jk} and T_{jk} are nonlinear.

We will also need an expression for the linearized operators $T'_{jk}[\phi]$. As it is described in [**12**], it is convenient to use an approximation to these operators which is easier to handle in the formal derivation of the adjoint scheme employed here. The approximated linearized operator is given by

$$T'_{jk}[\phi] : \Phi \longrightarrow Z_j \quad , \quad T'_{jk}[\phi]\delta\phi \;=\; M_j v_{jk} \tag{3.15}$$

where v_{jk} solves the linearized equation

$$\Delta v_{jk} \;+\; \left[a_k(\epsilon_{\rm b} + \epsilon_{\rm s})(x) + i b_k \sigma(x)\right] v_{jk} \;=\; -\, a_k \delta\epsilon_{\rm s}(x) u_{jk}(x), \tag{3.16}$$

with u_{jk} a solution of (3.11) and

$$\delta\epsilon_{\rm s}(x) \;=\; -\,[\hat{\epsilon}-\epsilon_{\rm b}(x)]\,\frac{\delta\phi(x)}{|\nabla\phi(x)|}C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x). \tag{3.17}$$

Here, $\chi_{B_\rho(\Gamma)}(x)$ is the characteristic function supported on the finite width neighborhood $B_\rho(\Gamma) = \cup_{y\in\Gamma}B_\rho(y)$, which is introduced in [**12**] as an approximation to the surface Dirac delta function concentrated on $\Gamma = \partial\Omega[\phi]$. We have $C_\rho(\Gamma) = L(\Gamma)/\mathrm{Vol}(B_\rho(\Gamma))$ where $L(\Gamma) = \int_{B_r(0)}\hat{\delta}_\Gamma(x)dx$ is the length of Γ, and $\mathrm{Vol}(B_\rho(\Gamma)) = \int_{B_r(0)}\chi_{B_\rho(\Gamma)}(x)dx$ is the volume of $B_\rho(\Gamma)$. For a very small ρ we will get a very large weight $C_\rho(\Gamma)$, whereas for increasing ρ this weight $C_\rho(\Gamma)$ decreases accordingly.

3.3. A nonlinear Kaczmarz-type approach. The algorithm works in a 'single-step fashion' as follows. Instead of using the data (2.11) for all sources and all frequencies simultaneously, we only use the data for one source and one frequency at a time while updating the linearized residual operator after each determination of the corresponding incremental correction $\delta\phi$.

To be more specific, let us assume that we are given a level set function $\phi^{(n)}(x)$ and a scattering permittivity $\epsilon_{\rm s}^{(n)}(x)$ such that $(\Omega^{(n)}, \epsilon_{\rm b}+\epsilon_{\rm s}^{(n)}, \phi^{(n)})$ forms an admissible triple in the sense of definition 3.2. Using a data set $\tilde{G}_{jk}$ corresponding to the fixed source position q_j and the frequency f_k, we want to find an update $\delta\phi^{(n)}$ to $\phi^{(n)}$ such that for the admissible triple

$$\left(\Omega^{(n+1)}, \epsilon_{\rm b}+\epsilon_{\rm s}^{(n+1)}, \phi^{(n+1)}\right) \;:= \tag{3.18}$$

$$\left(\Omega[\phi^{(n)}+\delta\phi^{(n)}], \epsilon_{\rm b}+\Lambda(\phi^{(n)}+\delta\phi^{(n)}), \phi^{(n)}+\delta\phi^{(n)}\right)$$

the residuals in the data corresponding to this source and this frequency vanish

$$T_{jk}(\phi^{(n+1)}) \;=\; T_{jk}(\phi^{(n)}+\delta\phi^{(n)}) \;=\; 0. \tag{3.19}$$

Applying a Newton-type approach, we get from (3.19) a correction $\delta\phi^{(n)}$ for $\phi^{(n)}$ by solving

$$T'_{jk}[\phi^{(n)}]\delta\phi^{(n)} \;=\; -\,T_{jk}(\phi^{(n)}) \;=\; -\,\left(M_j\,u_{jk}-\tilde{G}_{jk}\right) \tag{3.20}$$

where u_{jk} satisfies (3.11) with $\epsilon_{\rm s} = \Lambda(\phi^{(n)})$

$$\Delta u_{jk} \;+\; \left[a_k(\epsilon_{\rm b}+\Lambda(\phi^{(n)}))(x) + ib_k\sigma(x)\right]u_{jk} \;=\; q_j(x) \tag{3.21}$$

and

$$\epsilon_{\rm b}(x)+\Lambda(\phi^{(n)})(x) \;=\; \begin{cases} \hat{\epsilon} & , \quad x\in\Omega[\phi^{(n)}] \\ \epsilon_{\rm b}(x) & , \quad x\in\mathbb{R}^2\setminus\Omega[\phi^{(n)}]. \end{cases} \tag{3.22}$$

Since we have only few data given for one source and one frequency, equation (3.20) usually will have many solutions (in the absence of noise), such that we have to pick one according to some criterion. We choose to take that solution which minimizes the energy norm of $\delta\phi^{(n)}$

$$\mathrm{Min}\;\|\delta\phi^{(n)}\|_2 \quad \text{subject to} \quad T'_{jk}(\phi^{(n)})\delta\phi^{(n)} \;=\; -\left(M_j\,u_{jk}-\tilde{G}_{jk}\right). \tag{3.23}$$

This solution can be formulated explicitly. It is

$$\delta\phi_{\rm MN}^{(n)} \;=\; -\,T'_{jk}[\phi^{(n)}]^*\left(T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^*\right)^{-1}\left(M_j\,u_{jk}-\tilde{G}_{jk}\right), \tag{3.24}$$

where $T'_{jk}[\phi^{(n)}]^*$ denotes the adjoint operator to $T'_{jk}[\phi^{(n)}]$.

After correcting ϕ by $\phi \to \phi + \delta\phi_{jk}$, where $\delta\phi_{jk}$ is given by (3.24), we use the updated residual equation (3.20) to compute the next correction $\delta\phi_{j'k'}$. Doing this for one equation after the other, until each of the sources q_j and each of the frequencies f_k has been considered exactly once, will yield one complete sweep of the algorithm. This procedure is similar to the Kaczmarz method for solving linear systems, or the algebraic reconstruction technique (ART) in x-ray tomography [**20**] and the simultaneous iterative reconstruction technique (SIRT) as presented in [**9**]. Related approaches have also been employed in ultrasound tomography by Natterer and Wübbeling [**21**], in more general bilinear inverse problems by Natterer [**22**], in optical tomography by Dorn [**10**], and in 3D-electromagnetic induction tomography (EMIT) by Dorn *et alii* [**11**]. Because of the similarity of the derived method to the ART scheme in CT, we will often refer to it in short as 'levelART'.

3.4. The adjoint linearized operator $T'_{jk}[\phi^{(n)}]^*$. In order to calculate the minimal norm solution (3.24), we will need a practically useful expression for the adjoint of the linearized operator $T'_{jk}[\phi]$. We will present such an expression in this section. The calculation of the action of this operator will typically require to solve an adjoint Helmholtz problem. This explains the name 'adjoint field method' which is often used for the general inversion method employed here.

Let $\hat{C}_{jk}$ be defined as $\hat{C}_{jk} = T'_{jk}[\phi^{(n)}]T'_{jk}[\phi^{(n)}]^*$, or some approximation to it as it was derived in [**12**]. In order to calculate a correction $\delta\phi^{(n)}_{\mathrm{MN}}$ by (3.24) we have to apply the operator $T'_{jk}[\phi^{(n)}]^*$ to the vector

$$\zeta := \hat{C}_{jk}^{-1}\left(M_j u_{jk} - \tilde{G}_{jk}\right). \tag{3.25}$$

An explicit formula for $T'_{jk}[\phi^{(n)}]^*\zeta$ is given next.

$$T'_{jk}[\phi^{(n)}]^*\zeta = \frac{[\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)]}{a_k|\nabla\phi^{(n)}(x)|}\,\mathbf{Re}\left(\overline{u_{jk}(x)z_{jk}(x)}\right)\, C_\rho(\Gamma)\chi_{B_\rho(\Gamma)}(x), \tag{3.26}$$

where u_{jk} solves

$$\Delta u_{jk} + \kappa_k(x)u_{jk} = q_j(x), \tag{3.27}$$

and z_{jk} solves the 'adjoint equation'

$$\Delta z_{jk} + \kappa_k(x)z_{jk} = \sum_{d=1}^{D_j}\bar{\zeta}_d\delta(x - x_{jd}) \tag{3.28}$$

with

$$\kappa_k(x) = a_k\left[\epsilon_{\mathrm{b}}(x) + \Lambda(\phi^{(n)})(x)\right] + ib_k\sigma(x) \tag{3.29}$$

and a_k, b_k defined as in (2.8).

