

Adaptive parameterization of electric conductivity in inversion of electromagnetic data

Master's thesis in applied mathematics

by

Torbjørn Ek

June 1, 2012



Department of Mathematics
University of Bergen
Norway

Acknowledgements

I want to thank my supervisors Inga Berre and Martha Økland Lien for all their help, inspiring discussions and the enthusiasm they showed during this work. I have truly enjoyed the whole process. To my fellow students, thank you for creating a social and positive environment during the last five years. I especially thank Silje and Håkon for our discussions and memorable times together in room 620. Finally, I want thank my family.

*Torbjørn,
June 2012*

Abstract

We describe a methodology developed for 3-D parameter identification, with focus on large-scale applications such as monitoring subsea oil production and geothermal systems [7]. The methodology is designed to handle challenges related to low parameter sensitivity, nonuniqueness of the inverse solutions and costly numerical calculations. A reduced, composite parameter representation is chosen to meet these challenges. Our contributions to the methodology involves choosing a reduced representation with radial basis functions, to maintain a low number of parameters. We also propose the use of a 1. order selection measure in the refinement process to reduce the computational costs. The performance of the proposed changes in the methodology is illustrated in a series of examples for estimating the change in electric conductivity from time-lapse electromagnetic observations. The results show some limitations regarding the accuracy of the 1. order selection measure. For the investigated numerical examples, radial basis functions, together with the described methodology, effectively provide an estimated of the electric conductivity field using electromagnetic measurements.

Contents

Outline and motivation	1
1 Controlled source electromagnetics	3
1.1 Model parameters	3
1.2 Basic overview	4
1.3 Maxwell's equations	5
1.4 The Integral Equation method	7
1.4.1 Solution to Maxwell's equations	8
1.5 Numerical solution to the Fredholm integral equation	8
1.5.1 Numerical accuracy with respect to grid size	10
2 Inverse problems	13
2.1 Solving an inverse problem	14
2.2 Regularization	15
3 Parameter representation and adaptive refinement	17
3.1 Composite parameter representation	17
3.1.1 Exterior functions	17
3.1.2 Interior function	18
3.2 Basis functions	18
3.2.1 Trilinear basis functions	19
3.2.2 Radial basis functions	19
3.3 Multi-level estimation strategy	20
3.4 Adaptive refinement and selection measure	21
3.4.1 Predicted attainable objective function values	23
3.4.2 Objective function gradient steepness	24
3.4.3 The set of potential basis functions for refinement	26
3.5 Scaling of basis functions	27
4 Optimization and sensitivity calculations	31
4.1 Optimization	31
4.1.1 The Gauss-Newton method	32
4.1.2 The Levenberg-Marquardt method	34

4.2	Termination criterion	34
4.3	Sensitivity matrix	35
4.3.1	Numerical calculation of the sensitivity matrix	36
4.3.2	Quasi-analytical approximation	37
5	Numerical examples	39
5.1	Example setup	39
5.2	Scaling of basis functions	40
5.3	A multi-level strategy compared to a one-level strategy	41
5.4	Comparison of different basis functions in parameter representation	42
5.4.1	Example 1	43
5.4.2	Example 2	43
5.4.3	Example 3	47
5.4.4	Remarks	47
5.5	Comparison of first order and second order selection measure	49
5.5.1	Example 4	49
5.5.2	Example 5	51
5.5.3	Remarks	54
6	Summary and further work	55

Outline and motivation

Before and during reservoir production, knowledge of the reservoir characteristics is vital. The more knowledge we have, the better our choices will be, leading to a more efficient production process. It may concern placement of wells prior to production, or monitoring the reservoir during production.

Controlled source electromagnetic (CSEM) is a method that can help provide a more accurate description of a reservoir. CSEM data is sensitive to spatial changes in the electric conductivity in the reservoir. The electric conductivity of a saturated porous medium is partly dependent on the saturating fluid. Therefore, the method has the ability to distinguish between parts of a reservoir which is saturated with water and hydrocarbons respectively. This means that the method is promising for monitoring the interface of the injected water and oil during secondary production.

The aim of this thesis is to present a robust methodology for the inversion of CSEM data, in order to identify the electric conductivity field in the reservoir.

In Chapter 1 we present the method of controlled source electromagnetic, what type of information it provides us with, and what challenges that are associated with this method. Then we state the governing equations for the system, and describe the solution using the Integral Equation method. We present a numerical solution to the Fredholm integral equation, and finally investigate how the size of the computational grid effects the solution of the forward problem.

In Chapter 2 we describe the concept of inverse problems. We discuss why inverse problems typically are harder to solve than forward problems, and discuss the concept of ill-posedness. Regularization is presented as a strategy when solving ill-posed problems. Finally we present the regularization methods that we apply in our estimation process.

In Chapter 3 we parameterize the electric conductivity in the reservoir, and motivate our choice of a reduced representation. The choice of a composite parameterization with an exterior function and an interior function is motivated. The interior function consists of interior coefficients and basis functions. We apply two types of basis functions; trilinear basis functions and radial basis functions. A multi-level strategy for solving the parameterization problem is then presented, and we motivate an adaptive refinement strategy. For the adaptive refinement strategy we describe two different selection measures; a 1. order gradient based selection measure, and a 2. order selection measure based on an approximation of the Hessian matrix. We conclude the chapter by discussing different scalings of basis functions, and how scaling the basis functions may effect the parameter representation.

In Chapter 4 we present two basic ideas in optimization methodology, the *line-search* strategy, and the *thrust-region* strategy. We describe the line-search method of *Gauss-Newton*, before we replace it with the thrust-region method of *Levenberg-Marquardt*. We present our choice of solution criterion, as well as calculations associated with solving the sensitivity matrix.

In Chapter 5 numerical examples are presented. First, we consider the effect of scaling of basis functions on the chosen parameter representation, and then evaluate the use of a multi-level estimation strategy. Our two main results are presented next, which concern comparison of trilinear and radial basis functions, as well as comparison of the two selection measures (in the refinement process) presented in Chapter 3.

In Chapter 6 we give a summary of the results from Chapter 5, what consequences these results have, and consider some new questions that have emerged during this work.

Chapter 1

Controlled source electromagnetics

Controlled source electromagnetics (CSEM) has been in commercial use for more than a decade. Today it is regarded as an important geophysical tool in marine environments. The method is sensitive to changes in electric resistivity, which e.g. take place during the transition from water-filled sediments to more resistive oil-filled sediments. CSEM can be used both in exploration and reservoir monitoring. Although the resolution power of the data is not as high as for seismic methods, it provides the possibility of better reservoir estimates compared to methods such as gravity and magnetic surveying [8].

1.1 Model parameters

The properties of a reservoir depends on a set of physical quantities which we call model parameters. Before we go into details on the CSEM method, we present some important model parameters in this section.

Porosity

A porous medium is composed of solid material, called the matrix V_M , and void space V_P . The total volume is denoted by the bulk space $V_B = V_M + V_P$, and the porosity of a porous medium is defined as the volume fraction of the void space to the bulk space,

$$\Phi = \frac{V_P}{V_B}.$$

Normal methods to measure the porosity are based on the ability of a fluid to saturate the medium, so only interconnected pores are registered. Common practice is to consider the effective porosity, for which V_P is replaced by the volume of the interconnected pores.

Saturation

Consider a test volume of a porous medium, where at least one fluid phase is present, such as water, gas and oil. The fraction of pore volume occupied by a fluid phase α , is known

as the saturation of that phase;

$$S_\alpha = \frac{V_\alpha}{V_P}.$$

Given that there are n phases present and the pores are fully saturated, the following identities hold

$$\sum_{\alpha=1}^n S_\alpha = 1, \quad \sum_{\alpha=1}^n V_\alpha = V_P.$$

Electric conductivity

The ability of a medium to conduct electromagnetic signals is described both by electric conductivity, σ , and its reciprocal, the electric resistivity which is denoted by $\rho = \sigma^{-1}$. The units for these quantities are Siemens per meter (Sm^{-1}) and Ohm-meters (Ωm) respectively. The electric conductivity arises in Ohm's law;

$$\mathbf{J} = \sigma \mathbf{E},$$

where \mathbf{J} is current density and \mathbf{E} is the electric field intensity. For a saturated porous medium the electric conductivity depends mainly on the porosity of the medium, and the saturating fluid [1].

1.2 Basic overview

Figure 1.1 illustrates the marine CSEM method. The target reservoir is located at a certain depth below the seafloor. It consists of a distribution of water- and oil-saturated rocks. The method seeks to discover this distribution.

A set of electromagnetic (EM) field receivers are deployed on the seafloor. During a survey these receivers continuously record the electric and magnetic fields. These fields are originally broadcasted by a 50 – 300 meters long horizontal antenna that is towed approximately 20–100 meters above the seafloor. The EM signals generated by the antenna is spread in all directions - into the overlying water column, directly to the receiver, and downward into the seabed. By towing the antenna close to the seafloor the coupling with seafloor rocks and sediments is maximized [14].

As mentioned, the CSEM data is sensitive to changes in electric resistivity. Water-filled sediments represent good conductors with a resistivity value in the range 0.5 – 2 Ωm . Hydrocarbon-filled sediments are more resistive (30-500 Ωm) [13]. For heterogeneous rocks EM signals will be sensitive to variations in porosity and saturation. Furthermore, domains with high electric resistivity are not necessarily hydrocarbon-filled sediments. Different types of volcanic rocks, salts and carbonates represents resistive bodies that scatter EM signals. Such disturbances are known as *false positives* [10].

To reach potential targets as deep as 3000 meters below the seafloor, it is necessary to use frequency values in the range 0.1 – 1 Hz. With such low frequencies, the quality of the data makes it difficult to obtain a high resolution estimates. We can therefore not

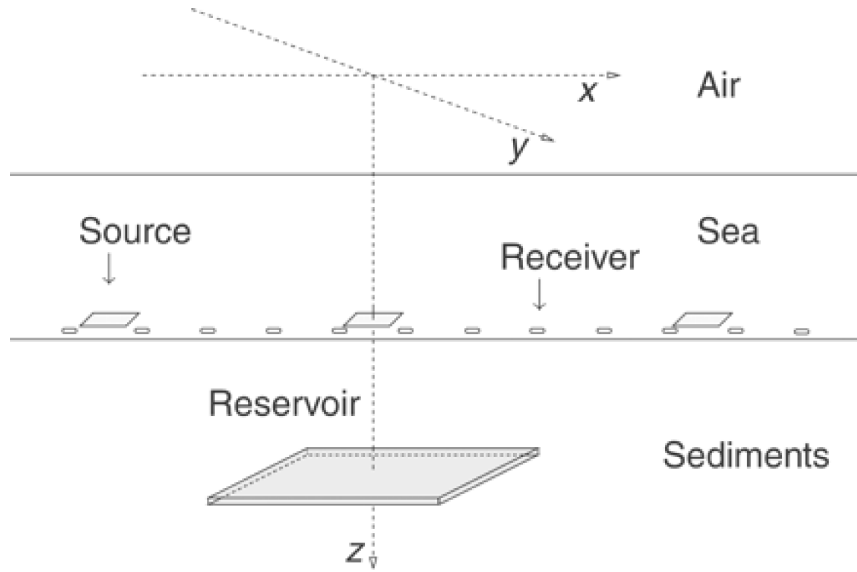


Figure 1.1: Sketch of a CSEM monitoring situation. Receivers are deployed at the seafloor, and records the electric and magnetic fields as the source emits electromagnetic waves. The data collected are thus a combination of EM waves directly from the source, and EM waves due to the reflection and scattering in oil-saturated and water-saturated sediments. Figure from [13]

expect any detailed structures to appear in our estimate, more than on a coarse-scale. Although this is a drawback, information on a coarse-scale is in many cases valuable in the reservoir characterization process. Seismic surveying mainly results in higher resolution, but seismics are sensitive to the elastic properties of the sediments, which are similar for saturating oil and water. Thus seismic signals distinguish gas and fluid with high resolution estimates, whilst CSEM is complimentary with its ability to distinguish between fluids.

In our work we have considered a monitoring setting, where a reservoir is injected with water through a well to increase the reservoir pressure. This is known as *secondary production*, as opposed to *primary production* where the oil is pumped out without any pressure support. The aim is to use CSEM to monitor the increasingly water-filled domain. As the EM signals are broadcasted by the antenna, the seafloor receivers start recording. How these recordings are coupled to the electric conductivity field in the reservoir is presented in the next section.

1.3 Maxwell's equations

In problems where electromagnetism plays a part, Maxwell's equations will typically arise. For situations with time variation $e^{-i\omega t}$ and a monochromatic field¹, Maxwell's equations

¹Monochromatic field - The frequency ω is constant

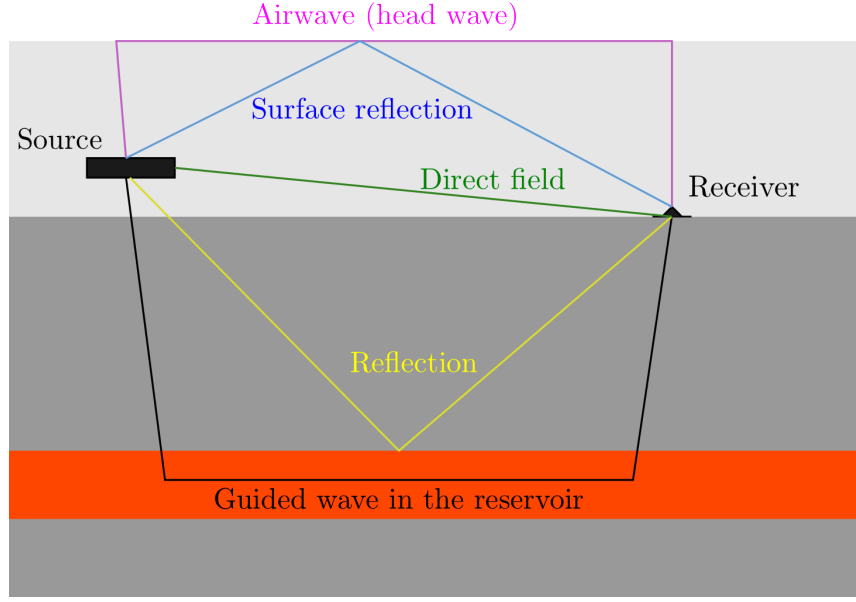


Figure 1.2: The EM signals broadcasted by the source is distributed in all directions, downward into the seafloor, directly to the receiver, and into the overlying water column. Figure from [14]

take the form

$$\begin{aligned}
 \nabla \times \mathbf{E} &= i\omega\mu\mathbf{H}, \\
 \nabla \times \mathbf{H} &= \tilde{\sigma}\mathbf{E} + \mathbf{j}^e, \\
 \nabla \cdot (\mu\mathbf{H}) &= 0, \\
 \nabla \cdot (\epsilon\mathbf{E}) &= q + q_e.
 \end{aligned} \tag{1.1}$$

Here \mathbf{E} and \mathbf{H} are the electric field and the magnetic field correspondingly, $\tilde{\sigma} = \sigma - i\omega\epsilon$ denotes the complex electric conductivity, ω denotes the frequency, μ denotes the magnetic permeability, ϵ denotes the dielectric permittivity, q is the spatial densities of free charges, and \mathbf{j}^e and q^e are the densities of extraneous electric currents and charges. It is shown in, e.g. [9] and [19] that in this case the *Helmholtz equations*

$$\begin{aligned}
 \nabla^2\mathbf{E} + \tilde{k}^2\mathbf{E} &= -i\omega\mu\mathbf{j}^e + (\nabla q^e)/\tilde{\epsilon}, \\
 \nabla^2\mathbf{H} + \tilde{k}^2\mathbf{H} &= -\nabla \times \mathbf{j}^e,
 \end{aligned} \tag{1.2}$$

must be satisfied. Here $\tilde{k} = i\omega\mu\tilde{\sigma} = i\omega\mu\sigma + \omega^2\mu\epsilon$ and the value of \tilde{k} is chosen to have a positive real part ($\text{Re}(\tilde{k}) > 0$). For situations with no free or extraneous charges, Maxwell's equations (1.1) reduce to

$$\begin{aligned}
 \nabla \times \mathbf{E} &= i\omega\mu\mathbf{H}, \\
 \nabla \times \mathbf{H} &= \tilde{\sigma}\mathbf{E} + \mathbf{j}^e, \\
 \nabla \cdot (\mu\mathbf{H}) &= 0, \\
 \nabla \cdot (\epsilon\mathbf{E}) &= 0.
 \end{aligned} \tag{1.3}$$

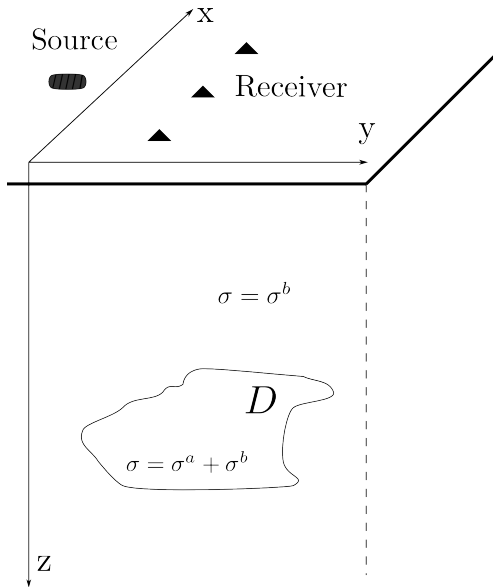


Figure 1.3: 3-D model of an inhomogeneity D with electric conductivity $\sigma = \sigma^a + \sigma^b$. The background electric conductivity is denoted $\sigma = \sigma^b$. The inhomogeneity structure is illuminated by electromagnetic signals broadcasted from the source.