3.5. Implementation: The levelART algorithm. In the levelART algorithm described in Table 1, η is a relaxation parameter for the update of the level set function which is determined empirically. The constant $C_\rho(\Gamma)$ could be calculated explicitly for the actual curve $\Gamma^{(n)}$, or it could be approximated by some value corresponding to a simple geometrical object (to give an example, in case of a single circle it would be $C_\rho(\Gamma) = (2\rho)^{-1}$). In our numerical experiments so far, however, it is simply considered as part of η. The same holds true for the constant c_1 which has been introduced in [**12**] as an approximation to $|\nabla\phi|$. The constant

Table 1: *The 'levelART' algorithm.*

Preparation step.

- Calculate $\hat{C}_{jk}$ and $D_{jk} = \hat{C}_{jk}^{-1}$ for each source q_j, $j = 1, \ldots, p$, and each frequency f_k, $k = 1, \ldots, K$, and store in memory for later use.
- Build groups of frequencies $G_m = \{f_1, \ldots, f_{K_m}\}$, $m = 1, \ldots, M$.

Initialization.

$n = 0$;

$(\Omega^{(0)}, \epsilon^{(0)}, \phi^{(0)})$ given from low-frequency STAF.

Reconstruction loop.

FOR $m = 1 : M$ march over frequency groups G_m

FOR $i = 1 : I_m$ perform I_m sweeps for frequency group G_m

FOR $k = 1 : K_m$ march over frequencies in G_m

FOR $j = 1 : p$ This loop needs only one LR-factorization

$\zeta_{jk} = D_{jk}(M_j u_{jk} - \tilde{G}_{jk})$; u_{jk} solves (3.27) with $\epsilon^{(n)}$, q_j, f_k

$\delta\phi_{jk} = -\frac{\hat{\epsilon} - \epsilon_{\mathrm{b}}(x)}{a_k} \mathbf{Re}(\overline{u_{jk} z_{jk}}) \chi_{B_\rho(\Gamma)}$; z_{jk} solves (3.28) with $\epsilon^{(n)}$ and ζ_{jk}

END

$\delta\phi^{(n)}(x) = \sum_{j=1}^{p} \delta\phi_{jk}(x)$;

$\phi^{(n+1)} = C_{\mathrm{LS}}^{(n)}(\phi^{(n)} + \eta \frac{C_\rho(\Gamma)}{c_1} \delta\phi^{(n)})$; update level set function

Optional step: 'curve shortening by diffusion'. See [**12**].

$\epsilon^{(n+1)} = \epsilon_{\mathrm{b}} + \Lambda(\phi^{(n+1)})$; $n = n + 1$; Reinitialization $n \to n + 1$

END

END alternatively, some stopping criteria can be used here

END

$(\Omega^{(N)}, \epsilon^{(N)}, \phi^{(N)}) = (\Omega[\phi^{(n)}], \epsilon_{\mathrm{b}} + \Lambda(\phi^{(n)}), \phi^{(n)})$; Final reconstruction.

ρ is in our numerical experiments chosen between 30-40 cm, which corresponds to 2-3 grid cells. The scaling factor $C_{\mathrm{LS}}^{(n)}$ is determined after each update to keep the global minimum (or maximum) of the level set function at a constant value. For more details see [**12**].

4. A simple source-type method (STAF)

In the framework of inverse scattering problems, the method we propose here can be considered as a 'source-type inversion method'. Roughly speaking, the general idea of source-type reconstruction methods in inverse scattering is to split a given nonlinear inverse scattering problem into two subproblems. The first one is linear, and tries to recover a virtual 'equivalent source' in the medium which would be able to fit the data if applied with the known background distribution. This equivalent source is related to the unknown scattering permittivity by a nonlinear 'constitutive' relation. Therefore, in the second part of the algorithm, a nonlinear inverse problem has to be solved to derive the scattering permittivity from the recovered equivalent source distribution.

This idea is not at all new. It has been applied for example in the Source-Type Integral Equation (STIE) method of Habashy *et al.* [**14**], or in the method presented by Chew *et al.* in [**5**]. More recently, similar ideas have been applied by Abdullah *et al.* [**1**], Caorsi *et al.* [**3**], and van den Berg *et al.* [**30, 31**].

All of these approaches have in common that they use the source-type method as a stand-alone inversion scheme. Such a method has the advantage that it is not as sensitive to strong nonlinearities in the inverse problem as for example perturbation methods or the Born or Rytov approximation are [**8, 13, 17**].

On the other hand, interpreting the inverse scattering problem as an inverse source problem is not without drawbacks. For example, the existence of so-called 'non-radiating sources' or 'invisible sources' gives rise to a nonuniqueness in the inverse source problem, which is difficult to deal with when solving the nonlinear part [**1, 7, 14**]. Moreover, it is not clear at all how to combine properly the information corresponding to different experiments, since each experiment creates its own 'equivalent sources' and its own 'invisible sources'. For more information about possible applications, advantages and drawbacks of the source-type scheme as a stand-alone inversion tool we refer to [**1, 2, 3, 5, 7, 14, 30, 31**].

Our approach is different from those mentioned above. We only want to find a good approximation to the shapes of the objects, and a corresponding initial level set function which can be used to start the levelART algorithm. In particular, we can make use of our prior information about the permittivity distribution. This will allow us to circumvent most of the problems of source-type schemes which have been mentioned above.

We will now describe this method, which we will call the Source-Type Adjoint Field (STAF) method, in more details.

4.1. Formulation as an inverse scattering problem. To start with we formulate the following inverse scattering problem.

Inverse Scattering Problem. Let us assume that we are given a bounded function $\epsilon_{\mathrm{b}} : \mathbb{R}^2 \to \mathbb{R}$, and some data $\tilde{G}$ as in (2.11). Find a bounded function $\tilde{\epsilon}_{\mathrm{s}} : \mathbb{R}^2 \to \mathbb{R}$ with $\mathrm{supp}(\epsilon_{\mathrm{s}}) \subset\subset B_r(0)$ such that $\tilde{\epsilon} = \epsilon_{\mathrm{b}} + \tilde{\epsilon}_{\mathrm{s}}$ reproduces the data, i.e. (2.10) holds with $\tilde{u}_{jk}$ given by (2.7) for $j = 1, \ldots, p,\ k = 1, \ldots, K$.

For a fixed frequency f_k and a source q_j, let $\tilde{u}_{jk}$ be the solution of

$$\Delta \tilde{u}_{jk} + \big(a_k(\epsilon_{\mathrm{b}} + \tilde{\epsilon}_{\mathrm{s}})(x) + i b_k \sigma(x)\big)\, \tilde{u}_{jk} = q_j(x), \tag{4.1}$$

and let u_{jk} be the solution of the 'unperturbed' equation

$$\Delta u_{jk} + (a_k \epsilon_{\mathrm{b}}(x) + i b_k \sigma(x))\, u_{jk} = q_j(x). \tag{4.2}$$

Define

$$\tilde{v}_{jk} := u_{jk} - \tilde{u}_{jk}. \tag{4.3}$$

Subtraction of (4.2) from (4.1) shows that $\tilde{v}_{jk}$ solves

$$\Delta \tilde{v}_{jk} + (a_k \epsilon_{\mathrm{b}}(x) + i b_k \sigma(x))\, \tilde{v}_{jk} = \tilde{Q}^s_{jk}(x), \tag{4.4}$$

where the 'scattering source' $\tilde{Q}^s_{jk}(x)$ is defined as

$$\tilde{Q}^s_{jk}(x) = a_k \tilde{\epsilon}_{\mathrm{s}}(x) \tilde{u}_{jk}(x). \tag{4.5}$$

Denoting by Y the Hilbert space of sources, we introduce a 'source type' forward operator A_{jk} by putting

$$A_{jk} : Y \longrightarrow Z_j \quad , \quad A_{jk} Q^s_{jk} = M_j v_{jk} \tag{4.6}$$

where M_j is the measurement operator defined in (2.9), and v_{jk} solves

$$\Delta v_{jk} + (a_k \epsilon_{\mathrm{b}}(x) + i b_k \sigma(x))\, v_{jk} = Q^s_{jk}(x). \tag{4.7}$$

The operator A_{jk} is linear.

Let us assume now that we apply the 'correct' scattering source $\tilde{Q}^s_{jk}(x)$ defined by (4.5) as argument of A_{jk}. Then we know from (2.10), (3.9), (3.10) that

$$A_{jk} \tilde{Q}^s_{jk} = M_j \tilde{v}_{jk} = M_j (u_{jk} - \tilde{u}_{jk}) = M_j u_{jk} - \tilde{G}_{jk} = R_{jk}(0). \tag{4.8}$$

The vectors $R_{jk}(0)$ are easily computed by solving a forward problem on the background distribution (4.2). Therefore, all we have to do to get back the scattering source $\tilde{Q}^s_{jk}$ from the data $\tilde{G}_{jk}$ is to solve (4.8) for $\tilde{Q}^s_{jk}$. Doing so amounts to solving an ill-posed but *linear* inverse problem.

Once we have recovered $\tilde{Q}^s_{jk}(x)$, we want to get back $\tilde{\epsilon}_s(x)$ by using the constitutive relation (4.5). This second part of the inversion scheme can be interpreted as solving a *nonlinear* inverse problem since $\tilde{u}_{jk}(x)$ depends on $\tilde{\epsilon}_s(x)$.

Notice that $\tilde{Q}^s_{jk}(x)$ varies with different sources and frequencies, but that $\tilde{\epsilon}_s(x)$ is the same for all sources and all frequencies (if we neglect dispersion). We will make use of this observation when we try to solve the nonlinear part (4.5). In the following, we describe the method which we will use to recover the scattering source $\tilde{Q}^s_{jk}(x)$ from a given data set $\tilde{G}_{jk}$ for a fixed source q_j and a fixed frequency f_k.

4.2. Looking for a scattering source. Since for a fixed (primary) source position and a fixed frequency we have only few data given to recover $\tilde{Q}^s_{jk}$, and since we have to take into account that also 'non-radiating' and 'invisible' sources have been generated in the experiment, we assume that there will be many solutions (in absence of noise) of (4.8). To pick one, we are looking for the solution with minimal norm

$$\text{Min } \|Q^s_{jk}\|_Y \quad \text{subject to} \quad A_{jk} Q^s_{jk} = R_{jk}(0). \tag{4.9}$$

It is given by

$$Q^s_{jk,\mathrm{MN}} = A^*_{jk} \left(A_{jk} A^*_{jk}\right)^{-1} R_{jk}(0), \tag{4.10}$$

where A^*_{jk} denotes the adjoint operator to A_{jk}. In our numerical experiments presented in this paper, we will simply replace the operator $A_{jk} A^*_{jk}$ by (a constant times) the identity which amounts to applying strong regularization.

The following theorem, which is proven in [**12**], tells us how to calculate the action of A^*_{jk} on a vector $\zeta \in Z_j$ in an efficient way.

THEOREM 4.1. *Let $\zeta = (\zeta_1, \ldots, \zeta_{D_j})^T \in Z_j$ and let x_{jd}, $d = 1, \ldots, D_j$ be the detector positions corresponding to the source q_j. Then, $A^*_{jk}\zeta$ is given by*

$$A^*_{jk}\zeta = \overline{z_{jk}}\, \chi_{B_r(0)}, \tag{4.11}$$

where z_{jk} solves

$$\Delta z_{jk} + (a_k \epsilon_b + i b_k \sigma)\, z_{jk} = \sum_{d=1}^{D_j} \overline{\zeta_d} \delta(x - x_{jd}). \tag{4.12}$$

4.3. Recovery of the scattering permittivity. After we have found a scattering source Q^s_{jk} which satisfies (4.9), we want to use the constitutive relation

$$Q^s_{jk}(x) \;=\; a_k \tilde{\epsilon}_s(x) \tilde{u}_{jk}(x), \tag{4.13}$$

which holds for the 'correct' scattering source $\tilde{Q}^s_{jk}$ according to (4.5), to find an approximation for $\tilde{\epsilon}_s(x)$.

Let $\tilde{u}_{jk}$ be a solution of (4.1) and u_{jk} a solution of (4.2). We decompose Q^s_{jk}, $\tilde{u}_{jk}$, u_{jk} and $\tilde{\epsilon}_s$ into amplitude and phase

$$Q^s_{jk}(x) \;=\; |Q^s_{jk}(x)|\, e^{ir(x)}, \qquad \tilde{u}_{jk} \;=\; |\tilde{u}_{jk}|\, e^{i\tilde{s}(x)}, \tag{4.14}$$

$$u_{jk} \;=\; |u_{jk}|\, e^{is(x)}, \qquad \epsilon_s(x) \;=\; |\epsilon_s(x)|\, e^{it(x)}, \tag{4.15}$$

where we have omitted the subscripts jk in the argument functions r, $\tilde{s}$, s, and t for simplicity in the notation. Making use of the fact that $\tilde{\epsilon}_s(x) \in \mathbb{R}$ we see that

$$t(x) \;\in\; \{0, \pi\} \quad \text{for all } x \in \mathbb{R}^2. \tag{4.16}$$

With (4.14),(4.15) equation (4.9) decomposes into two equations, one for the amplitude and one for the phase. They are

$$|Q^s_{jk}(x)| \;=\; a_k\, |\tilde{\epsilon}_s(x)|\, |\tilde{u}_{jk}|, \tag{4.17}$$

$$r(x) \;=\; \tilde{s}(x) \;+\; t(x). \tag{4.18}$$

The observation in our numerical experiments is that, although $s(x)$ and $\tilde{s}(x)$ might be quite different from each other for large perturbations $\tilde{\epsilon}_s(x)$, the amplitudes $|u_{jk}(x)|$ and $|\tilde{u}_{jk}(x)|$ do not differ too much from each other in the scattering region. Therefore, in our applications it is a reasonable approximation to assume that

$$|\tilde{u}_{jk}(x)| \;\approx\; |u_{jk}(x)| \quad \text{in } B_r(0). \tag{4.19}$$

With this approximation, (4.17) yields the following estimate for $|\tilde{\epsilon}_s(x)|$

$$|\tilde{\epsilon}^{(jk)}_s(x)| \;\approx\; \frac{|Q^s_{jk}(x)|}{a_k |u_{jk}(x)|} \quad \text{in } B_r(0). \tag{4.20}$$

We have added the indices j and k on the left hand side of (4.20) to indicate that we have used only the data $\tilde{G}_{jk}$ corresponding to source q_j and frequency f_k for its determination.

Notice that the step (4.19), (4.20) is *nonlinear* since taking the amplitude of a complex number is a nonlinear operation. Therefore, the approach presented here is quite different from the usual Born approximation which approximates $\tilde{u}_{jk}$ by u_{jk}.

For the purposes of the present paper, the determination of $|\tilde{\epsilon}_s(x)|$ is already sufficient in order to get a good first guess for the shapes of the unknown inclusions, since we can now make use of our prior information about the correct value of $\hat{\epsilon}$ in (2.12).

In source-type reconstruction schemes it is generally not clear what the best way is to combine information corresponding to different source positions q_j, $j = 1, \ldots, p$, in the reconstruction process. We have chosen to use the following formula

Table 2: *The STAF algorithm.*

Select a frequency f_k.

FOR $j = 1 : p$ This loop needs only one LR-factorization

$R_{jk}(0) = M_j u_{jk} - \tilde{G}_{jk}, \qquad u_{jk}$ solves (4.2)

$Q^s_{jk} = A^*_{jk} R_{jk}(0) = \overline{z_{jk}}\, \chi_{B_r(0)}, \quad z_{jk}$ solves (4.12) with $\zeta_{jk} = R_{jk}(0)$

$|\tilde{\epsilon}^{(jk)}_s(x)| = \frac{|Q^s_{jk}(x)|}{a_k |u_{jk}(x)|}$

END

$|\tilde{\epsilon}_s(x)| = \frac{1}{p}\sum_{j=1}^{p} |\tilde{\epsilon}^{(jk)}_s(x)|, \qquad x \in B_r(0)$

Fix a constant $C_l > 0$. The reconstructed shape is $\{x \in B_r(0) \ : \ |\tilde{\epsilon}_s(x)| \geq C_l\}$.

for this purpose:

$$|\tilde{\epsilon}^{(k)}_s(x)| \ = \ \frac{1}{p}\sum_{j=1}^{p} |\tilde{\epsilon}^{(jk)}_s(x)|. \tag{4.21}$$

We mention that the reconstruction formula (4.21) was chosen here mainly because of its apparent similarity to the backprojection technique without filtering in Computerized Tomography (CT), with the lines replaced by more complicated patterns given by $|\tilde{\epsilon}^{(jk)}_s(x)|$. It should be emphasized, however, that the mathematical derivation of the reconstruction scheme applied here is completely different from the derivation of the backprojection scheme in CT. Whereas the latter one admits the interpretation as an adjoint scheme for approximately inverting the linear Radon transform [**15**], the derivation of (4.21) involves a nonlinear step (taking the absolute values of the fields), and tries to approximately invert a *nonlinear* forward operator.