These are the governing equations for our problem. Consequently the Helmholtz equations are reduced to

$$\begin{aligned}\nabla^2 \mathbf{E} + \tilde{k}^2 \mathbf{E} &= i\omega\mu \mathbf{j}^e, \\ \nabla^2 \mathbf{H} + \tilde{k}^2 \mathbf{H} &= -\nabla \times \mathbf{j}^e.\end{aligned}\quad (1.4)$$

Maxwell's equations (1.3) can be solved with a variety of methods. In our work we apply the Integral Equation (IE) method, which we reproduce from [19] in the next section.

1.4 The Integral Equation method

In the following procedure, the complex conductivity $\tilde{\sigma}$ in a local inhomogeneity D is regarded as a sum of a background complex conductivity $\tilde{\sigma}^b$ and the complex conductivity change $\tilde{\sigma}^a$, that is $\tilde{\sigma} = \tilde{\sigma}^b + \tilde{\sigma}^a$. Outside this inhomogeneity the conductivity is given as $\tilde{\sigma} = \tilde{\sigma}^b$. This is shown in Figure 1.3. Consequently we write the electric and magnetic fields as a sum

$$\mathbf{E} = \mathbf{E}^b + \mathbf{E}^a, \quad \mathbf{H} = \mathbf{H}^b + \mathbf{H}^a, \quad (1.5)$$

where the background field is due to the background complex conductivity, and the anomalous field is caused by the anomalous conductivity distribution $\tilde{\sigma}^a$.

We know that the total electromagnetic field must satisfy equations (1.3). Thus for the background fields \mathbf{E}^b and \mathbf{H}^b we have that

$$\begin{aligned}\nabla \times \mathbf{E}^b &= i\omega\mu_0 \mathbf{H}^b, \\ \nabla \times \mathbf{H}^b &= \tilde{\sigma}^b \mathbf{E}^b + \mathbf{j}^e,\end{aligned}\quad (1.6)$$

and for the anomalous fields \mathbf{E}^a and \mathbf{H}^a we have

$$\begin{aligned}\nabla \times \mathbf{E}^a &= i\omega\mu_0 \mathbf{H}^a, \\ \nabla \times \mathbf{H}^a &= \tilde{\sigma}^b \mathbf{E}^a + \mathbf{j}^a.\end{aligned}\quad (1.7)$$

Here \mathbf{j}^a is the density of excess anomalous electric currents within the inhomogeneity D ,

$$\mathbf{j}^a = \tilde{\sigma}^a(\mathbf{r})\mathbf{E}(\mathbf{r}) = \tilde{\sigma}^a(\mathbf{r})[\mathbf{E}^b(\mathbf{r}) + \mathbf{E}^a(\mathbf{r})]. \quad (1.8)$$

If (1.6) and (1.7) are added, the top two equations of Maxwell's equations (1.3) are obtained. As this thesis considers the CSEM problem where only low frequency EM signals are broadcasted from the antenna, the displacement currents can be neglected and the conductivity will satisfy $\tilde{\sigma} = \sigma$ in equations (1.6) and (1.7).

1.4.1 Solution to Maxwell's equations

The IE solutions to equations (1.6) and (1.7) can be written as

$$\mathbf{E}^b(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\mathbf{j}^e(\mathbf{r})dV \quad (1.9)$$

$$\mathbf{E}^a(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})(\mathbf{E}^b(\mathbf{r}) + \mathbf{E}^a(\mathbf{r}))dV \quad (1.10)$$

$$\mathbf{H}^b(\mathbf{r}') = \int_{\Omega} \mathcal{G}_H(\mathbf{r}'|\mathbf{r})\mathbf{j}^e(\mathbf{r})dV \quad (1.11)$$

$$\mathbf{H}^a(\mathbf{r}') = \int_{\Omega} \mathcal{G}_H(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})(\mathbf{E}^b(\mathbf{r}) + \mathbf{E}^a(\mathbf{r}))dV \quad (1.12)$$

where $\mathcal{G}_{E,H}$ denote Green's tensors (electric or magnetic) for the background conductivity, calculated as solutions to the differential system

$$\begin{aligned} \nabla \times \mathcal{G}_H &= \sigma^b \mathcal{G}_E + \delta I \\ \nabla \times \mathcal{G}_E &= i\omega\mu \mathcal{G}_H, \end{aligned} \quad (1.13)$$

where I denote the identity tensor. Derivations of these equations are outlined in, e.g., [19]. Note that for a monitoring setting, even for a computationally demanding calculation of Green's tensors, system (1.13) needs to be solved only once for the inverse problem. This applies to the background terms of both the electric field \mathbf{E}^b and the magnetic field \mathbf{H}^b , so the only calculations that must be repeated are those for finding the anomalous terms \mathbf{E}^a and \mathbf{H}^a . Once \mathbf{E}^a is found from equation (1.10), however, finding \mathbf{H}^a from (1.12) is straight-forward. Thus, the only significant computational cost when applying the IE method to our problem, is obtaining \mathbf{E}^a . Equation (1.10) is known as a Fredholm integral equation of the second-kind.

1.5 Numerical solution to the Fredholm integral equation

We will next solve the Fredholm integral equation using the same approach as by [11]. We discretize the reservoir into N_g cells D_i , where $i = 1, 2, \dots, N_g$. For each grid cell D_i we

assume the electric conductivity and electromagnetic fields to be constant. We define $\widehat{\mathbf{G}}_D$ as the $3N_g \times 3N_g$ matrix

$$\widehat{\mathbf{G}}_D = \begin{bmatrix} \Gamma_{xx}^{11} & \dots & \Gamma_{xx}^{1N_g} & \Gamma_{xy}^{11} & \dots & \Gamma_{xy}^{1N_g} & \Gamma_{xz}^{11} & \dots & \Gamma_{xz}^{1N_g} \\ & & & & \vdots & & & & \\ \Gamma_{xx}^{N_g 1} & \dots & \Gamma_{xx}^{N_g N_g} & \Gamma_{xy}^{N_g 1} & \dots & \Gamma_{xy}^{N_g N_g} & \Gamma_{xz}^{N_g 1} & \dots & \Gamma_{xz}^{N_g N_g} \\ \Gamma_{yx}^{11} & \dots & \Gamma_{yx}^{1N_g} & \Gamma_{yy}^{11} & \dots & \Gamma_{yy}^{1N_g} & \Gamma_{yz}^{11} & \dots & \Gamma_{yz}^{1N_g} \\ & & & & \vdots & & & & \\ \Gamma_{yx}^{N_g 1} & \dots & \Gamma_{yx}^{N_g N_g} & \Gamma_{yy}^{N_g 1} & \dots & \Gamma_{yy}^{N_g N_g} & \Gamma_{yz}^{N_g 1} & \dots & \Gamma_{yz}^{N_g N_g} \\ \Gamma_{zx}^{11} & \dots & \Gamma_{zx}^{1N_g} & \Gamma_{zy}^{11} & \dots & \Gamma_{zy}^{1N_g} & \Gamma_{zz}^{11} & \dots & \Gamma_{zz}^{1N_g} \\ & & & & \vdots & & & & \\ \Gamma_{zx}^{N_g 1} & \dots & \Gamma_{zx}^{N_g N_g} & \Gamma_{zy}^{N_g 1} & \dots & \Gamma_{zy}^{N_g N_g} & \Gamma_{zz}^{N_g 1} & \dots & \Gamma_{zz}^{N_g N_g} \end{bmatrix} \quad (1.14)$$

where $\Gamma_{\alpha\beta}^{ij}$ is the electric Green's tensor integrals

$$\Gamma_{\alpha\beta}^{ij} = \int_{D_j} \mathcal{G}_E^{\alpha\beta}(\mathbf{r}_i|\mathbf{r}_j)dV, \quad \alpha, \beta = x, y, z, \quad \text{and } i, j = 1, 2, \dots, N_g. \quad (1.15)$$

Let \mathbf{e}^b and \mathbf{e}^a be the $3N_g \times 1$ column vectors of the background and anomalous field,

$$\begin{aligned} \mathbf{e}^b &= [E_{x,1}^b \dots E_{x,N_g}^b, \quad E_{y,1}^b \dots E_{y,N_g}^b, \quad E_{z,1}^b \dots E_{z,N_g}^b]^T, \\ \mathbf{e}^a &= [E_{x,1}^a \dots E_{x,N_g}^a, \quad E_{y,1}^a \dots E_{y,N_g}^a, \quad E_{z,1}^a \dots E_{z,N_g}^a]^T, \end{aligned} \quad (1.16)$$

and let Σ^a denote the $3N_g \times 3N_g$ matrix with the anomalous conductivities along the diagonal;

$$\Sigma^a = \text{diag}([\Delta\sigma_1^a, \dots, \Delta\sigma_{N_g}^a, \Delta\sigma_1^a, \dots, \Delta\sigma_{N_g}^a, \Delta\sigma_1^a, \dots, \Delta\sigma_{N_g}^a]). \quad (1.17)$$

The discretized Fredholm integral equation will be

$$\mathbf{e}^a = \widehat{\mathbf{G}}_D \Sigma^a (\mathbf{e}^a + \mathbf{e}^b), \quad (1.18)$$

which we rearrange and obtain

$$\mathbf{A} \mathbf{e}^a = \widehat{\mathbf{G}}_D \Sigma^a \mathbf{e}^b, \quad (1.19)$$

and where $\mathbf{A} = \mathbf{I} - \widehat{\mathbf{G}}_D \Sigma^a$. In general, the matrix \mathbf{A} may be ill-conditioned. This is typically the case for large models with high electric conductivity contrast. This might cause slow convergence, or even divergence, for the iterative algorithms. In order to reduce the condition number we use preconditioning, where we replace the matrix \mathbf{A} with $\mathbf{M}_1 \mathbf{A} \mathbf{M}_2$, where the left and right preconditioners \mathbf{M}_1 and \mathbf{M}_2 are defined as

$$\mathbf{M}_1 = \sqrt{\Sigma^b}, \quad \mathbf{M}_2 = (2\Sigma^b + \Sigma^a)^{-1} 2\sqrt{\Sigma^b}, \quad (1.20)$$

where Σ^b is the $3N_g \times 3N_g$ matrix containing the background conductivities along its diagonal. We solve the resulting discretized Fredholm integral equation with the Biconjugate Gradient Stabilized (BICGSTAB) method [17, 3].

1.5.1 Numerical accuracy with respect to size of computational grid

From the outline above on solving the Fredholm integral equation, it is clear that our choice of computational grid may affect the final estimate. By choosing a discretized computational grid, we introduce numerical error. As we do not want the computational grid to influence the solution, this numerical error should not exceed the measurement errors.

We use the forward solver presented by [19, 11] to solve the forward problem. The solver has two important restrictions regarding the computational grid; its horizontal extent has to be quadratic and the number of cells in x - and y - direction must be equal. Thus $N_g = N_s \times N_s$ where N_s is the number of grid cells in both x - and y -direction. The value N_s is introduced to facilitate the following presentation of the optimal grid size.

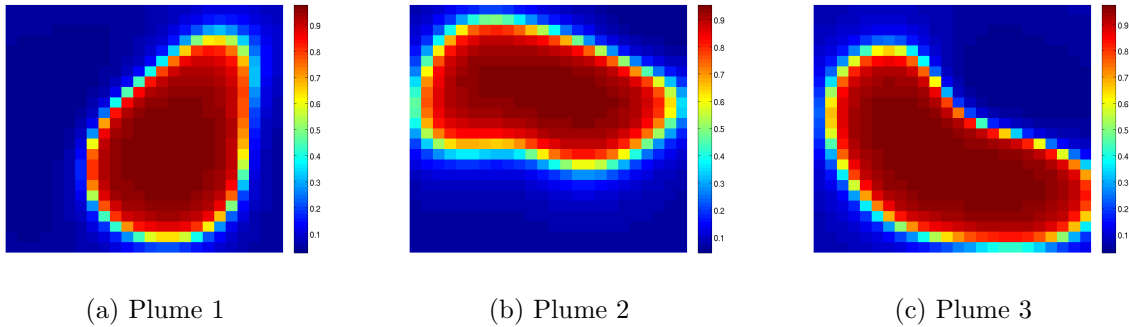


Figure 1.4: Reference time-lapse conductivity fields represented exactly for $N_s = 24, 48, 96, 192, 384$.

Consider the three examples of change in electric conductivity in the reservoir seen in Figure 1.4. These three electric conductivity fields are designed so that they can be represented on a computational grid with $N_s = 24$. It can also be represented exactly on grids with $N_s = 48, 96, 192, 384, \dots$. The forward solver results in a vector \mathbf{d} , which consists of the measurements of the electric and magnetic fields at the receivers. As the number of observations depends on the number of sources receivers, field components and frequencies used in the estimation, and not the size of the computational grid, these observation vectors can be compared. Consider a case with $N_s = 384$ as a reference case. Denote the observations for this case \mathbf{d}^{ref} . Then we compare the relative error in the observations with $N_s = 24, 48, 96, 192$, according to the formula

$$E(N_s) = \frac{\|\mathbf{d}^{ref} - \mathbf{d}^{N_s}\|}{\|\mathbf{d}^{ref}\|}, \quad N_s = 24, 48, 96, 192. \quad (1.21)$$

Here \mathbf{d}^{N_s} is the observations from a problem with grid size $N_s \times N_s$. As the number of observations depends on the number of sources receivers, field components and frequencies

used in the estimation, and not the size of the computational grid, these observation vectors can be compared. We expect that the finer a grid is, the smaller the relative error will be. In terms of accuracy, we expect that there is a value for N_s at which there is little to gain in increasing the grid size even further.