4.4. Implementation: The STAF algorithm. The Source Type Adjoint Field (STAF) scheme, which we have employed in our numerical experiments presented in section 5, is in brief algorithmic form described in Table 2.

5. Numerical Experiments

5.1. Discretization of the computational domain. In our numerical experiments, we use a Finite-Differences Frequency Domain (FDFD) code written in MATLAB for solving (2.1)-(2.3). The code uses appropriately designed perfectly matched layers (PML) to avoid reflections at the artificial computational boundaries [**24**].

The physical domain is partitioned into 100×100 elementary cells (pixels) in the first numerical example, and into 180×110 elementary cells in the second example. Each of these grid cells has a physical size of approximately 0.14×0.14 m^2, such that the total computational domain in the first example covers an area of 14×14 m^2, and in the second example of 15×25 m^2. The eight layers which are closest to the boundaries of the computational domain are used as a PML.

We will refer to the first numerical example as the 'full-view' situation, and to the second numerical example as the 'limited-view', 'cross-borehole' or 'geophysical' situation. This terminology is motivated by the source and receiver geometries used, which are as follows.

In the full-view example, we have 64 sources and 64 receivers given which surround the domain of interest. Each source position is at the same time a receiver position and vice versa. The distance of two adjacent sources or receivers from each other is four pixels or 55 cm. The area enclosed by these sources and receivers has a size of 10×10 m^2.

In the two limited-view examples, 74 sources and receivers are positioned equally spaced in two boreholes. The distance of two adjacent sources or receivers from each other is again 4 pixels or 55 cm, and the distance of the two boreholes from each other is about 10 m.

We mention that, in all of our numerical examples, the regions beyond the source and receiver positions are part of the inversion problem, too. This means, the area which has to be recovered from the data is the whole area situated between the PML boundaries. In some of our numerical experiments, artifacts can be observed developping in the outer areas during the early stages of the reconstruction process.

We apply time-harmonic dipole sources of the form (2.5) with frequencies of $f = 5, 10, 15, 20, 25$, or 30 MHz. In our examples, this corresponds to wavelengths in the background medium between 2 meters for $f = 30$ MHz and 13 meters for $f = 5$ MHz. The size of an individual grid cell is chosen such that each of these wavelengths is sampled by at least 16 pixels in order to avoid numerical artifacts due to undersampling.

The data in our numerical examples are generated by running the FDFD forward modelling code on the correct permittivity and conductivity distributions. Using the same forward code for creating the data and for doing the reconstruction is usually called 'inverse crime'. Therefore, to make sure that the situations we model in our experiments are as realistic as possible, we have tested the forward modelling code thoroughly, and add Gaussian noise with signal-to-noise ratios of up to 10 dB to the real and imaginary parts of the generated data.

5.2. A full-view example. Our first numerical example demonstrates the performance of the STAF algorithm as well as the levelART algorithm in a situation where a relatively complicated shape has to be recovered from data which correspond to receiver positions completely surrounding the domain of interest. All data are noise-free in this idealized example. The geometry is shown in Figure 1. The positions of the sources and receivers are indicated by dots in the Figure. The background medium in this example consists of a homogeneous conductivity distribution $\sigma_b = 3.0 \times 10^{-4}$ Siemens/m, and a homogeneous permittivity distribution $\epsilon_{\mathrm{b}} = 20$. Inside the object, the permittivity is $\hat{\epsilon} = 15$, having a moderate contrast to the background distribution. Notice that an interesting feature of this geometry is the 'hole' in the body of the object which is difficult to reconstruct.

First we apply the STAF reconstruction scheme by using the data corresponding to one fixed frequency only. We do this for six different frequencies, namely 5, 10, 15, 20, 25 and 30 MHz. The constant C_l in the reconstruction scheme (see Table 2) which determines the shapes is chosen to be 0.7 times the maximal value of the reconstructed function $|\tilde{\epsilon}_s(x)|$. Figure 2 shows the different reconstructions. As it was expected the resolution of the reconstructions increases if we go to higher frequencies. The lowest frequency (5 MHz) only gives us a large blop at the correct position of the sought object. At $f = 15$ MHz the triangular shape of the roof of the house becomes visible, and the hole in the bottom of the house is correctly recovered at a frequency of $f = 25$ MHz. The highest frequencies ($f = 25$ MHz

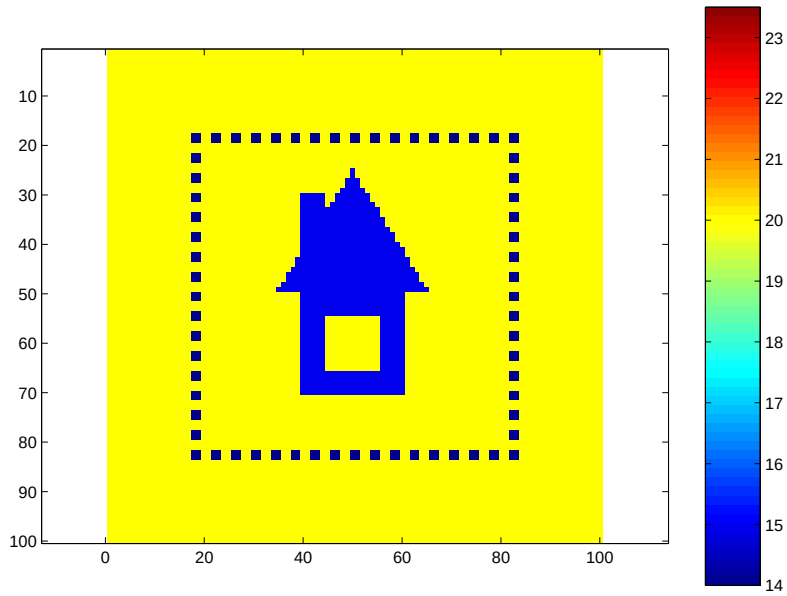


FIGURE 1. Original object for the example with full view. The dots in the figure indicate the source and receiver positions. The permittivity in the background is $\epsilon_b = 20$, and in the object $\hat{\epsilon} = 15$.

and $f = 30$ MHz) give us a quite decent reconstruction of the correct shape of the unknown object.

Figure 3 demonstrates how the levelART algorithm performs in this situation when we use the lowest frequency (5 MHz) STAF reconstruction as an initial guess. We use three groups of frequencies for this reconstruction (see Table 1), each of them containing only one single frequency, namely $G_1 = \{10 \text{ MHz}\}$, $G_2 = \{20 \text{ MHz}\}$, and $G_3 = \{30 \text{ MHz}\}$. We start the reconstruction with the lowest frequency $f = 10$ MHz and perform $I_1 = 30$ sweeps as described in Table 1. Continuing from that result, we perform $I_2 = 30$ sweeps with $f = 20$ MHz, and finally $I_3 = 30$ sweeps with $f = 30$ MHz. We believe that marching in this way over the frequencies, starting with the lower frequencies and continuing with the higher frequencies, is helpful in dealing with the strong nonlinearity (high contrast) of the problem, since the lower frequencies seem to be less sensitive to the high contrast in the parameters than the higher frequencies are. This is in agreement with related observations reported earlier by Chen [**4**].

Comparing the final result of the levelART algorithm with the STAF reconstructions in Figure 2, we conclude that the iterative levelART scheme can be used to improve each of the STAF reconstructions.

5.3. A cross-borehole situation with multiple objects. In our second numerical example, we consider a more realistic situation which is typical for geophysical applications. Comparable situations occur for example when we wish to monitor pollutant plumes at environmental cleanup sites from cross-borehole EM data. The geometry of this second example is shown in Figure 4.

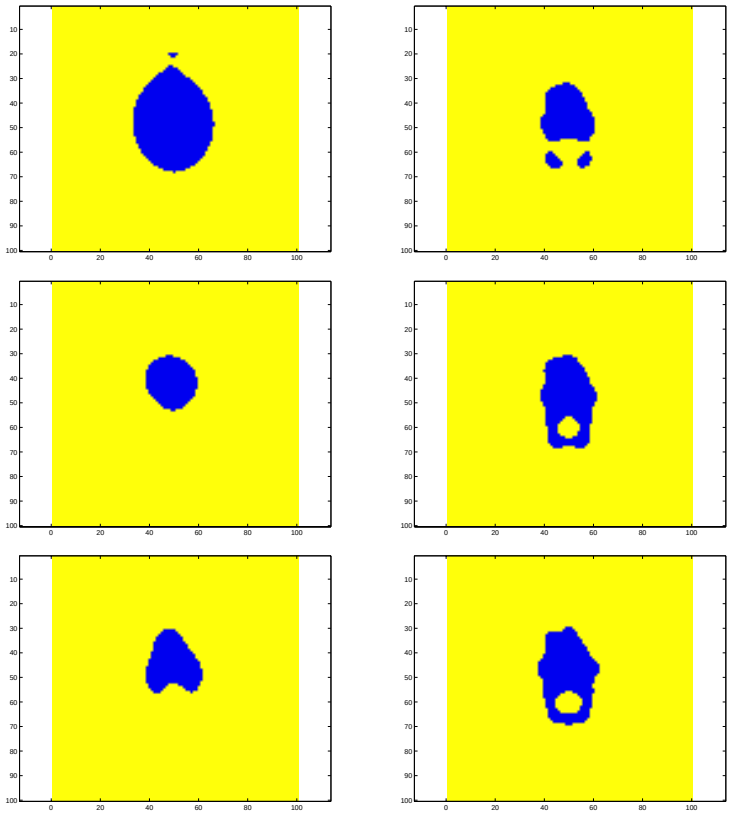


FIGURE 2. Different STAF reconstructions for the full-view example using data with a fixed frequency. Left column from top to bottom: $f = 5$ MHz, $f = 10$ MHz, $f = 15$ MHz; Right column from top to bottom: $f = 20$ MHz, $f = 25$ MHz, $f = 30$ MHz. The data are noise-free.

The background permittivity distribution in this example consists of four tilted layers with values of $\epsilon_{\rm b} = 21$ in the top layer, and then continuing downwards with 20 , 19, and again 21 for the deepest layer. The conductivity distribution σ_b is homogeneous with a value of $\sigma_b = 3.0 \times 10^{-4}$ S/m everywhere.

Embedded in this background are three compact inclusions as it can be seen in Figure 4. The permittivity inside these inclusions is $\hat{\epsilon} = 5$, having a high contrast to the background values. The three inclusions are oriented such that there are two 'channels' of background material between them, one of them in the vertical and one in the horizontal direction. The difficulty in this example is to separate the three inclusions from each other from the limited-view data. In particular, the reconstruction of the vertical channel is critical since we expect that the resolution in the horizontal direction will suffer from the missing data.

As in the full-view example, we first apply the STAF resonstruction scheme by using only the data corresponding to one fixed frequency. We do this for each of the six frequencies 5, 10, 15, 20, 25, and 30 MHz. Again, the constant C_l in the

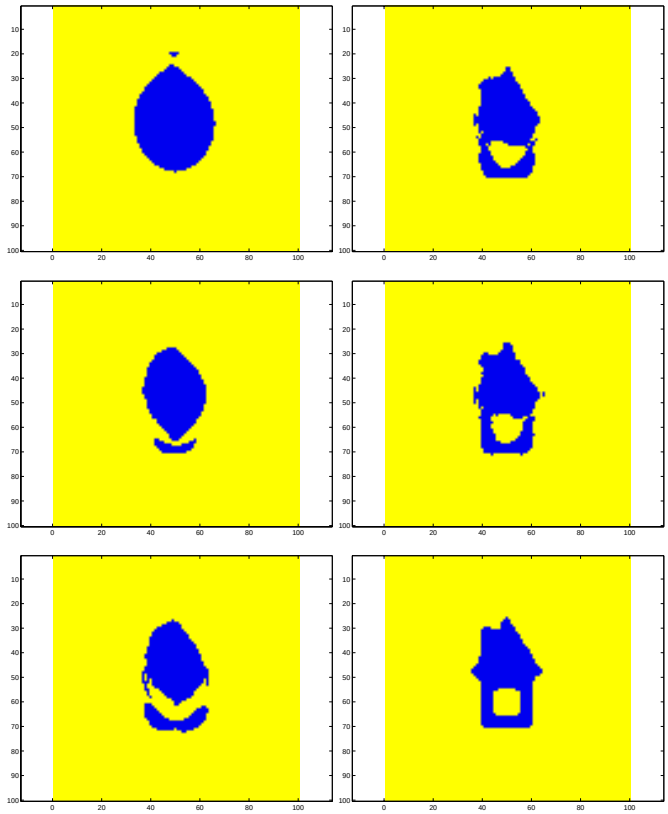


FIGURE 3. The levelART process: Evolution of permittivity $\epsilon^{(n)}$. Left column from top to bottom: STAF reconstruction of $\epsilon^{(0)}$ for 5 MHz (top left); This is the starting guess for the following reconstruction using levelART. After 10 steps of levelART with 10 MHz; After 30 steps with 10 MHz; Right column from top to bottom: After 10 more steps with 20 MHz; After 30 steps with 20 MHz; Final reconstruction after 30 steps of levelART with 30 MHz (bottom right). The algorithm used noise-free data.

reconstruction scheme is chosen to be 0.7 times the maximal value of the reconstructed function $|\tilde{\epsilon}_s(x)|$. Figure 5 shows the different reconstructions in the case where the data are noise-free, and Figure 6 shows the corresponding reconstructions in the case where the data are contaminated by additive Gaussian noise with a signal-to-noise ratio (SNR) of 10 dB.

We observe in the Figures 5 and 6 that, in contrast to our first example, the STAF reconstructions do *not* improve with increasing frequency in this limited-view situation. The reconstruction corresponding to the lowest frequency (5 MHz) is similar to the reconstruction which we get in the full-view situation, which means that the lower frequencies are less sensitive to the large amount of missing data. When we increase the frequency, the reconstructions look like deteriorated images which are oscillatory in the horizontal direction, due to the missing data on the

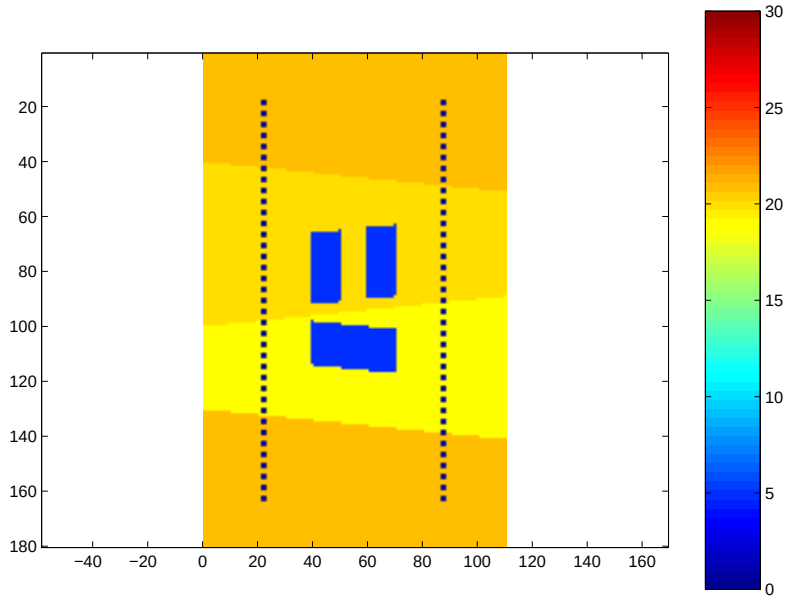


FIGURE 4. True permittivity distribution in the cross-borehole example. The dots in the figure indicate the source and receiver positions. The permittivity in the background layers is (from top to bottom) $\epsilon_b = 21$, 20, 19, and 21. Inside the object it is $\hat{\epsilon} = 5$.

top and on the bottom of the reconstruction domain. In the case of noisy data in Figure 6, the higher frequency reconstructions look even worse due to additional blurring which is caused by the noise in the data. In short, the figures show that the sensitivity of the STAF reconstructions to limited view and noise in the data increases with increasing frequency.

Combining the observations so far we conclude that low frequency data are fairly insensitive to noise in the data and to limited view, but the corresponding reconstructions have low resolution. On the other hand, the high frequency data yield in the ideal situation of complete data sets high resolution reconstructions, but these reconstructions deteriorate dramatically when only noisy limited-view data are available.