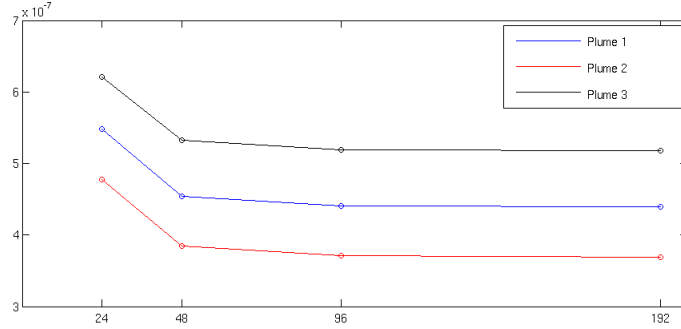


Figure 1.5: The relative error $E(N_s)$ versus for grid sizes $N_s = 24, 48, 96, 192$ with respect to a reference grid size of $N_s = 384$.

The results are presented in figure 1.5. The relative numerical error is of the order 10^{-7} , thus measurement error of the order 10^{-2} (specified in Chapter 5) will be the dominating source of error. Furthermore, there is a flattening of the relative error in the area between $N_s = 48$ and $N_s = 96$. From this figure it is clear that we should choose one of these two values for N_s . As there is still an improvement in accuracy associated with increasing grid size from $N_s = 48$ to $N_s = 96$, we choose this $N_s = 96$. In the case of three dimensions, we pick $N_g^{3D} = 48 \times 48 \times 5$.

Chapter 2

Inverse problems

We have already established some of the governing equations in the CSEM problem, and further information on the electric conductivity field in the reservoir enable us to approximately predict the electric and magnetic fields measured by the receivers. More accurate and comprehensive knowledge will result in better predictions.

Expressed mathematically, our knowledge about some parameters m should result in good predictions of our observations or data d . This knowledge is embedded in the forward operator G . It maps the parameters from the parameter space \mathbb{P} to the data space \mathbb{D} , as in the mathematical model

$$G(m) = d. \quad (2.1)$$

Any extra information incorporated in the forward operator should result in better predictions of the data. For an inverse problem the data d are known, we want to determine the parameter m . It is well-known that inverse problems are generally harder to solve than forward problems where you calculate d given m . In Section 2.1 inverse problems with certain complicating properties are classified as *ill-posed* inverse problems. One such complicating feature could be that real-life data include some amount of noise. This noise can be caused by for example measurement error. The data is a sum of the true data from the “perfect” experiment, d_{true} , and the noise component η ,

$$\begin{aligned} d &= G(m_{true}) + \eta \\ &= d_{true} + \eta. \end{aligned} \quad (2.2)$$

Here m_{true} are the parameters that corresponds to d_{true} , given that the forward modeling is exact. The noise component η is unknown, otherwise we would simply exclude it from our calculations. We realize that the true parameters m_{true} , from which our observed data arises, are not the parameters that “solves” our problem (2.1), considering that our data are contaminated by noise.

When solving an inverse problem, we want to obtain an inverse operator F that maps the data d back to the parameter space, and preferable not “very far” from the true parameters, see Figure 2.1. For some simple problems, F can be found as an exact inverse operator, $F(G(m)) = m$. For more realistic problems, the inverse operator F can be

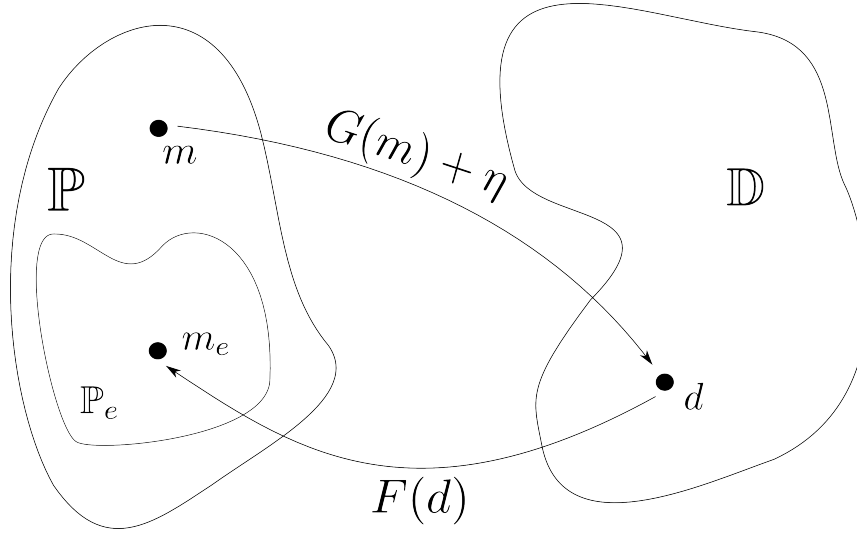


Figure 2.1: The parameters m are mapped by the forward operator from parameter space \mathbb{P} to data space \mathbb{D} . Our aim is to obtain an inverse operator F that maps the observed data d back to the parameter space \mathbb{P}_e . In some sense regularization can be used to modify this space \mathbb{P}_e so that nonphysical or unlikely solutions are excluded. Figure from [4]

found as a process involving many steps, resulting in an approximation $F(d) = m_e$. The estimated parameters m_e are included in the parameter space \mathbb{P}_e , that is the attainable parameter space for data d from \mathbb{D} . In our work F is regarded as a process more than an operator, where the process involves utilizing different techniques for obtaining m_e .

Before we present this estimation process, we modify our problem slightly. For the CSEM problem both the parameters and observations are finite sets. The observations consist of measurements of the electric and magnetic field components at the receivers at a given time. By discretizing the reservoir, the parameters becomes the electric conductivities in the N_g grid cells in the computational grid. The observations can thus be represented as a vector \mathbf{d} . The discrete representation of the model parameters is denoted by $m(\mathbf{r}; \mathbf{a})$. The general mathematical model (2.1) can therefore be rewritten as a system of equations,

$$\mathbf{G}(m(\mathbf{r}; \mathbf{a})) = \mathbf{d}. \quad (2.3)$$

Here $\mathbf{G}(m(\mathbf{r}; \mathbf{a}))$ is a vector of length N_d , and \mathbf{a} are the coefficients in the discrete parameterization. Problem (2.3) is called a parameter estimation problem.

2.1 Solving an inverse problem

Upon solving an inverse problem, three issues should be considered. Those are the

- existence,

- uniqueness, and
- stability,

of the attained solution. Because of the data contamination, we will not pursue an exact solution, as it can not be identified. For problems where an exact solution exists and we are looking for it, there might be numerous solutions besides $m(\mathbf{r}; \mathbf{a})_{true}$ that exactly satisfy $\mathbf{G}(m(\mathbf{r}; \mathbf{a})) = \mathbf{d}_{true}$. That is a matter of uniqueness. The third issue is the stability of a solution. Computing a solution estimate to an inverse problem can be an extremely unstable process, with respect to small changes in observed data. A small uncontrollable change in noise contribution might lead to a great change in the estimated parameters. Inverse problems with these described features are referred to as ill-posed problems [2].

As no exact solution estimate $m(\mathbf{r}; \mathbf{a})_{true}$ can be found, our aim is to identify parameters that are “sufficiently close” to the true parameters. To obtain such estimates, we use the standard strategy involving the minimization problem

$$\mathbf{a} = \arg \min_{\{\mathbf{a} | m(\mathbf{r}; \mathbf{a}) \in \mathbb{P}_e\}} J(m(\mathbf{r}; \mathbf{a})), \quad (2.4)$$

where the objective function $J(m(\mathbf{r}; \mathbf{a}))$ is defined as

$$J(m(\mathbf{r}; \mathbf{a})) = (\mathbf{d} - \mathbf{G}(m(\mathbf{r}; \mathbf{a})))^H \mathbf{C}^{-1} (\mathbf{d} - \mathbf{G}(m(\mathbf{r}; \mathbf{a}))) + \alpha \phi(m(\mathbf{r}; \mathbf{a})). \quad (2.5)$$

Here \mathbf{C} is the covariance matrix for measurement errors, $\phi(\mathbf{m})$ is a general regularization term with weight α , and H is the Hermitian operator as the calculations include complex numbers. We replace the parameter estimation problem (2.3) with the minimization problem (2.4). The idea is that a set of parameters that solve (2.4) will be a good approximation to the true parameters $m(\mathbf{r}; \mathbf{a})_{true}$. This process involving replacement of the original problem with a minimization problem, is an example of regularization.

2.2 Regularization

Regularization is a collective term used to describe methods where we include *a priori* information to reduce the set of attainable solutions \mathbb{P}_e . It is a process where we change the problem to a similar and better posed problem. Some solutions will be preferred to others based on this a priori information. But assuming that the changes are reasonable, we might produce an usable solution to an otherwise unsolvable problem.

Modifying our problem from the mathematical model (2.1) to the parameter estimation problem (2.3), regularization is applied as we restrict the solution space. Our parameters consist of the value of the electric conductivity at a number of spatial grid points in our reservoir. The reservoir itself is by the continuum hypothesis¹ considered a continuous medium, however, so we are not solving the original problem, but an approximation.

¹Any matter is discontinuous or discrete at microscopic scales. The continuum hypothesis states that as long as the system under consideration is much larger than the mean free part of the molecules in the system, it is possible to ignore the discrete molecular structure of the matter and replace it by a continuous distribution, called a continuum [12]

We note that $\phi(\mathbf{a}) = \phi(m(\mathbf{r}; \mathbf{a}))$ is a different example of regularization presented in the objective function (2.5). This regularization term can be chosen to favour different types of solutions. One standard choice of $\phi(\mathbf{a})$ is known as Tikhonov regularization. It is a widely used method applied to stabilize parameter estimation problems. If reliable a priori knowledge exist it can be incorporated into the regularization term

$$\phi(m(\mathbf{a})) = (m(\mathbf{a}) - m(\mathbf{a})_{prior})^H \mathbf{C}_{prior}^{-1} (m(\mathbf{a}) - m(\mathbf{a})_{prior}), \quad (2.6)$$

and penalize solutions that deviate from this prior model \mathbf{m}_{prior} . Here \mathbf{C}_{prior} denote the positive definite weighting matrix which in a Bayesian framework corresponds to the prior covariance matrix of error in $m(\mathbf{a})_{prior}$ [5]. The better your prior model fits the true solution, the more effective this regularization term is. If the prior model $m(\mathbf{a})_{prior}$ is unreliable for some reason, however, and the weighting variable α ensures that the regularization term $\phi(m(\mathbf{a}))$ is included when minimizing the objective function (2.5), there is a risk of a solution that is significantly different from the true solution.

In this thesis, the estimated parameters solves the minimization problem (2.4), where the objective function is defined as in equation (2.5) with the regularization term $\phi(\mathbf{a}) = 0$. For a realistic situation, the number of parameters N_g is significantly higher than the number of data N_d . A high number of unknowns and restricted resolution power of the data, enhance the impact of the prior model on the estimates, and gives a considerable risk of obtaining an unstable estimation. To overcome this, and the fact that a high number of unknowns is computational demanding, we suggest to regularize the problem by reparameterization. We choose a low-dimensional representation of the model parameters $m(\mathbf{a})$, and seek to obtain a correspondence to the available data \mathbf{d} . Such a reduced representation will bias the solution, however, hence we seek a representation with a high flexibility with respect to expected structures in $m(\mathbf{a})$.

Chapter 3

Parameter representation and adaptive refinement

For the CSEM problem we have some knowledge about the initial electric conductivity distribution, as the reservoir is oil-filled. During secondary production, water is injected. It will spread out in the reservoir creating a new, water-flooded, connected area, whose shape we wish to monitor using EM signals. We assume that the electric conductivity field mainly consists of two regions of nearly constant conductivity. Between these regions, the conductivity changes rapidly. If we know the location of this area of rapid change, we know the main characteristics of the whole domain. The data we measure, however, is not sufficient to discover fine-scale variations of the separation front, but the reduced representation will nevertheless represent the front as a rapid change in estimated conductivity. We choose a reduced representation that meet these requirements, first presented by [6].

3.1 Composite parameter representation

The parameter function $m(\mathbf{r}; \mathbf{a})$, representing the time-lapse electric conductivity change in the reservoir, is given as

$$m(\mathbf{r}; \mathbf{a}) = c_1 E_1(I(\mathbf{r}; \mathbf{a})) + c_2 E_2(I(\mathbf{r}; \mathbf{a})), \quad (3.1)$$

where the interior function is denoted by I , E_j are the exterior functions and c_j are called the exterior coefficients for $j = 1, 2$. We assume the exterior coefficients to be known and equal to the conductivity change in the water-flooded and oil-flooded areas of the reservoir respectively.

3.1.1 Exterior functions

The exterior functions are given as smoothened Heaviside-like functions as follows;

$$E_1(I) = \tilde{H}(I), \quad E_2(I) = 1 - \tilde{H}(I), \quad \tilde{H}(I) = \frac{1}{\pi} \tan^{-1}(I) + \frac{1}{2}, \quad (3.2)$$

where the functional value of $\tan^{-1}(I)$ is in the range $(-\pi/2, \pi/2)$. This representation makes it possible to handle large and almost constant-valued regions, but also transition regions in between, with a varying degree of smoothness. For a gentle $\nabla I(\mathbf{r}; \mathbf{a})$ across the zero isocontour of $I(\mathbf{r}; \mathbf{a})$, there will be a wide transition region between domains of constant electric conductivity. From figure 3.1 it is clear that a more rapidly varying $I(\mathbf{r}; \mathbf{a})$ will result in a narrower transition region.

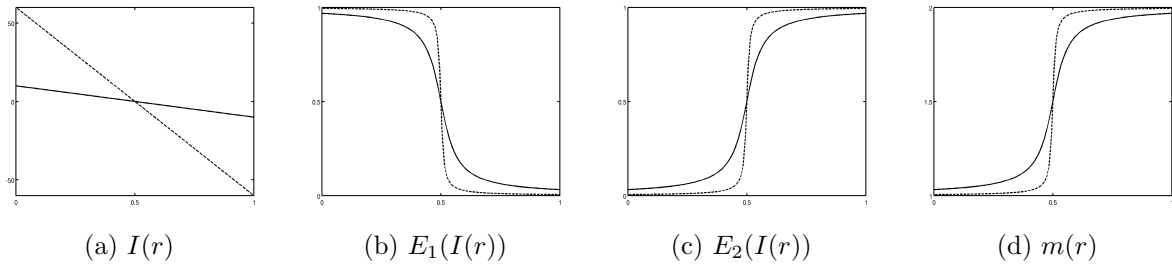


Figure 3.1: Figures showing the effects of steepness of $I(r)$ across the zero isosurface, for $E_1(I(r)), E_2(I(r))$ and $m(r)$, for two different interior functions $I(r)$ where $r \in [0, 1]$, and where coefficients $c_1 = 1$ and $c_2 = 2$.

3.1.2 Interior function

From the composite parameter representation (3.1), where we assume that the exterior coefficients c_j are known, the degrees of freedom in the estimation process are now related to $I(\mathbf{r}; \mathbf{a})$. We represent the interior function by

$$I(\mathbf{r}; \mathbf{a}) = \sum_{k=1}^{N_a} a_k \theta_k(\mathbf{r}), \quad (3.3)$$

where the functions $\{\theta_k\}_{k=1}^{N_a}$ denote a set of basis functions, $\{a_k\}_{k=1}^{N_a}$ the associated interior coefficients, and N_a is the number of basis functions. Note that in order to find the *model parameter* m , we must choose a set of basis functions and find the interior coefficients. Before we proceed to consider these interior coefficients, we present the basis functions θ_k .

3.2 Basis functions

From the presentation of our parameter representation (3.1) it is clear that the shape of the interior function $I(\mathbf{r}; \mathbf{a})$, and thus the basis functions θ_k , play an important role in the estimation process. The basis functions should be able to represent the structures we expect to be present in the reservoir. In our work we consider two types of basis functions.

3.2.1 Trilinear basis functions

Consider the computational grid with N_g cells. For trilinear basis functions (TBFs), the cells in the computational grid are partitioned into a coarser parameterization grid. Each cell in the parameterization grid consists of one or more cells from the computational grid. For each vertex in the parameterization grid, there is a TBF with its center in it. A TBF has value one in its center, and value zero in any other vertex. The shape of a TBF is, as the name suggests, linear. The set of TBFs provides a continuous representation of $I(\mathbf{r}; \mathbf{a})$.

For a parameterization grid with only one cell, in which the entire computational grid is embedded, eight basis functions exist, see Figure 3.2. Each basis function has center in one of the eight vertices. If the parameterization grid has more than eight cells the trilinear basis functions has local support, as each vertex has at most eight adjacent cells.

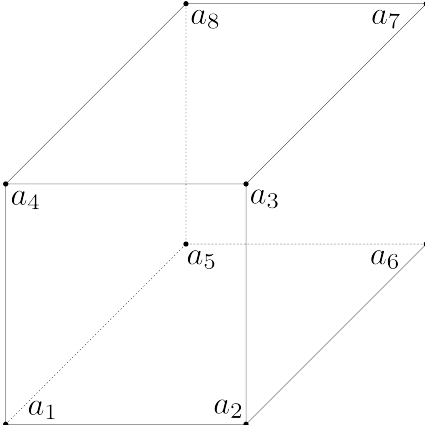
$$\theta_k(x, y, z) = \begin{cases} (1-x)(1-y)(1-z) & \text{if } k = 1 \\ x(1-y)(1-z) & \text{if } k = 2 \\ xy(1-z) & \text{if } k = 3 \\ (1-x)y(1-z) & \text{if } k = 4 \\ (1-x)(1-y)z & \text{if } k = 5 \\ x(1-y)z & \text{if } k = 6 \\ xyz & \text{if } k = 7 \\ (1-x)yz & \text{if } k = 8 \end{cases}$$


Figure 3.2: A reference parameterization cell, and the eight trilinear basis functions $\{\theta_k\}_{k=1}^8$ with support on this cell, $D_r = [0, 1] \times [0, 1] \times [0, 1]$, and corresponding interior coefficients, $\{a_k\}_{k=1}^8$.

3.2.2 Radial basis functions

A radial basis function (RBF) is characterised by that its value in any point \mathbf{r} is given as a function of the distance from that point to the center \mathbf{r}_k , see figure 3.3. We use a Gaussian type¹ RBF

$$\theta_k(\mathbf{r}) = \exp\left(-\frac{1}{2\sigma_k^2}(\mathbf{r} - \mathbf{r}_k)^2\right), \quad k = 1, 2, \dots, N_a \quad (3.4)$$

where σ_k is a shape parameter², corresponding to the standard deviation in this case. An RBF has global support, so is applies to the computational grid where the value of an RBF in a cell depends on the distance from the cell to the RBF's center. For the TBF it is the partitioning of the computational grid that restricts what functions you can approximate.

¹Gaussian RBFs are on the form $\rho(r) = \exp(-[\epsilon r]^2)$.

²Note that the shape parameter σ_k not is related to the model parameters σ (electric conductivity).

For RBFs the shape parameters σ_k and center location \mathbf{r}_k determine what forms can be approximated on the computational grid.

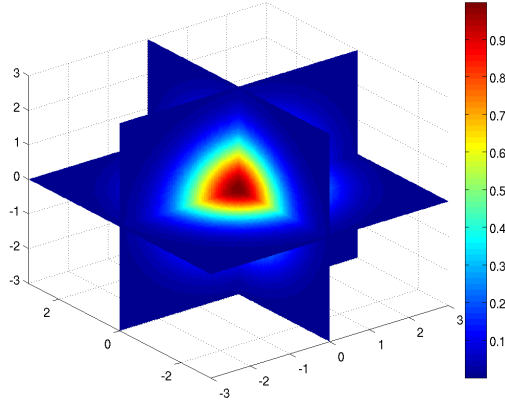


Figure 3.3: A radial basis function with center $\mathbf{r} = (0, 0, 0)$ and shape parameter $\sigma = 1$.

3.3 Multi-level estimation strategy

Now consider the interior coefficients a_k . For a reduced parameterization, N_a is much smaller than N_g . A coarse resolution will increase parameter sensitivity, and consequently can stabilize the solution. It is also shown by [6] that for some inverse problems a coarse representation reduce model nonlinearity. However, a coarse resolution might not have the flexibility required to obtain the various structures of the solution. But as we know the CSEM problem to have limited resolution power of the observed data, too many parameters will easily lead to over-parameterization or trapping us in a local minimum due to nonuniqueness of the solution. For these reasons, finding the optimal set of basis functions $\{\theta_k\}_{k=1}^{N_a^*}$ with respect to center location and shape parameter, is imperative. This set can not be known prior to the estimation process.

To meet the challenges described, we select a sequential multi-level estimation strategy [7]. In this procedure, the resolution of the representation is found during the estimation process. For a coarse initial representation at level $s = 1$, the coefficients a_k are optimized in order to minimize the objective function (2.5). If the solution criterion is met, the final estimate is found. Otherwise, we continue to the next level, where we add one or more basis functions to the existing set. The process of adding basis functions to the existing representation in order to increase resolution, is known as a refinement process. With a new representation, and an extended structure matrix, the coefficients are again optimized to minimize the objective function. This whole process continues until the solution criterion is met, or the maximum number of levels is exceeded, see Figure 3.4.

The refinement process is presented next, while the optimization process follows in Chapter 4.

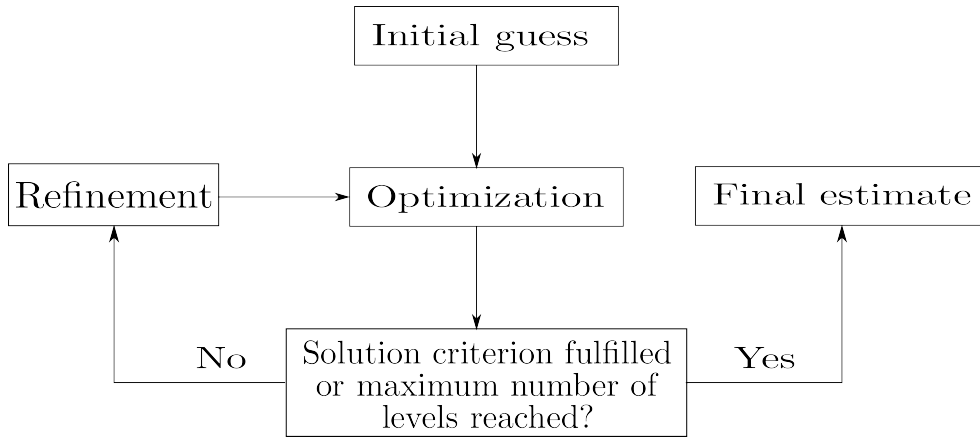


Figure 3.4: The multi-level estimation strategy

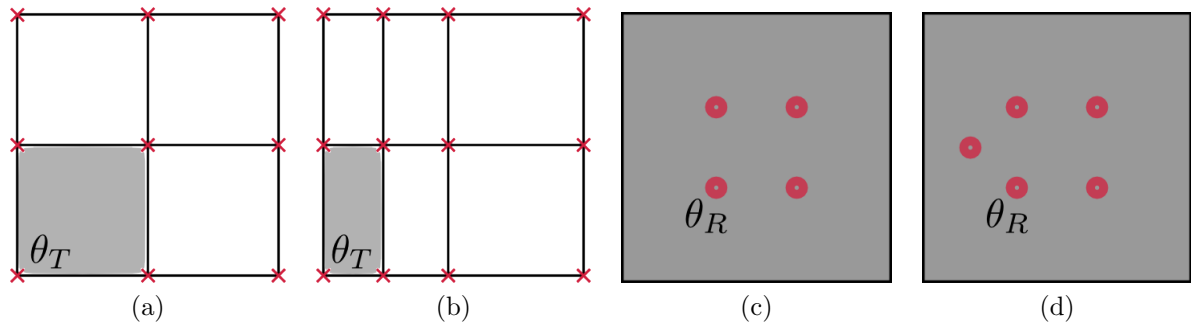


Figure 3.5: An illustration of the effects on a basis function due to refinement. For a TBF θ_T in figure (a) and (b) refinement might lead to a decrease in domain of support (shaded area). For an RBF θ_R as in figure (c) and (d) there is no change, as RBFs are global. Figure (a) and (b) are also examples of the higher increase in coefficients associated with TBFs than RBFs.

3.4 Adaptive refinement and selection measure

Finding an efficient rule for refinement is not trivial. We need a set of guidelines to aid us in the execution of this successive increase of basis functions. One straight-forward rule is to apply *regular refinement*, which means that basis functions are added so that the resolution increases but remains constant throughout the computational grid. *Adaptive refinement* is an alternative refinement strategy, which enables us to increase the resolution in sensitive areas, while keeping a low resolution in the non-sensitive areas.

Refinement was presented as including basis functions to an existing parameterization. The refinement process is different for the two types of basis functions. For TBFs, a new basis function can only arise as a result of subdivision of the parameterization grid, for which at least one more cell is created. In two dimensions this corresponds to including one or more lines, see Figure 3.5. By adding new lines, we create a basis function not only in each new grid node, but we also affect the support of the “old” basis functions.

As the “steepness” of TBFs are determined by the support of the function, the structure matrix Θ has to be calculated for each refinement candidate in the refinement process. For RBFs, a new basis function can be included in any location with any shape parameter value, without affecting the existing basis functions. Thus the only change for the structure matrix is the extra basis vector(s) corresponding to the new basis function(s).

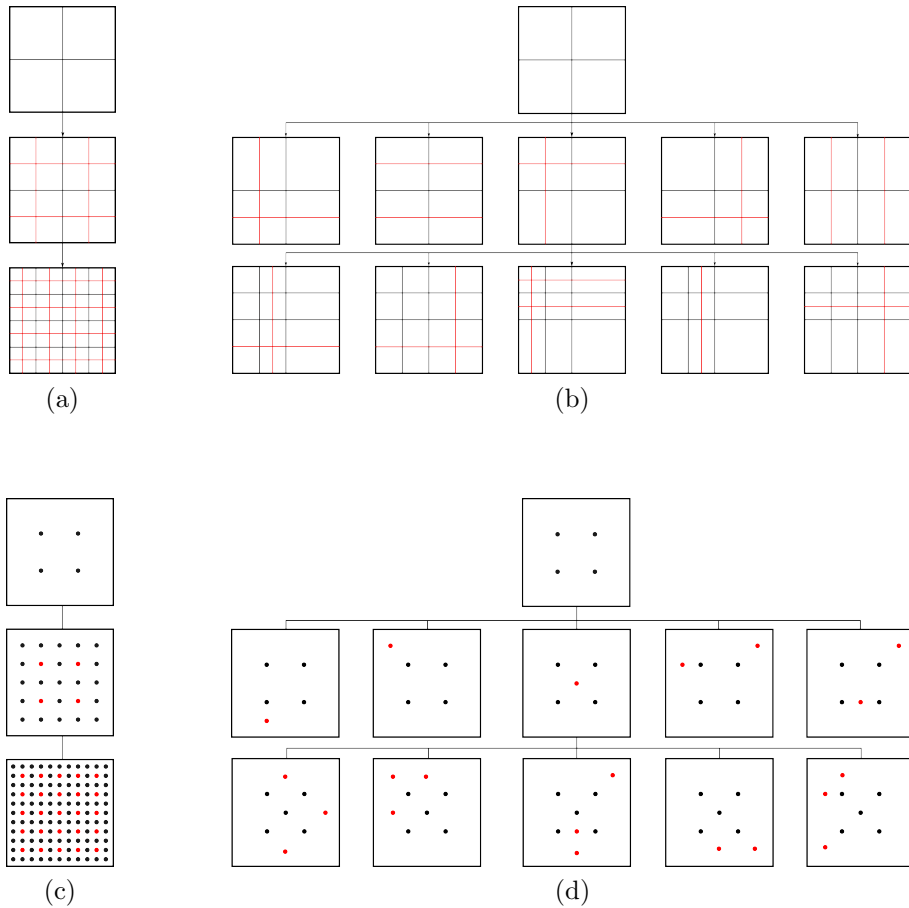


Figure 3.6: Two-dimensional examples of regular and adaptive refinement. Figure (a) and (b) shows examples of regular and adaptive refinement with TBFs, whilst (c) and (d) shows the same for RBFs.

Having addressed the differences between TBFs and RBFs in the refinement process, we consider the differences regarding regular and adaptive refinement. Regular refinement ensure a parameter estimation $m(\mathbf{r}; \mathbf{a})$ for which the resolution is constant over the entire domain, see Figures 3.6a and 3.6c. This leads to an exponential growth in the number of interior coefficients. This strategy might lead to over-parameterization, and we might obtain an unstable solution.

We assumed that the electric conductivity is more or less constant for the different regions in our domain. The adaptive refinement strategy allows us to increase the resolution

in areas where we expect a change in the electric conductivity, without having to increase the resolution in areas where only a few basis functions are needed. The difference between the two refinement strategies is illustrated in Figure 3.6. It is clear that the use of RBFs can be more conservative compared to TBFs, as TBFs might include basis functions in non-sensitive areas due to “crossing lines”. We proceed to present a multi-level estimation strategy with adaptive refinement for RBFs. The use of TBFs in the same process is presented by [7].

Assume that the optimization at level s is finished without fulfilling the solution criterion, and we consider the refinement process at level $s+1$. A large number of basis functions or refinement candidates, can be added. By combining two or more such basis functions, new potential candidates are created, and can be added to the existing representation.

To avoid over-parameterization, we impose a maximum number of added basis functions at each level, ΔN^{max} , such that the number of added basis functions ΔN is lower than the maximum, $\Delta N \leq \Delta N^{max}$. We also assume that we have some parameterization candidates $\{P^l\}_{l=1}^L$. Here $P^l = \{\{a\}_{j=1}^{N_a^l}, \{\theta\}_{j=1}^{N_\theta^l}\}$, where L is the total number of candidates, and N_a^l denotes the number of interior coefficients for candidate P^l . Given a refinement candidate P^1 , how can we be sure that this is a refinement candidate we prefer over refinement candidate P^2 ? For the adaptive refinement strategy to work, we need a selection measure to choose between the different candidates. In this thesis, we consider two different selection measures, *predicted attainable objective function value* and the *steepness of the objective function gradient*. The predicted attainable objective function values is a 2. order strategy. This is because it includes an approximation of the 2. order derivative of the objective function. The steepness of the objective function gradient is a 1. order selection measure, as it compares the gradient (first order derivative) of the objective function for the compared candidates. We expect the 2. order selection measure to make better choices in the refinement process, and thus result in better estimates compared to the 1. order selection measure. The question is whether or not the 1. order measure can produce usable estimates, compared to the estimates from the 2. order measure. The computational time associated with the 1. order measure is negligible compared to that of the 2. order measure. Thus a result indicating that the 1. order measure is sufficient would be of importance. The fact that we do not expect any detailed information about the separation front reinforce the hope that a 1. order selection measure can be used.