In the levelART approach we try to make use of the favourable features of the low frequency data as well as of the high frequency data, but in a way such that the unwanted features of each of them are compensated for by the other one. In order to achieve this goal, we define two (or in general even more) groups of frequencies (see Table 1), namely $G_1 = \{15 \text{ MHz}\}$, and $G_2 = \{20, 25, \text{and } 30 \text{ MHz}\}$. As initial guess we use the STAF reconstruction of $f = 5$ MHz. We make this choice because of our observation that the lower frequency reconstructions are less sensitive to the limited-view geometry and to the noise in the data. Now we perform $I_1 = 20$ sweeps of the levelART algorithm with the frequency 15 MHz, corresponding to the first frequency group. Since we have already a good initial guess available, the nonlinearity in the problem has been reduced and it causes no problems to use that frequency in this stage of the reconstruction process.

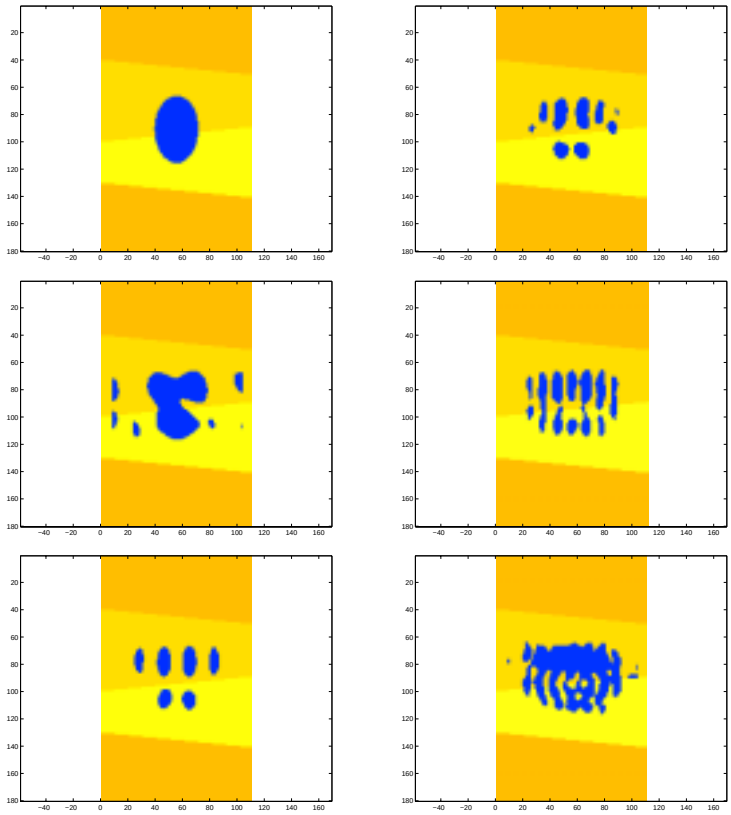


FIGURE 5. Different STAF reconstructions for the cross-borehole example using limited-view data with a fixed frequency. Left column from top to bottom: $f = 5$ MHz, $f = 10$ MHz, $f = 15$ MHz; Right column from top to bottom: $f = 20$ MHz, $f = 25$ MHz, $f = 30$ MHz. The data are noise-free.

We see in Figure 7 that after 20 sweeps with 15 MHz we almost have the three reconstructed objects separated from each other. However, it can be observed that it becomes more and more difficult for the algorithm to improve the reconstruction with only the data corresponding to this frequency. Therefore, we continue with the second frequency group, and perform $I_2 = 30$ sweeps with the three frequencies 20, 25, and 30 MHz. Each of these sweeps uses first the lowest frequency (20 MHz) data for calculating one update for the shape, then uses the next higher frequency (25 MHz) data for the next update, and finally the highest frequency (30 MHz) data for the final update of the given sweep. The succeeding sweep starts again with the lowest frequency (20 MHz) data of that group G_2.

We believe that, in this combination, the updates from the lower frequency data in the given group act as regularizing and stabilizing elements in the combined reconstruction process, and compensate for the deteriorating effects due to the missing data and the noise in the highest frequencies. The resolution of the final reconstruction nevertheless seems to be mainly determined by the highest frequency data in the group, as it can be seen in Figure 7.

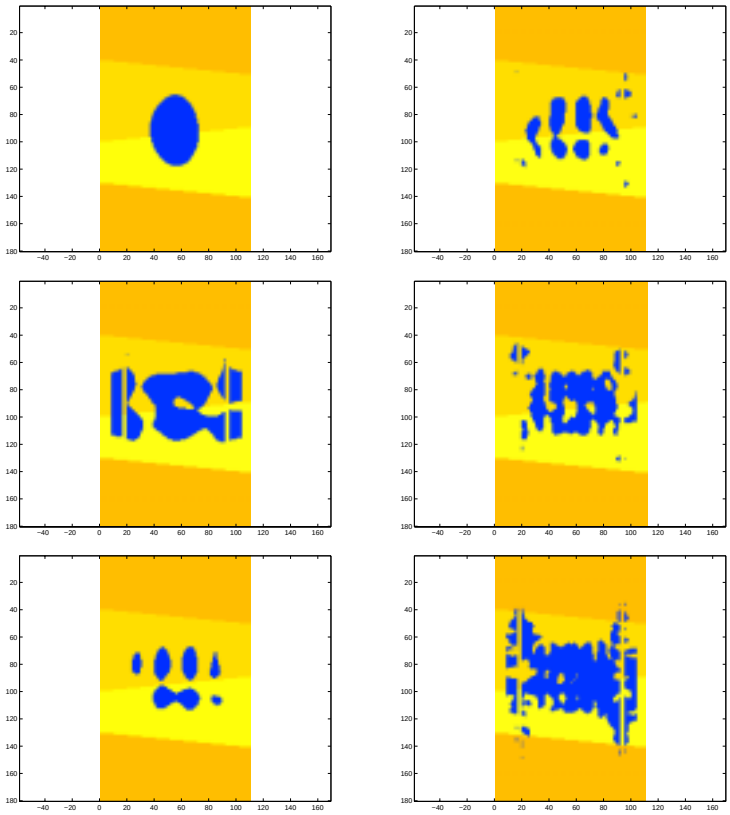


FIGURE 6. Different STAF reconstructions for the cross-borehole example using limited-view data with a fixed frequency. Left column from top to bottom: $f = 5$ MHz, $f = 10$ MHz, $f = 15$ MHz; Right column from top to bottom: $f = 20$ MHz, $f = 25$ MHz, $f = 30$ MHz. The data are noisy with a signal-to-noise ratio of 10 dB.

The regularizing effect of lower frequency data becomes even more visible in the following experiment. Here, we want to find out whether it would be sufficient just to run the levelART algorithm with data corresponding to only one of the higher frequencies and using the STAF reconstruction with 5 MHz as an initial guess. We use the same geometry as in the above cross-borehole experiment, but do not add any noise to the data. Starting with the STAF result for 5 MHz as displayed in Figure 8, we perform for each of the frequencies 15 MHz, 20 MHz, 25 MHz, and 30 MHz 110 sweeps with levelART using only that single frequency. No additional regularization (as for example 'curve shortening by diffusion') is applied. The results we get are displayed in Figure 8. Then, to compare with, we do an experiment where we first run levelART using the data with frequency 15 MHz 40 times, and then run 25 additional sweeps using the data with frequencies 20 MHz, 25 MHz, and 30 MHz repeatedly in a cyclic order. In this way, all experiments are comparable with each other in terms of total computational cost. The result of this last experiment is also displayed in Figure 8.