3.4.1 Predicted attainable objective function values

We group the refinement candidates $\{P^l\}_{l=1}^L$ into subsets $\{\{P_W^l\}_{l=1}^{L_W}\}_{W=1}^{\Delta N^{max}}$, according to the number of basis functions added. For the candidate P_W^l , we associate a predicted attainable objective function value $\tilde{J}(P_W^l)$. This value is an approximation to $\min_{\mathbf{a}_W^l} J(P_W^l)$, that follows from linearization of the model output;

$$\tilde{J}(P_W^l) \triangleq \min_{\mathbf{a}_W^l} (\mathbf{d} - L(\mathbf{G}(\mathbf{a}_W^l)))^H \mathbf{C}^{-1}(\mathbf{d} - L(\mathbf{G}(\mathbf{a}_W^l))), \quad (3.5)$$

where $\mathbf{G}(\mathbf{a}) = \mathbf{G}(m(\mathbf{r}; \mathbf{a}))$ and the linearization operator is $L(\mathbf{G}(\mathbf{a}_W^l)) \triangleq \mathbf{G}(\mathbf{a}_W^l) + \mathbf{S}_W(\mathbf{a}_W^l - \mathbf{a}_W^{l,I})$. Here $\mathbf{a}_W^{l,I}$ denotes the initial estimate of \mathbf{a}_W^l at the start of level $s + 1$ and \mathbf{S}_W denotes the sensitivity matrix of $\mathbf{G}(\mathbf{a}_W^l)$ evaluated at $\mathbf{a}_W^{l,I}$. Observe that the right hand side of approximation (3.5) is quadratic with respect to \mathbf{a}_W^l . The minimizer $\mathbf{a}_W^{l,*}$ is found to be

$$\mathbf{a}_W^{l,*} = \mathbf{a}_W^{l,I} + (\mathbf{S}_W^H \mathbf{C}^{-1} \mathbf{S}_W)^{-1} \mathbf{S}_W^H \mathbf{C}^{-1} (\mathbf{d} - \mathbf{G}(\mathbf{a}_W^{l,I})). \quad (3.6)$$

When we insert the minimizer (3.6) into equation (3.5), we obtain the predicted attainable objective function value

$$\tilde{J}(P_W^l) = (\mathbf{d} - \mathbf{G}(\mathbf{a}_W^{l,I}))^H (\mathbf{C}^{-1} - [\mathbf{C}^{-1} \mathbf{S}_W (\mathbf{S}_W^H \mathbf{C}^{-1} \mathbf{S}_W)^{-1} \mathbf{S}_W^H \mathbf{C}^{-1}]) (\mathbf{d} - \mathbf{G}(\mathbf{a}_W^{l,I})). \quad (3.7)$$

Let $P_W^{l,*}$ be the parameterization within subset $\{P_W^l\}_W$ with the lowest associated value of \tilde{J} . Then $\tilde{J}(P_{W+1}^{l,*}) \leq \tilde{J}(P_W^{l,*})$ follows from the fact that P_{W+1} is in a larger solution space than P_W . Thus, in a sense, it is always advantageous to include more basis functions. On the other hand, we need the improvement on the predicted attainable objective function value to be significant. We do not choose P_{W+1} over P_W if

$$\tilde{J}(P_{W+1}^{l,*}) + \zeta(P_{W+1}^{l,*}) \geq \tilde{J}(P_W^{l,*}) \quad (3.8)$$

where

$$\zeta(P_W^l) = \sqrt{4\tilde{J}(P_W^l) - 2(N_d - N_a^l)}. \quad (3.9)$$

The selected parameterization $P_W^{l,*}$ will then be the lowest dimensional parameterization that also satisfies the criterion (3.8).

3.4.2 Objective function gradient steepness

Predicting the attainable objective function values involves calculating the sensitivity matrix. This is a computational demanding operation. Thus the number of potential candidates we consider for refinement is restricted. We now present a method where the actual calculation of the sensitivity matrix is not needed, and thus allows us to evaluate more candidates. Our aim is still to discover the potential each basis function has in reducing the objective function, and then to choose the one(s) with the best potential.

Let

$$\nabla_a J = \left(\frac{\partial J}{\partial a_1}, \dots, \frac{\partial J}{\partial a_{N_a}} \right)^T \quad (3.10)$$

denote the coarse-scale gradient of the objective function, that is the gradient with respect to the interior coefficients. Further, let

$$\nabla_I J = \left(\frac{\partial J}{\partial I_1}, \dots, \frac{\partial J}{\partial I_{N_g}} \right)^T \quad (3.11)$$

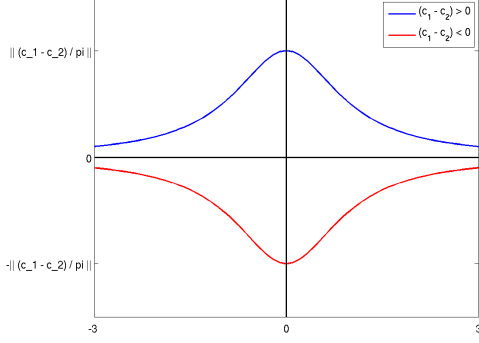


Figure 3.7: The partial derivative $\partial m/\partial I$ given positive or negative constant $(c_1 - c_2)$. It displays that the amount of change in $m(\mathbf{r})$ per change in $I(\mathbf{r})$ depends strongly on the value of the interior function.

denote the fine-scale gradient of the objective function with respect to the interior function in each grid cell. For our composite representation (3.1) the relation between the coarse- and fine-scale gradient is given as

$$\nabla_a J = \Theta \nabla_I J, \quad (3.12)$$

$$= [\theta_1, \theta_2, \dots, \theta_{N_a}]^T \nabla_I J, \quad (3.13)$$

where Θ is the $N_a \times N_g$ structure matrix. Relation (3.12) is found with the use of the chain rule for functions of several variables [4]. An element of the coarse-scale gradient can be written as

$$\frac{\partial J}{\partial a_k} = \frac{\partial J}{\partial m} \frac{\partial m}{\partial I} \frac{\partial I}{\partial a_k}. \quad (3.14)$$

Here

$$\frac{\partial m}{\partial I} = (c_1 - c_2) \frac{1}{\pi(1 + I^2)} \quad (3.15)$$

and

$$\frac{\partial I}{\partial a_k} = \theta_k. \quad (3.16)$$

Consider the partial derivative (3.15). It contains information of the amount of change in model parameters $m(\mathbf{r}; \mathbf{a})$ per change in the interior function $I(\mathbf{r}; \mathbf{a})$. From the composite representation (3.1) we know that, unless $I(\mathbf{r}; \mathbf{a})$ is close to zero, there is not much change in $m(\mathbf{r}; \mathbf{a})$, even for a considerable change in I . This is illustrated in figure 3.7.

Consider again the relation between the fine-scale and coarse-scale gradients (3.12). The fine-scale gradients $\nabla_I J$ are generally computational demanding to calculate, but they are independent of the reduced representation. This observation allows us to calculate it only once per level. Given the fine-scale gradient, we only need the partial derivative (3.16) to obtain the coarse-scale gradient $\nabla_a J$, and this is found to be exactly the structure vector θ_k for the given coefficient, which is known to us.

We know the coarse-scale gradient, and we can use it as a refinement selection measure. Consider the norm of the coarse-scale gradient, $\|\nabla_a J(a)\|$. The idea is that the basis

function which results in the largest norm, will be located in the area where the objective function has largest potential of decreasing. Let Θ_{old} be the structure matrix from the previous level. The structure matrix on the new level will then be on the form $\Theta_{new} = [\Theta_{old}, \Theta_{test}]$. Correspondingly the norm of the coarse-scale gradient can be split in two parts;

$$\|\nabla_a J(\mathbf{a}_{new})\| = \|[\nabla_a J(\mathbf{a}_{old})^T, \nabla_a J(\mathbf{a}_{test})^T]\|. \quad (3.17)$$

As $\|\nabla_a J(\mathbf{a}_{old})\|$ is equal for all parameterizations, we need to find the basis functions that gives the largest $\|\nabla_a J(\mathbf{a}_{test})\|$. If we consider one basis function at the time, the basis function with the largest $|\partial J/\partial a_{test}|$ is chosen. In the case where you want to add more basis functions at each level, it is common either to add the two or three basis functions with the highest norm, or add all basis functions for which $|\partial J/\partial a|$ is within 95% of the highest value. The drawback with this strategy is that once you choose one basis function, the resolution is no longer as on the previous level. Thus the sensitivities will change, and the second best basis function is perhaps no longer the best of the remaining potential basis functions. Consider the case where all potential basis functions are duplicated, and the duplications also considered as basis functions. Then the two best basis functions with respect to the value of the norm are identical. But once one of them is included in the representation the duplication will at best have no effect, and might even lead to an unstable estimate. In our work we only consider the case where one basis function is added at the time.

We mentioned earlier that the fine-scale gradient $\nabla_I J$ has to be calculated once per level. As $\partial m/\partial I$ is known, the remaining partial derivative is the change in objective function value per change in electric conductivity $\nabla_m J$. This expression is obtained from the objective function 2.5 directly:

$$\nabla_m J = \mathbf{F}_m^* \mathbf{R}, \quad (3.18)$$

where $\mathbf{R} = \mathbf{C}^{-1}(\mathbf{G}(m(\mathbf{r}; \mathbf{a})) - \mathbf{d})$ is the residual vector and \mathbf{C}^{-1} is the weighting matrix of the data. \mathbf{F}_m is known as the Frechet derivative matrix, or the sensitivity matrix. The preceding section presents a method where the sensitivity matrix is found explicitly. In the sensitivity matrix is known, a simple matrix-gradient multiplication with the residual vector would provide us with the exact gradient. However, as this is just as computational demanding as the 2. order method, we propose a different method to obtain an approximated “quick” gradient, called *quasi-analytical* (QA) approach. The QA approach is presented in Section 4.3.2.

3.4.3 The set of potential basis functions for refinement

The success of the adaptive refinement strategy is strongly dependent on the performance of the set of potential refinement candidates. The 2. order selection measure can combine basis functions to make new candidates, while the 1. order selection measure considers only single basis functions. However, for both measures the set of potential basis functions is of fundamental importance. The characteristics of a radial basis function is the shape

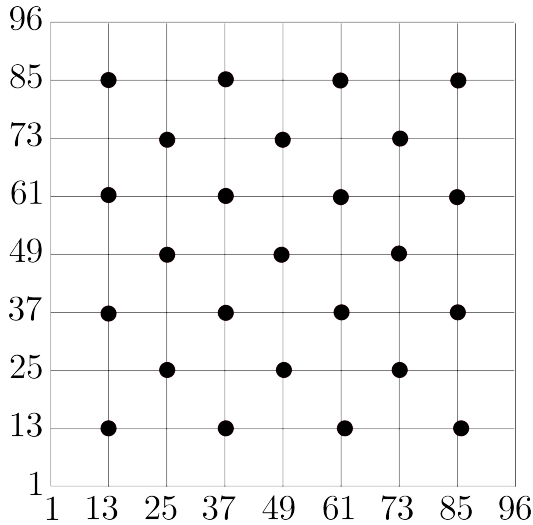


Figure 3.8: The location and shape of potential basis functions: We have three different values for the shape parameter σ of the basis functions. No basis function is placed at the boundary. The black dots mark the location of a basis function with the largest shape parameter. Basis functions with the two smallest shape parameter values are placed in each vertex, including the ones where there already is an overlap with a basis function with large shape parameter.

parameter σ_k and the center location \mathbf{r}_k . Thus we need a strategy for how these characteristics are varied and combined to make a set with a conservative number of RBFs which represent a broad range of shapes in the solution. For the two-dimensional computational grid with 96×96 grid cells, we choose 7 potential centers for the RBFs in the x - and y -direction. Thus we have a 7×7 grid of potential centers \mathbf{r}_k . For the three dimensional computational grid of size $48 \times 48 \times 5$ we choose a $7 \times 7 \times 5$ grid of potential centers \mathbf{r}_k . In all examples we exclude centers on the boundary in the x - and y -direction.

Regarding the shape parameter σ_k we should keep in mind that we do not know the shape of the separation front, and thus should be able to represent both large and tiny areas of changing electric conductivity. From the literature regarding the use of RBFs in function interpolation, we have found some general formulas for how the shape parameters can be determined. We have not, however, found anything regarding the use of RBFs in parameter representation in inverse problems. Inspired by some of the formulas from the interpolation-theory, and some trial and error with different shape parameters, we end up with three different shape parameters. One of the shape parameter values is chosen to be the total number of computational grid cells in on direction, divided by the number of potential basis function centers in that direction. The two other shape parameter values are created by multiplying or dividing by $\sqrt{2}$ respectively.

By combining the center locations and shape parameter we obtain a set of 123 potential basis functions for both selection measures, shown in figure 3.8.

3.5 Scaling of basis functions

From the relation between the fine-scale gradient and the coarse-scale gradient (3.12), the coarse-scale gradient can be viewed as a weighted sum of the fine-scale gradient. The weighting is determined by the structure matrix Θ . As the coarse-scale gradient is used in the refinement process for the 1. order selection measure, the “structure” of the structure

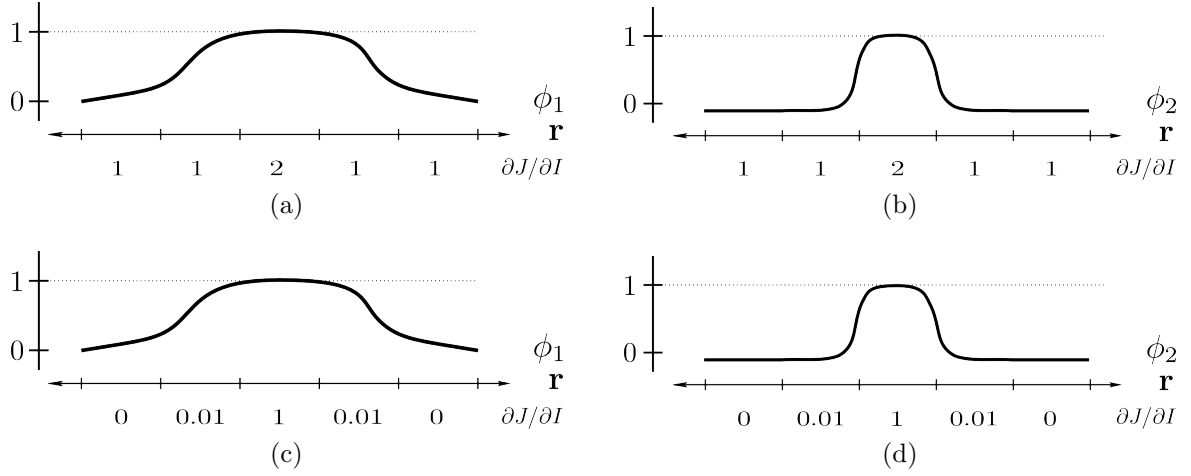


Figure 3.9: Figures (a) and (c) shows a wide radial basis function ϕ_1 scaled by the max-norm, on two domains with different sensitivities $\partial J/\partial I$. Figures (b) and (c) shows the same for a narrower radial basis function ϕ_2 , on the same domains. $\phi_k = \phi(\sigma_k, \mathbf{r}_k)$ with corresponding coefficient a_k .

matrix has a direct effect when choosing a new parameterization.

Both the trilinear basis functions and the radial basis functions has a maximum value $\theta_k = 1$ at their center. If we scale the basis functions, will this affect the coarse-scale gradient and thus the choices made in the refinement process? The answer is yes. In fact, we already applied scaling to these basis functions. Consider the formula 3.4 for the radial basis functions, with a slight modification;

$$\phi_k(\mathbf{r}) = \frac{1}{\lambda_k} \exp\left(-\frac{1}{2\sigma_k^2}(\mathbf{r} - \mathbf{r}_k)^2\right) = \frac{1}{\lambda_k}\theta_k(\mathbf{r}), \quad (3.19)$$

where λ_k is a scaling coefficient. Let λ_k be the max-norm

$$\lambda_k = \max_{\mathbf{r}} |\theta_k(\mathbf{r})|. \quad (3.20)$$

Thus the scaled basis function 3.19 with the scaling given by equation 3.20 ($\lambda_k = 1$) is the same basis function as 3.4. In that sense, the radial basis functions as we presented them are scaled by the max-norm. Consider the case illustrated in Figure 3.9. Let $\phi_k = \phi_k(\mathbf{r})$ be a radial basis function with center \mathbf{r}_k , shape parameter σ_k and corresponding coefficient a_k , where $k = 1, 2$. In the situation on (a) and (b), the weighted sum $|\nabla_{a_k} J|$ will be larger for function ϕ_1 than for ϕ_2 , thus ϕ_1 will be preferred. This is a desired outcome, as the sensitivity is high throughout the domain. However, in the situation in (c) and (d), the sensitivity is concentrated in the center interval, thus the narrowest function ϕ_2 should be preferred. But as in the first situation, $|\nabla_{a_1} J|$ is larger than $|\nabla_{a_2} J|$, and thus ϕ_1 is chosen. We therefore conclude that for basis functions scaled by the max-norm the 1. order selection measure will prefer wide basis functions, even in situations where it is clearly not the best choice.

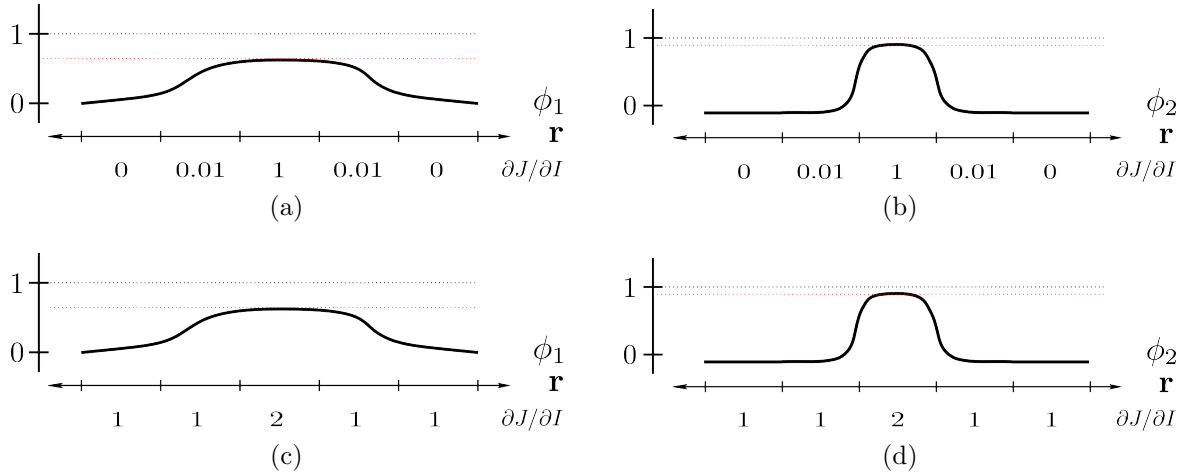


Figure 3.10: Figures (a) and (c) shows a wide radial basis function ϕ_1 scaled by the L_1 -norm, on two domains with different sensitivities $\partial J/\partial I$. Figures (b) and (c) shows the same for a narrower radial basis function ϕ_2 , on the same domains. $\phi_k = \phi(\sigma_k, \mathbf{r}_k)$ with corresponding coefficient a_k .

Consider an alternative scaling, the L_1 -norm:

$$\lambda_k = \sum_{\mathbf{r}} |\theta_k(\mathbf{r})|. \quad (3.21)$$

We scale the new radial basis functions with the sum of the absolute values of the basis functions for all grid cells. This will ensure that a wide basis function as in figure 3.9 is not preferred over a narrow basis function, when they cover the same areas of sensitivity. Consider Figure 3.10, (a) and (b). As ϕ_1 and ϕ_2 are scaled with the L_1 -norm, the weighted sum $|\nabla_{a_2} J|$ is larger than $|\nabla_{a_1} J|$, and ϕ_2 will be preferred as a new basis function. This is the best choice, but for (c) and (d), the situation is different. Because of a wider sensitivity domain, we would prefer the wide basis function ϕ_1 , but as the weighted sum $|\nabla_{a_2} J|$ dominates, ϕ_2 is chosen.

Clearly none of these two norms are perfect, in the sense that we came up with simple problems for which both scaling factors favour the “wrong” basis function. In a case where both basis functions cover the area of most interest, the max-norm favours wide basis functions, while the L_1 -norm as a scaling factor favours narrow basis functions.

As the L_1 -norm and the max-norm are extreme norms in a certain sense, see Figure 3.11, we know of other norms that could perhaps result in better choices. The L_p -norms are obvious choices, defined as

$$\lambda_k = \left[\sum_{\mathbf{r}} \theta_k(\mathbf{r})^p \right]^{1/p}, \quad p = 1, 2, 3, \dots \quad (3.22)$$

In fact, any norm would suffice as scaling factor, but we limit ourselves to compare the L_1 -, L_2 - and max-norm in several numerical examples, see also comment in Chapter 5.

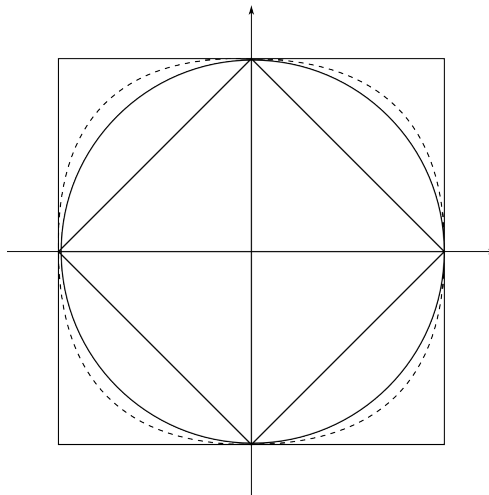


Figure 3.11: Unit balls in \mathbb{R}^2 for different norms. From center and outwards: L_1 -norm, L_2 -norm, L_p -norm and max-norm. The L_p -norm is shown with a dashed line as its shape depends on the value of p .

Chapter 4

Optimization and sensitivity calculations

Chapter 3 presented the parameter representation, the multi-level strategy, and different methods for refinement. In this chapter we consider the two remaining subjects that were sketched in Figure 3.4; the optimization process and the termination criterion. We also consider the sensitivity matrix calculations, and present the quasi-analytical approximation involved in the 1. order selection measure.

4.1 Optimization

We have a set of coefficients $\{a_k\}_{k=1}^{N_a}$ from which we wish to recreate the observations $\{d_j\}_{j=1}^{N_d}$. The coefficients and the data are related by the nonlinear system of equations (2.3) and the composite parameter representation (3.1). Our goal is to find coefficients whose output $\mathbf{G}(m(\mathbf{r}; \mathbf{a}))$ best fit the data \mathbf{d} . This is done by finding the interior coefficients that minimize the objective function $J(m(\mathbf{r}; \mathbf{a}))$ from equation (2.5). As mentioned we set the regularization term $\phi(m(\mathbf{r}; \mathbf{a}))$ to zero, and for simplicity, let $J(\mathbf{a}) = J(m(\mathbf{r}; \mathbf{a}))$ and $J_h = J(m(\mathbf{r}; \mathbf{a}_h))$.

We start out with a set of initial coefficients \mathbf{a}_0 , and use the optimization routine to generate a sequence of iterates $\{\mathbf{a}_h\}_{h=1}^{\infty}$, and then terminate when a solution criterion has been met or there has been no update in objective function value. To generate the iterate \mathbf{a}_{h+1} , the optimization routine may use information from the previous iterates, \mathbf{a}_h , or even earlier iterates $\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{h-1}$. In general, all new iterates will result in a decreased objective function value, $J(\mathbf{a}_{h+1}) < J(\mathbf{a}_h)$. Before we present our optimization routine, we consider the two fundamental strategies for choosing new iterates.

For a *line-search strategy*, the routine chooses a direction \mathbf{q}_h , and search along this direction to find a new iterate

$$\mathbf{a}_{h+1} = \mathbf{a}_h + \alpha \mathbf{q}_h. \quad (4.1)$$

The step length is chosen by approximately solving the minimization problem

$$\min_{\alpha > 0} J(\mathbf{a}_h + \alpha \mathbf{q}_h). \quad (4.2)$$

Equation (4.2) can be solved exactly, but due to computational costs it is common to do some trial steps, and choose an approximation to this minimizer. The process is repeated until it terminates.

The *thrust-region strategy* collects information about the objective function in order to make a *model function* p_h that approximates the behaviour of the objective function J in the vicinity of the previous iterate \mathbf{a}_h . Both elliptic, box-shaped and spherical thrust regions can be used. We consider a spherical thrust region. This region is defined by a radius $\Delta > 0$, and within it we find the best minimizing iterate $\mathbf{a}_h + \mathbf{q}$, that is we solve the minimization problem

$$\min_{\mathbf{q}} p_h(m(\mathbf{r}; \mathbf{a}_h + \mathbf{q})) \quad \|\mathbf{q}\| < \Delta. \quad (4.3)$$

If the candidate \mathbf{q} does not produce sufficient decrease in J , we shrink the thrust region, and re-solve the minimization problem.

One way to describe the difference between these two methods is to say that while the line-search method chooses the direction first, and then find the step length, the thrust-region determines the maximum “distance” Δ first, and then decides on the direction and actual step length. If this is unsatisfactory, we reduce the distance measure Δ , and try again [16].

We use the *Levenberg-Marquardt method* as our optimization routine. It is a thrust-region strategy, but in order to describe it, we start out with the outline of a line-search method, the *Gauss-Newton method*.

4.1.1 The Gauss-Newton method

We wish to minimize the objective function $J(\mathbf{a})$. Assume that the coefficients can be written as $\mathbf{a} = \mathbf{a}_0 + \Delta \mathbf{a}$, where \mathbf{a}_0 is some initial guess. If J is twice differentiable, its Taylor series expansion is

$$J(\mathbf{a}_0 + \Delta \mathbf{a}) \approx J_0 + (\nabla J_0)^T \Delta \mathbf{a} + \frac{1}{2} \Delta \mathbf{a}^T (\nabla^2 J_0) \Delta \mathbf{a} \quad (4.4)$$

where

$$\nabla J_0 = (\partial J_0 / \partial a_1, \dots, \partial J_0 / \partial a_{N_d})^T, \quad (4.5)$$

and $\nabla^2 J_0$ is the Hessian matrix

$$\nabla^2 J_0 = \begin{bmatrix} \frac{\partial^2 J_0}{\partial a_1^2} & \dots & \frac{\partial^2 J_0}{\partial a_1 \partial a_{N_d}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 J_0}{\partial a_{N_d} \partial a_1} & \dots & \frac{\partial^2 J_0}{\partial a_{N_d}^2} \end{bmatrix}. \quad (4.6)$$

Let \mathbf{a}^* be the minimizer for J . Thus $\nabla J(\mathbf{a}^*) = \mathbf{0}$. By Taylor expansion around \mathbf{a}^* , it follows that

$$\nabla J(\mathbf{a}^*) = \nabla J(\mathbf{a}_0 + \Delta\mathbf{a}) \approx \nabla J_0 + \nabla^2 J_0 \Delta\mathbf{a}. \quad (4.7)$$

As the approximate gradient in this equation is zero, we rearrange it and obtain

$$\nabla^2 J_0 \Delta\mathbf{a} = -\nabla J_0. \quad (4.8)$$

The arbitrary initial guess \mathbf{a}_0 is replaced by a coefficient at a general step h , \mathbf{a}_h , and equation (4.8) may be written as;

$$\nabla^2 J_h \Delta\mathbf{a} = -\nabla J_h. \quad (4.9)$$

The iterative method

$$\mathbf{a}_{h+1} = \mathbf{a}_h + \Delta\mathbf{a}, \quad (4.10)$$

where $\Delta\mathbf{a}$ is given by (4.9), is known as the Newtons method for minimizing function $J(\mathbf{a})$.

Next, rewrite the objective function (2.5) as

$$J = \sum_{j=1}^{N_d} R_j(m(\mathbf{r}; \mathbf{a}))^2, \quad (4.11)$$

where the R_j is referred to as the residuals. The residuals form the vector $\mathbf{R}(\mathbf{a}) = \mathbf{R}(m(\mathbf{r}; \mathbf{a}))$;

$$\mathbf{R}(\mathbf{a}) = (R_1(m(\mathbf{r}; \mathbf{a})), R_2(m(\mathbf{r}; \mathbf{a})), \dots, R_{N_d}(m(\mathbf{r}; \mathbf{a})))^T. \quad (4.12)$$

Here

$$R_j(m(\mathbf{r}; \mathbf{a})) = \frac{\mathbf{G}(m(\mathbf{r}; \mathbf{a}))_j - d_j}{\gamma_j} \quad j = 1, 2, \dots, N_d, \quad (4.13)$$

where γ_j is the standard deviation in measurement j . We now collect the derivatives of $J(\mathbf{a})$ in the $N_d \times N_g$ Jacobian matrix $\mathbf{S}(m(\mathbf{r}; \mathbf{a}))$ of partial derivatives:

$$\mathbf{S}(m(\mathbf{r}; \mathbf{a})) = \left[\frac{\partial R_j}{\partial m_i} \right]_{\substack{j=1,2,\dots,N_d \\ i=1,2,\dots,N_g}}. \quad (4.14)$$

Let $\mathbf{S} = \mathbf{S}(m(\mathbf{r}; \mathbf{a}))$ and $\mathbf{S}_h = \mathbf{S}(m(\mathbf{r}; \mathbf{a}_h))$. The gradient $\nabla J(\mathbf{a})$ and Hessian matrix $\nabla^2 J(\mathbf{a})$ of the objective function are given by

$$\nabla J(\mathbf{a}) = 2 \sum_{j=1}^{N_d} R_j(m(\mathbf{r}; \mathbf{a})) \nabla R_j(m(\mathbf{r}; \mathbf{a})) = \mathbf{S}^H \mathbf{R}(m(\mathbf{r}; \mathbf{a})), \quad (4.15)$$

$$\begin{aligned} \nabla^2 J(\mathbf{a}) &= 2 \left(\sum_{j=1}^{N_d} \nabla R_j(m(\mathbf{r}; \mathbf{a})) \nabla R_j(m(\mathbf{r}; \mathbf{a}))^H + \sum_{j=1}^{N_d} R_j(m(\mathbf{r}; \mathbf{a})) \nabla^2 R_j(m(\mathbf{r}; \mathbf{a})) \right) \\ &= 2\mathbf{S}^T \mathbf{S} + 2 \sum_{j=1}^{N_d} R_j(m(\mathbf{r}; \mathbf{a})) \nabla^2 R_j(m(\mathbf{r}; \mathbf{a})). \end{aligned} \quad (4.16)$$

We ignore the second term in the Hessian matrix, by assuming near-linearity of the model near the solution (small $\nabla R_j(m(\mathbf{r}; \mathbf{a}))$), or small residuals $R_j(m(\mathbf{r}; \mathbf{a}))$. Thus the approximated Hessian matrix includes only the first term of the exact Hessian matrix (4.16), which is the matrix product of the Jacobian with itself, and consequently can be calculated explicitly for our problem. Insert equation (4.15) and (4.16) into (4.9) for a step h and obtain

$$\mathbf{S}_h^H \mathbf{S}_h \Delta \mathbf{a} = -\mathbf{S}_h^H \mathbf{R}(\mathbf{a}_h), \quad (4.17)$$

and then solve the system of equations (4.17) and insert for $\Delta \mathbf{a}$ in the iterative step (4.10) to obtain the Gauss-Newton method.