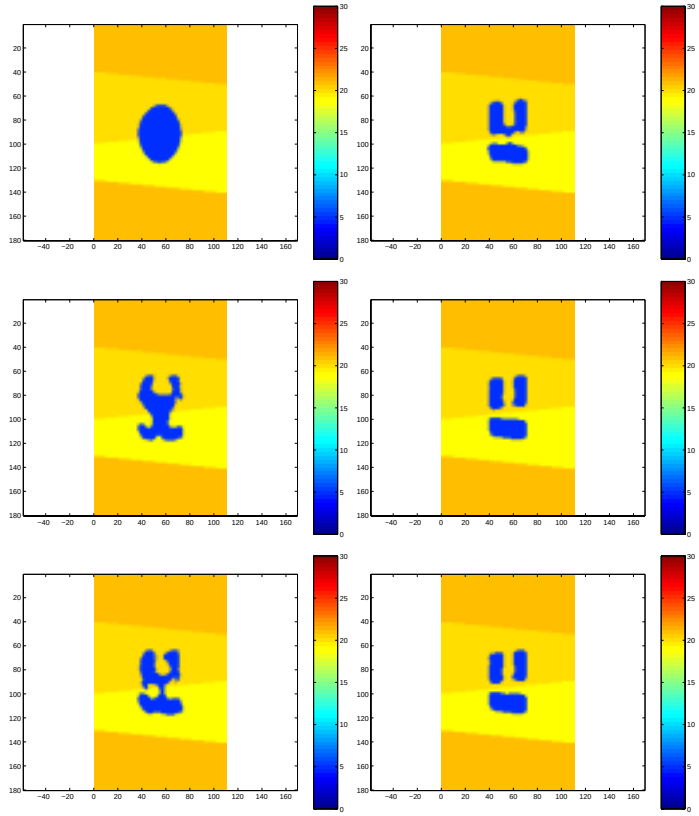


FIGURE 7. The levelART process: Evolution of permittivity $\epsilon^{(n)}$. Left column from top to bottom: STAF reconstruction $\epsilon^{(0)}$ for 5 MHz (top left); This is the starting guess for the following reconstruction using levelART. After 10 steps of levelART with 15 MHz; After 20 steps with 15 MHz; Right column from top to bottom: After 10 more sweeps with $20+25+30$ MHz; After 20 sweeps with $20+25+30$ MHz; Final reconstruction after 30 sweeps of levelART with $20+25+30$ MHz (bottom right). The algorithm used noisy data with a SNR of 10 dB and curve shortening by diffusion.

These results show that the reconstructions using only the data corresponding to one single frequency are quite useful as long as this frequency is not too high. The reconstructions for 15 MHz and 20 MHz data clearly approximate the correct geometry of the three objects, but these objects are not yet completely separated from each other. In addition, the resolution in these reconstructions is still not very good. Increasing the frequency should improve the resolution, but the figures corresponding to 25 MHz and 30 MHz show that instead the algorithms break down in the sense that the reconstructions get trapped far away from the correct ones. The combined reconstruction in the bottom right image of Figure 8, on the other hand, behaves perfectly stable and shows an improved resolution due to the incorporated higher frequency data.

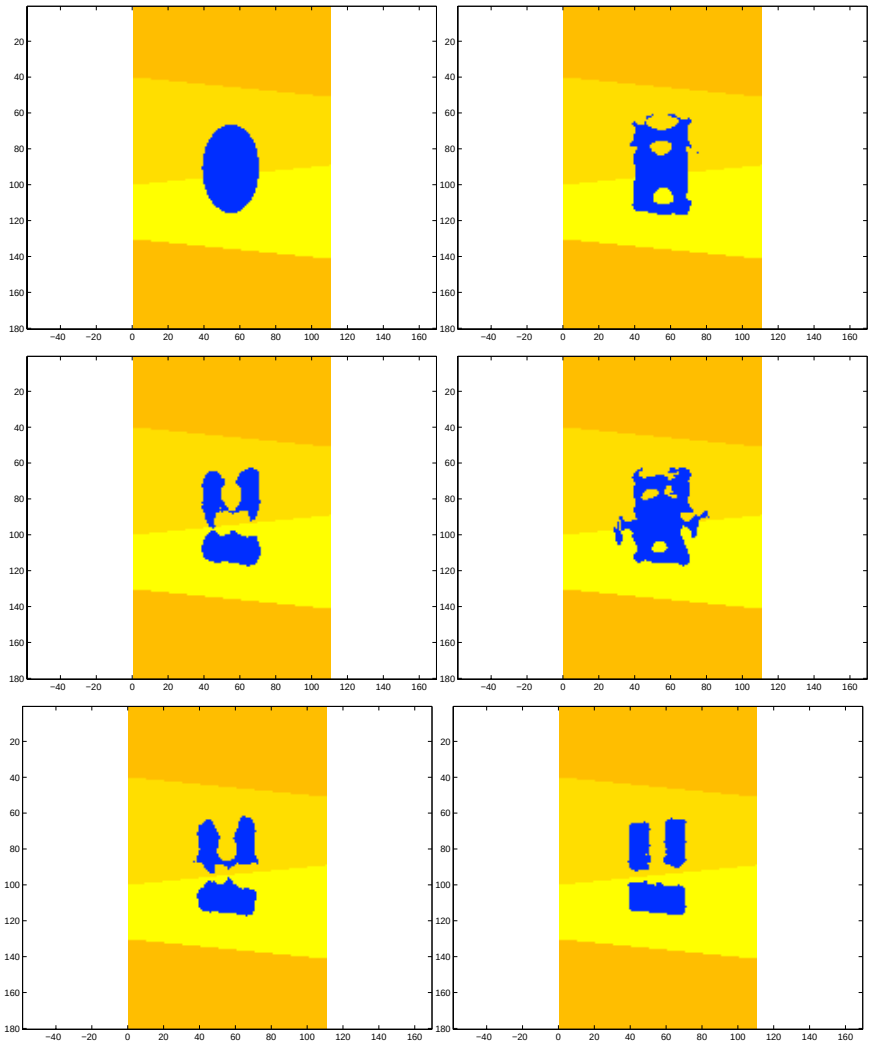


FIGURE 8. The levelART algorithm: Final reconstructions of permittivity ϵ. Left column from top to bottom: STAF reconstruction $\epsilon^{(0)}$ for 5 MHz (top left), which is the starting guess for each of the following reconstructions: final reconstruction after 110 iteration steps using only 15 MHz data; using only 20 MHz data. Right column from top to bottom: final reconstruction after 110 steps using only 25 MHz data; using only 30 MHz data; bottom right: final reconstruction after first using 40 steps with 15 MHz data, and then 25 more sweeps using 20, 25 and 30 MHz data in a cyclic order. The data are noise-free and no additional regularization is applied.

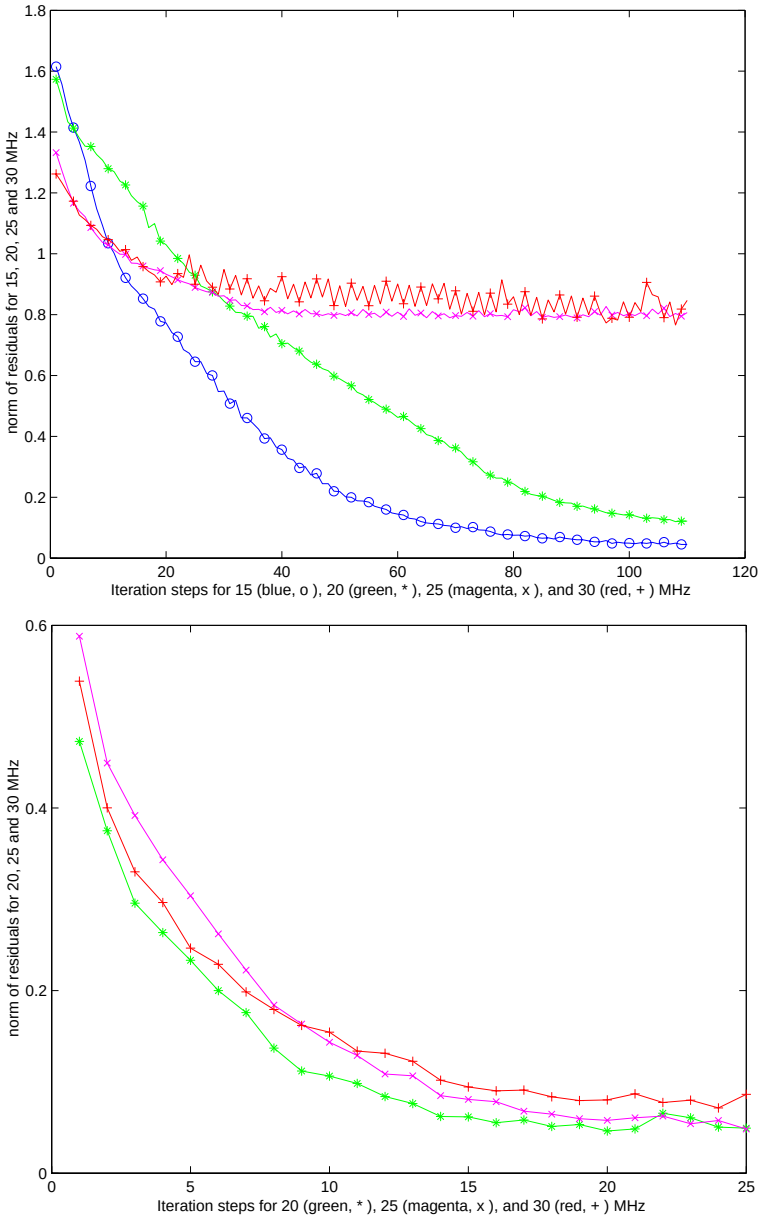


FIGURE 9. Top: norms of residuals in the reconstructions which use only one single frequency (110 iterations). Bottom: norms of the residuals in the reconstruction which starts with the frequency 15 MHz (40 iterations, not shown in the figure), and then continues in a cyclic order with 20, 25 and 30 MHz data, using each of these exactly 25 times.