4.1.2 The Levenberg-Marquardt method

One weakness with the Gauss-Newton method is its performance when the Jacobian matrix S is rank-deficient. Replacing the line-search strategy with a thrust-region strategy may increase the performance in that situation. We still ignore the second term in the Hessian matrix (4.16), of the same reason as above.

We must reformulate our problem to accommodate a thrust-region method. Our new minimization problem is

$$\min_{\mathbf{q}} \frac{1}{2} \|\mathbf{S}_h \mathbf{q} + \mathbf{R}(\mathbf{a}_h)\|_2^2 \quad \|\mathbf{q}\| < \Delta_h, \quad (4.18)$$

rather than minimizing the objective function J , which is equivalent to a model function p_h given as

$$p_h(q) = \frac{1}{2} \|\mathbf{R}(\mathbf{a}_h)\|_2^2 + \mathbf{q}^T \mathbf{S}_h^T \mathbf{R}(\mathbf{a}_h) + \frac{1}{2} \mathbf{q}^T \mathbf{S}_h^T \mathbf{S}_h \mathbf{q}. \quad (4.19)$$

For simplicity we suppress the iteration counter h . From [15] we know that when a solution to the Gauss-Newton equations (4.17) lies strictly inside the thrust region, it also solves the thrust-region problem (4.18), and if that is not the case, there exists a $\lambda > 0$ such that $\mathbf{q} = \mathbf{q}^{LM}$ satisfies $\|\mathbf{q}\| = \Delta$, and

$$(\mathbf{S}^T \mathbf{S} + \lambda \mathbf{I}) \mathbf{q} = -\mathbf{S}^T \mathbf{R}(\mathbf{a}). \quad (4.20)$$

This is known as the Levenberg-Marquardt method [16].

4.2 Termination criterion

The multi-level strategy needs a termination criterion. As the measurements contain noise that we cannot exclude from the calculations, we need a target value for the objective function for which to terminate when reached. We assume the measurements error to be normally distributed, and that $\mathbf{G}(P^l)$ is close to linear in the region around the minimiser of $J(P^l)$. Remember that P^l is a parameterization candidate, defined in Section 3.4. The minimum value of the objective function at the solution of the inverse problem $J^*(P^l)$ is

then χ^2 -distributed, with $N_d - N_a$ degrees of freedom. From the theory of χ^2 -distributions we know that

$$E(J^*(P^l)) = N_d - N_a \quad (4.21)$$

$$\zeta(J^*(P^l)) = \sqrt{2(N_d - N_a)} \quad (4.22)$$

where E is the expectation value and ζ the standard deviation. In our work the termination criterion is set to be the expectation value plus one standard deviation,

$$J(P^l) < N_d - N_a + \sqrt{2(N_d - N_a)}. \quad (4.23)$$

4.3 Sensitivity matrix

The sensitivity matrix plays an important role in this thesis. The 2. order selection measure presented in Section 3.4.1 need it to estimate how sensitive the model output is to the interior coefficients. The Levenberg-Marquardt method uses it in the optimization process. In this section we will derive the sensitivity matrix for the IE method presented in Section 1.4.

Recall that $\{(\mathbf{G}(m(\mathbf{r}; \mathbf{a})))_j\}_{j=1}^{N_d}$ and $\{a_k\}_{k=1}^{N_a}$ are model output and interior coefficients respectively. Let \mathbf{F} denote both \mathbf{E} and \mathbf{H} , so that $\mathbf{G}(m(\mathbf{r}; \mathbf{a})) = \Delta\mathbf{F}(\mathbf{r})$. Upon solving Maxwell's equations (1.9) - (1.12), we consider the expression for the time-lapse signal \mathbf{F} :

$$\begin{aligned} \Delta\mathbf{F} &= \int_D \mathcal{G}_F \sigma_{n+1}^a (\mathbf{E}_{n+1}^a + \mathbf{E}_{n+1}^b) dV - \int_D \mathcal{G}_F \sigma_n^a (\mathbf{E}_n^a + \mathbf{E}_n^b) dV \\ &= \int_D \mathcal{G}_F [\sigma_{n+1}^a (\mathbf{E}_{n+1}^a + \mathbf{E}_{n+1}^b) - \sigma_n^a (\mathbf{E}_n^a + \mathbf{E}_n^b)] dV \\ &= \int_D \mathcal{G}_F [\sigma_n^a (\mathbf{E}_{n+1}^a + \mathbf{E}_n^a) - \Delta\sigma (\mathbf{E}_{n+1}^a + \mathbf{E}^b)] dV. \end{aligned} \quad (4.24)$$

We used that $\mathbf{E}_{n+1}^b = \mathbf{E}_n^b = \mathbf{E}^b$, as the background electric field component is constant. We also inserted for the electric conductivity $\sigma_{n+1}^a = \sigma_n^a + \Delta\sigma$. The subscript n refers to a survey at time t , and subscript $n + 1$ refers to a survey conducted at a time $t + \Delta t$.

The expression for sensitivity of time-lapse signal with respect to time-lapse conductivity is now given as

$$\begin{aligned} \frac{\partial \Delta\mathbf{F}}{\partial \Delta\sigma} &= \frac{\partial}{\partial \Delta\sigma} \int_D \mathcal{G}_F [\sigma_n^a (\mathbf{E}_{n+1}^a + \mathbf{E}_n^a) - \Delta\sigma (\mathbf{E}_{n+1}^a + \mathbf{E}^b)] dV \\ &= \int_D \mathcal{G}_F \left[\sigma_n^a \frac{\partial}{\partial \Delta\sigma} (\mathbf{E}_{n+1}^a + \mathbf{E}_n^a) - \frac{\partial}{\partial \Delta\sigma} (\Delta\sigma (\mathbf{E}_{n+1}^a + \mathbf{E}^b)) \right] dV \\ &= \int_D \mathcal{G}_F \left[\sigma_n^a \frac{\partial (\mathbf{E}_{n+1}^a + \mathbf{E}_n^a)}{\partial \Delta\sigma} - (\mathbf{E}_{n+1}^a + \mathbf{E}^b) - \frac{\partial \mathbf{E}_{n+1}^a}{\partial \Delta\sigma} \right] dV \\ &= \int_D \mathcal{G}_F \left[(\sigma_n^a + \Delta\sigma) \frac{\partial (\mathbf{E}_{n+1}^a - \mathbf{E}_n^a)}{\partial \Delta\sigma} + (\mathbf{E}_{n+1}^a + \mathbf{E}^b) \right] dV, \end{aligned} \quad (4.25)$$

where we have used that $\partial \mathbf{E}_n^b / \partial \Delta \sigma = 0$, $\partial \mathbf{E}_n^a / \partial \Delta \sigma = 0$ and inserted for $\partial \mathbf{E}_{n+1}^a / \partial \Delta \sigma = \partial(\mathbf{E}_{n+1}^a - \mathbf{E}_n^a) / \partial \Delta \sigma$. The expression (4.25) is in fact the general sensitivity matrix, $\partial \Delta \mathbf{F} / \partial \Delta \sigma$. Given our composite representation (3.1), we let $\Delta \sigma = m(\mathbf{r}; \mathbf{a})$ to obtain

$$\frac{\partial \Delta \mathbf{F}}{\partial a_k} = \int_D \mathcal{G}_F[(\sigma_n^a + m(\mathbf{r}; \mathbf{a})) \frac{\partial(\mathbf{E}_{n+1}^a - \mathbf{E}_n^a)}{\partial a_k} + (\mathbf{E}_{n+1}^a + \mathbf{E}^b) \frac{\partial m(\mathbf{r}; \mathbf{a})}{\partial a_k}] dV. \quad (4.26)$$

Recall from Section 3.4.2, that the partial derivative $\partial m / \partial a_k$ can be found from the identities

$$\frac{\partial m}{\partial I} = (c_1 - c_2) \frac{1}{\pi(1 + I^2)}, \quad \text{and} \quad \frac{\partial I}{\partial a_k} = \theta_k. \quad (4.27)$$

so that we are able to calculate each term in the sensitivity matrix (4.26).

4.3.1 Numerical calculation of the sensitivity matrix

We found an expression for the sensitivity matrix through equation (4.26) above. The Green's tensors and \mathbf{E}_{n+1} are available from the forward simulations, and $\partial m(\mathbf{r}; \mathbf{a}) / \partial a_k$ can be found by matrix multiplication. The computationally demanding part is to find $\partial(\mathbf{E}_{n+1}^a - \mathbf{E}_n^a) / \partial a_k$ within D [7].

Let $\widehat{\mathbf{S}}$ be the $3N_g \times N_a$ sensitivity matrix,

$$\widehat{\mathbf{S}} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{N_a}], \quad (4.28)$$

where each \mathbf{s}_k is a $3N_g \times 1$ column vector with entries $\partial(\mathbf{e}_{n+1}^a - \mathbf{e}_n^a) / \partial a_k$, and \mathbf{e}^a is given as a numerical discretization of \mathbf{E}^a as defined by equation (1.16). Consider $\mathbf{F} = \mathbf{E}$, and let \mathbf{v}_k be a $3N_g \times 1$ column vector defined as

$$\mathbf{v}_k = [E_{x,1} \frac{\partial m_1}{\partial a_k} \quad \dots \quad E_{x,N_g} \frac{\partial m_{N_g}}{\partial a_k}, \\ E_{y,1} \frac{\partial m_1}{\partial a_k} \quad \dots \quad E_{y,N_g} \frac{\partial m_{N_g}}{\partial a_k}, E_{z,1} \frac{\partial m_1}{\partial a_k} \quad \dots \quad E_{z,N_g} \frac{\partial m_{N_g}}{\partial a_k}]^T. \quad (4.29)$$

Finally remember $\widehat{\mathbf{G}}_D$ as defined in Section 1.5. Then the sensitivity matrix (4.26) can be written as

$$\mathbf{s}_k = \widehat{\mathbf{G}}_D(\Sigma_{n+1}^a \mathbf{s}_k + \mathbf{v}_k), \quad k = 1, 2, \dots, N_a. \quad (4.30)$$

Rearrange to get

$$(\mathbf{I} - \widehat{\mathbf{G}}_D \Sigma_{n+1}^a) \mathbf{s}_k = \widehat{\mathbf{G}}_D \mathbf{v}_k, \quad k = 1, 2, \dots, N_a, \quad (4.31)$$

hence the $3N_g \times 3N_g$ coefficient matrix $\mathbf{A} = \mathbf{I} - \widehat{\mathbf{G}}_D \Sigma_{n+1}^a$ from the discretized Fredholm integral equation (1.19) appears, and we write our system as

$$\mathbf{A} \mathbf{s}_k = \widehat{\mathbf{G}}_D \mathbf{v}_k \quad k = 1, 2, \dots, N_a. \quad (4.32)$$

This enables us to consider the sensitivity calculations during the forward calculations. For solving equations (4.32), we utilize the BICGSTAB method with preconditioning (Section 1.5). The extra calculations will increase the costs with about 20% of a forward simulation, for each sensitivity coefficient [7]. Thus the computational cost increase rapidly with the number of coefficients in the representation.

4.3.2 Quasi-analytical approximation

We now present the QA approximations associated with the 1.order selection measure. Recall the Integral Equation solution 1.10 for the anomalous part of the electric field:

$$\mathbf{E}^a(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})(\mathbf{E}^b(\mathbf{r}) + \mathbf{E}^a(\mathbf{r}))dV. \quad (4.33)$$

The quasi-analytical approximation as presented by [18] is based on the assumption that the anomalous field \mathbf{E}^a inside the anomalous domain is linearly proportional to the background field through some tensor \hat{Q} ;

$$\mathbf{E}^a(\mathbf{r}) \approx \hat{Q}(\mathbf{r})\mathbf{E}^b(\mathbf{r}). \quad (4.34)$$

Substituting the approximation (4.34) into equation (4.33), we obtain

$$\mathbf{E}^a(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})(\mathbf{I} + \hat{Q}(\mathbf{r}))\mathbf{E}^b(\mathbf{r})dV, \quad (4.35)$$

where \mathbf{I} is the identity tensor. So far we have presented what is known as the quasi-linear (QL) approximation. If we further assume the tensor to be a scalar one,

$$\hat{Q} = Q\mathbf{I}, \quad (4.36)$$

then the anomalous field can be written as

$$Q(\mathbf{r})\mathbf{E}^b(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})(1 + Q(\mathbf{r}))\mathbf{E}^b(\mathbf{r})dV. \quad (4.37)$$

For a slowly varying $Q(\mathbf{r})$ in D , it can be moved outside the integration

$$Q(\mathbf{r})\mathbf{E}^b(\mathbf{r}) \approx (1 + Q(\mathbf{r}))\mathbf{E}^B(\mathbf{r}), \quad (4.38)$$

where $\mathbf{E}^B(\mathbf{r})$ denotes the Born approximation:

$$\mathbf{E}^B(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r})\sigma^a(\mathbf{r})\mathbf{E}^b(\mathbf{r})dV. \quad (4.39)$$

Multiply with the complex conjugate of \mathbf{E}^b on both sides of equation 4.38 to obtain

$$Q(\mathbf{r})\mathbf{E}^b(\mathbf{r}) \cdot \mathbf{E}^{b*}(\mathbf{r}) \approx (1 + Q(\mathbf{r}))\mathbf{E}^B(\mathbf{r}) \cdot \mathbf{e}^{b*}, \quad (4.40)$$

from which we obtain an expression for $Q(\mathbf{r})$:

$$Q(\mathbf{r}) = \frac{g(\mathbf{r})}{1 - g(\mathbf{r})}. \quad (4.41)$$

where

$$g(\mathbf{r}) = \frac{\mathbf{E}^B(\mathbf{r}) \cdot \mathbf{E}^{b*}(\mathbf{r})}{\mathbf{E}^b(\mathbf{r}) \cdot \mathbf{E}^{b*}(\mathbf{r})}. \quad (4.42)$$

Substituting the expression (4.41) into Equation (4.35), we finally determine

$$\mathbf{E}_{QA}^a(\mathbf{r}') = \int_{\Omega} \mathcal{G}_E(\mathbf{r}'|\mathbf{r}) \frac{\sigma^a(\mathbf{r})}{1 - g(\mathbf{r})} \mathbf{E}^b(\mathbf{r}) dV. \quad (4.43)$$

Note that this QA approximation is similar to the Born approximation (4.39) apart from the scalar function $(1 - g(\mathbf{r}))^{-1}$. The computational cost of generating the QA approximation and the Born approximation is practically the same, and negligible compared to the full IE forward solver. We use this QA approximation to obtain an approximated sensitivity matrix, and solve the system (3.18) with an adjoint method.

Chapter 5

Numerical examples

The theory presented in previous chapters will now be employed in several numerical examples. In the examples we consider identification of time-lapse electric conductivity changes corresponding to saturation changes within oil reservoir structures from time-lapse CSEM data.

The various example settings are presented in Section 5.1, followed by a discussion regarding the effect of scaling of the basis functions follows, as well as the use of a multi-level strategy. Our two main results involving the performance of radial basis functions, and the introduction of a 1. order selection measure, are presented in Section 5.4 and 5.5 respectively.

The last section present the two main results, the performance of radial basis functions (RBFs) in parameter representation in CSEM inversion, and the performance of a first order selection measure in the refinement process.