In Figure 9 we display the evolution of the norms of the residuals during the reconstruction process. When using only data corresponding to one low frequency (15 MHz and 20 MHz), the residuals decrease continuously and the reconstruction is performed stably. However, when higher frequency data (25 MHz and 30 MHz) are used instead, the residuals get 'trapped' on their way and start oscillating around some intermediate value. On the other hand, when using a combination of lower and higher frequency data, all residuals (including those corresponding to higher frequencies) decrease continuously until the reconstruction task is completed.

We want to mention that other choices for building frequency groups are certainly possible, and might even yield better results than those presented here. So far it is not clear whether there is an 'optimal' way of selecting these groups.

6. Summary and future directions

We have investigated the use of multifrequency electromagnetic data for shape reconstruction in a limited-view source-receiver geometry. We have compared the results of two recently introduced shape reconstruction methods, namely a very simple one-step source type method (STAF) on the one hand, and an iterative method (levelART) on the other hand which uses adjoint fields and level sets for the inversion. It was shown that, in the full view situation, the STAF algorithm yields low resolution reconstructions if applied to the data corresponding to the lower frequencies, and decent higher resolution reconstructions if applied to the higher frequency data. The reconstructions from the lower frequency data are fairly insensitive to noise and missing data in a limited-view geometry, whereas the higher frequency reconstructions deteriorate dramatically in that situation.

The levelART algorithm, if applied with a STAF reconstruction as initial guess, is able to yield improved reconstructions in all cases. It uses the lower frequency data for stabilizing and regularizing the reconstruction process, and for overcoming the problems caused by the noisy data and the high contrasts in the perturbations. Furthermore, in all examples which we have considered so far, the resolution of the final reconstructions is determined by the highest frequency data which are used for the inversion. Some additional regularization (for example 'curve shortening by diffusion') might be useful in order to increase the smoothness of the reconstructed boundaries. This feature of the levelART algorithm is discussed elsewhere [**12**].

So far we have only considered a somehow idealized 2D situation. However, we believe that the generalization of the method to a more realistic 3D situation is possible and doable. This is an interesting problem for future research.

References

[1] H. Abdullah and A. K. Louis *The approximate inverse for solving an inverse scattering problem for acoustic waves in an inhomogeneous medium*, Inverse Problems **15**, 1999, 1213–1229.

[2] T. S. Angell, W. Jiang and R. E. Kleinman *A distributed source method for inverse acoustic scattering*, Inverse Problems **13**, 1997, 531–46.

[3] S. Caorsi and G. L. Gragnani *Inverse-scattering method for dielectric objects based on the reconstruction of the nonmeasurable equivalent current density*, Radio Science **34** (1), 1999, 1–8.

[4] Y. Chen *Inverse scattering via Heisenberg's uncertainty principle*, Inverse Problems **13**, 1997, 253–282.

[5] W. C. Chew, Y. M. Wang, G. Otto, D. Lesselier, and J. Ch. Bolomey *On the inverse source method of solving inverse scattering problems*, Inverse Problems **10**, 1994, 547–553.

[6] D. L. Colton and R. Kress *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd edn, Berlin, Springer, 1998.

[7] A. J. Devaney and G. C. Sherman *Nonuniqueness in inverse source and scattering problems*, IEEE Trans. Antennas Propag. **30**, 1984, 1034–1037.

[8] A. J. Devaney *Reconstructive tomography with diffracting wavefields*, Inverse Problems **2**, 1986, 161–83.

[9] K. A. Dines and R. J. Lytle *Computerized geophysical tomography*, Proc. IEEE **67**, 1979, 1065–1073

[10] O. Dorn *A transport-backtransport method for optical tomography*, Inverse Problems **14**, 1998, 1107–1130

[11] O. Dorn, H. Bertete-Aguirre, J. G. Berryman and G. C. Papanicolaou *A nonlinear inversion method for 3D electromagnetic imaging using adjoint fields*, Inverse Problems **15**, 1999, 1523–1558.

[12] O. Dorn, E. L. Miller and C. M. Rappaport *A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets*, Inverse Problems **16** (5), 2000, 1119–1156.

[13] T. M. Habashy, R. W. Groom and B. R. Spies *Beyond the Born and Rytov Approximations: A Nonlinear Approach to Electromagnetic Scattering*, Journal of Geophysical Research **98** (B2), 1993, 1759–1775.

[14] T. M. Habashy, M. L. Oristaglio and A. T. de Hoop *Simultaneous nonlinear reconstruction of two-dimensional permittivity and conductivity*, Radio Science **29** (4), 1994, 1101–1118.

[15] S. Helgason *The* Radon Transform, Birkhauser, 1980.

[16] F. Hettlich and W. Rundell *Recovery of the support of a source term in an elliptic differential equation*, Inverse Problems **13**, 1997, 959–976.

[17] K. T. Ladas and A. J. Devaney *Iterative methods in geophysical diffraction tomography*, Inverse Problems **8**, 1992, 119–132.

[18] A. Litman, D. Lesselier and F. Santosa *Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set*, Inverse Problems **14**, 1998, 685–706.

[19] E. Miller, M. Kilmer and C. Rappaport *A New Shape-Based Method for Object Localization and Characterization from Scattered Field Data*, submitted to IEEE Trans. Geoscience and Remote Sensing.

[20] F. Natterer *The* Mathematics of Computerized Tomography, Stuttgart, Teubner, 1986.

[21] F. Natterer and F. Wübbeling *A propagation-backpropagation method for ultrasound tomography*, Inverse Problems **11**, 1995, 1225–1232.

[22] F. Natterer *Numerical Solution of Bilinear Inverse Problems*, Preprints "Angewandte Mathematik und Informatik" 19/96-N, 1996, Münster. This preprint is electronically available at http://wwwmath.uni-muenster.de/math/num/Preprints/1997/natterer_3.

[23] S. Osher and J. Sethian *Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations*, Journal of Computational Physics **56**, 1988, 12–49.

[24] C. M. Rappaport, M. Kilmer and E. Miller *Accuracy Considerations in Using the PML ABC with FDFD Helmholtz Equation Computation*, Journal of Numerical Modelling (in press).

[25] C. Rozier, D. Lesselier, D. Angell and R. E. Kleinman *Shape retrieval of an obstacle immersed in shallow water from single frequency fields using a complete family method*, Inverse Problems **13**, 1996, 487–508.

[26] F. Santosa *A Level-Set Approach for Inverse Problems Involving Obstacles*, ESAIM: Control, Optimization and Calculus of Variations **1**, 1996, 17–33.

[27] J. A. Sethian *Numerical algorithms for propagating interfaces: Hamilton-Jacobi equations and conservation laws*, J. Diff. Geom. **31**, 1990, 131–61.

[28] J. A. Sethian *Level Set Methods and Fast Marching Methods*, 2nd ed., Cambridge University Press, 1999.

[29] L. Souriau, B. Duchene, D. Lesselier and R. E. Kleinman *Modified gradient approach to inverse scattering for binary objects in stratified media*, Inverse Problems **12**, 1996, 463–481.

[30] P. M. Van den Berg and R. E. Kleinman *A contrast source inversion method*, Inverse Problems **13**, 1997, 1607–20.

[31] P. M. Van den Berg, A. L. van Broekhoven and A. Abubakar *Extended contrast source inversion*, Inverse Problems **15**, 1999, 1325–1344.

[32] A. von Hippel *D*ielectric Materials and Applications, Wiley, New York, 1953, 3–4.

Dept of Computer Science, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada
E-mail address: dorn@cs.ubc.ca

Center for Subsurface Sensing and Imaging Systems, Northeastern University, Boston, MA 02115, USA
E-mail address: elmiller@ece.neu.edu

Center for Subsurface Sensing and Imaging Systems, Northeastern University, Boston, MA 02115, USA
E-mail address: rappaport@neu.edu