5.1 Example setup

We use the forward solver based on [19, 11] for the forward simulations using equations on (1.9) - (1.12). The background electric conductivity is set to $\sigma = 3.33$ S/m in the sea water column, and $\sigma = 1$ S/m in the water saturated subsediments. The electric conductivity of hydrocarbon-filled and water-filled areas of the reservoir is dependent on the value of the interior function and the asymptotic values for the parameter representation. For the hydrocarbon saturated parts of the reservoir the “target value” is $\sigma = 0.001$ S/m, whilst the corresponding value for water-filled parts is $\sigma = 1$ S/m, just as for the water saturated subsediments. The reservoir is located 1055 meters below the seafloor, and it covers the volume $D = \{\mathbf{r} | x \in [0, 2400], y \in [0, 2400], z \in [2020, 2120]\}$ m³. The forward simulations are made on the computational grid that was presented in Section 1.5.1.

The receivers are distributed at the seafloor, 995 meters below the surface. They measure the field components E_x, E_y, E_z, H_x and H_y . The 100 meters long sources are towed 40 meters above the seafloor. The fields are created by a current of 1000 A and 0.5 Hz frequency. We use two different setups for the sources and receivers, as presented

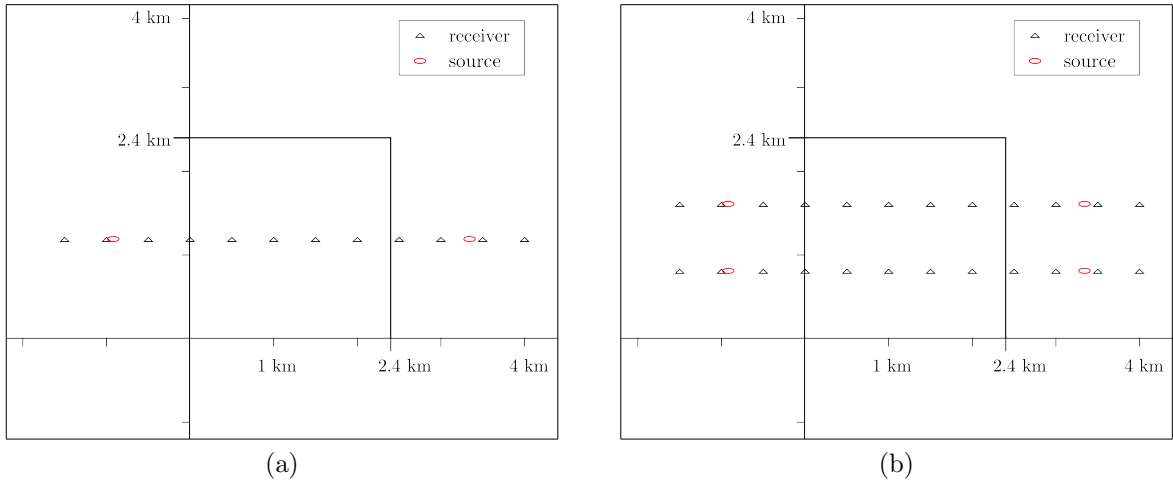


Figure 5.1: In setup (a) the 12 receivers are distributed in a line at $y = 1200$ meters, from $x = -1500$ meters to $x = 4000$ meters, with an equidistance of 500 meters. The two sources are located along the same line, at approximately $x = -1000$ meters and $x = 3400$ meters. For setup (b) we have two parallel lines with 12 receivers at $y = 800$ meters and $y = 1600$ meters, with 500 meters between each receiver. Correspondingly there are four sources, allocated along these lines, with the same distance to the reservoir as the sources in setup (a).

in Figure 5.1.

For the examples in this thesis we include some Gaussian noise of 1% in the reference conductivity field, to represent small variations in e.g. porosity or saturation. To represent measurement error in the observations, we include 5% Gaussian noise.

We use the multi-level estimation strategy outlined in Section 3.3. If the solution criterion is met, or if the objective function value does not improve in three successive levels, the estimation process will be terminated.

5.2 Scaling of basis functions

Since there is a possibility of significant effect from scaling, as indicated in Section 3.5, we include an investigation of the performance when using these different scalings. Due to low sensitivity of the interior coefficients with respect to time-lapse variations, we seek a representation with broad support. However, the representation should have the flexibility to represent different structures. We include an example to illustrate some trends in our results.

As Figure 5.2 shows, the basis functions scaled with the max-norm or the L_2 -norm provide better results than when the L_1 -norm. We know employing the max-norm could result in “wider” basis functions even though a “narrower” choice would lead to better results, so we choose the L_2 -norm as a scaling factor for our basis functions.

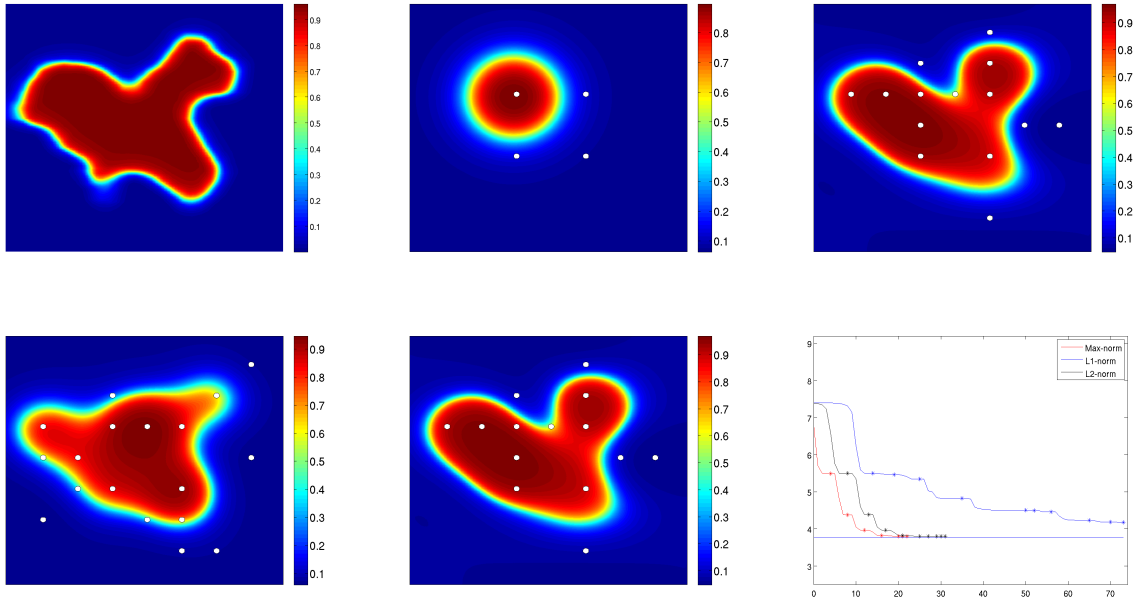


Figure 5.2: Top left to bottom right: Reference time-lapse conductivity field, initial time-lapse conductivity, final time-lapse conductivity estimates for RBFs scaled with the max-norm, L_1 -norm and L_2 -norm respectively, and objective function values versus number of iterations for all choices of scalings. White points mark the center of each basis function.

5.3 A multi-level strategy compared to a one-level strategy

We have chosen a multi-level strategy for solving the parameter estimation problem. This strategy keeps the final number of basis functions N_a unknown until the solution criterion is met or the maximum number of levels is exceeded. Our aim is to find the N_a basis functions best suited to solve the problem. This is a combination of two aspects; Not only do we seek the N_a basis functions that best corresponds to the information content of the available data, we need a suitable choice of basis functions with respect to center location and shape parameters. The optimization routine seeks to obtain the best coefficient values, given the parameterization.

In this section we compare the use of a multi-level strategy with a one-level strategy, for which all potential basis functions, described in Section 3.4.3, are included in the optimization. We include two figures to illustrate the results of the multi-level and one-level strategies. Figure 5.3 shows the results when a setup as in Figure 5.1a is used. The multi-level approach provides a better solution than the one-level approach, both for the objective function value and the field estimate.

Figure 5.4 shows the results when more data is included. Here we use 4 sources and 144 receivers. The receivers are distributed in a 12×12 grid on the seafloor, and the sources

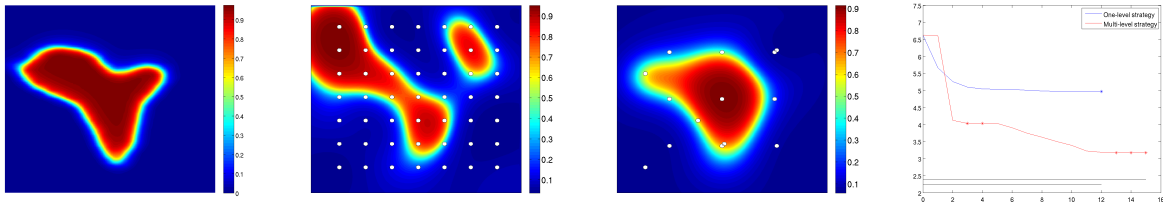


Figure 5.3: Left to right: Reference time-lapse conductivity field, final time-lapse conductivity estimates for one-level strategy, final time-lapse conductivity estimates for multi-level strategy, and objective function values versus number of iterations. White points mark the center of each basis function.

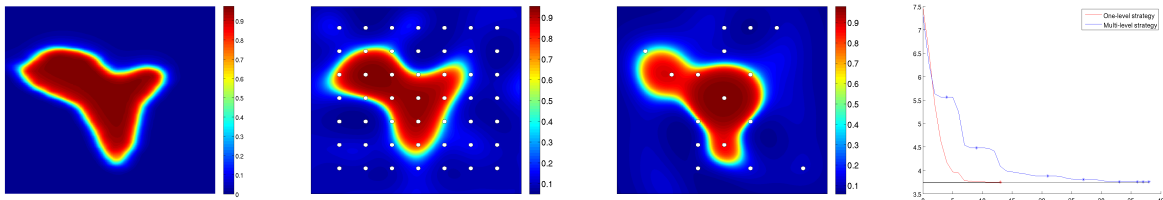


Figure 5.4: Left to right: Reference time-lapse conductivity field, final time-lapse conductivity estimates for one-level strategy, final time-lapse conductivity estimates for multi-level strategy, and objective function values versus number of iterations. White points mark the center of each basis function.

placed on all four sides. The one-level strategy results in a visual conductivity estimate nearly identical to the reference conductivity field, and it also meets the solution criterion.

Figures 5.3 and 5.4 show the importance of the quality of the data we consider for inversion. They also show the importance of using an appropriate set of N_a basis functions. The choice of this set can generally not be determined before the inversion. The multi-level strategy with its possibility to determine N_a during the process, can be a robust alternative to other strategies in inversion of CSEM observations.

5.4 Comparison of different basis functions in parameter representation

In this section, we compare the use of trilinear basis functions to radial basis functions in parameter representation. We consider three examples, one of which is a three-dimensional reservoir. This is included to show that our solution methodology works in three dimensions as well as in two dimensions, and also to show that results obtained in two dimensional cases can be generalized into three dimensions. For the adaptive refinement process, we use the 2. order selection measure as presented in Section 3.4.1. The initial guess consists of one

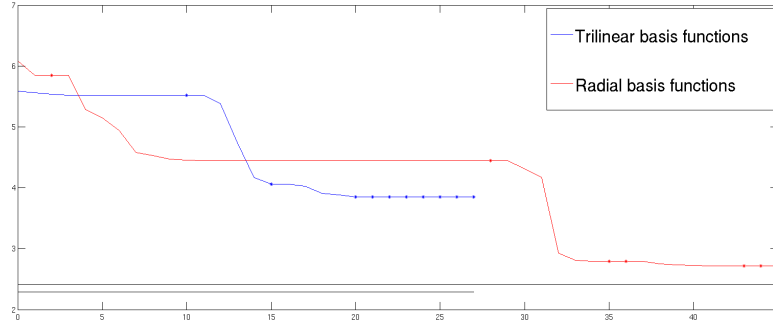


Figure 5.5: The log-objective function values versus number of iterations, for both choices of basis functions. The inverse problem solution criterion for the last level is indicated by a horizontal line where the length corresponds to the number of iterations in the optimization algorithm. A star marks the end of the optimization process at a given level.

basis function with center in the injection well. We consider two types of basis functions, and the initial guesses are as similar as possible given the different parameterizations. The effect of different initial guesses is not believed to have a decisive impact on the final estimates, so we treat the basis functions equally in the comparison.

5.4.1 Example 1

In the first example, we consider a two-dimensional reservoir with a time-lapse conductivity change illustrated in Figure 5.6. We distribute the sources and receivers as sketched in Figure 5.1a.

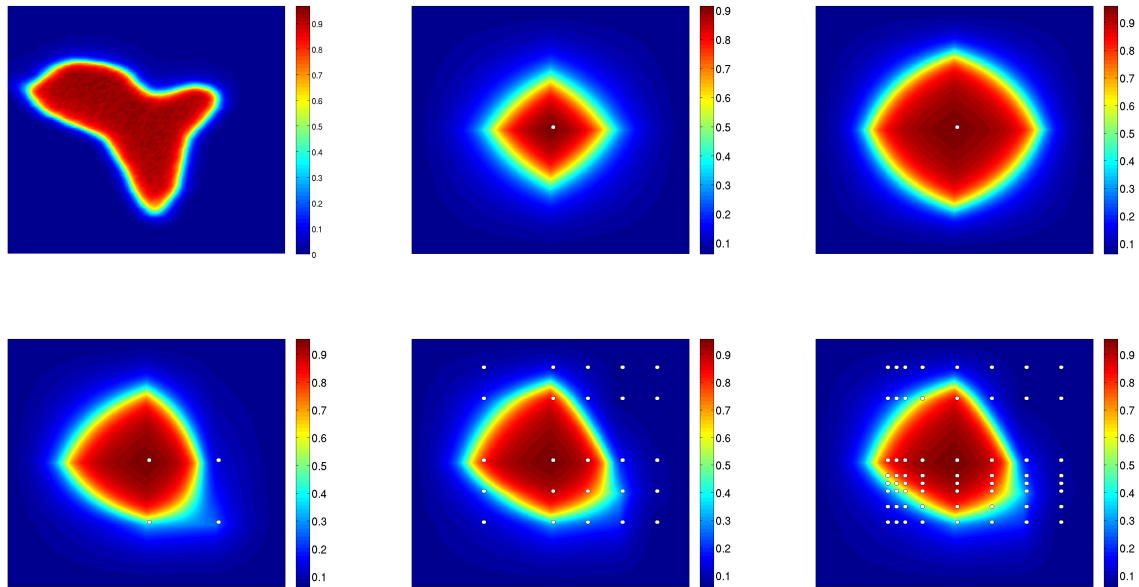
Figure 5.5 shows the resulting objective function values at each iteration for different levels. Both types of basis functions give a significantly improvement in the objective function value from level $s = 1$ to $s = 2$. The TBFs provide a slight decrease in objective function value from level $s = 2$ to level $s = 3$, but after that only minor improvements are made in the remaining levels. This is illustrated in Figure 5.6. From $s = 4$ to $s = 10$ the number of basis functions N_a is increased from 25 to 64, without any visible change in the conductivity field.

For the estimation process with RBFs, the refinement choice at level $s = 3$ results in an objective function value close to the solution criterion. The refinement processes at the final four levels do not result in any major improvement. Note that for the last three levels, a total of three extra basis functions are added. This illustrates the larger increase in number of coefficients for TBFs than for RBFs.

5.4.2 Example 2

For the second example, we consider a reservoir with a time-lapse conductivity change shown in Figure 5.8, where sources and receivers are located as shown in Figure 5.1b.

Trilinear basis functions:



Radial basis functions:

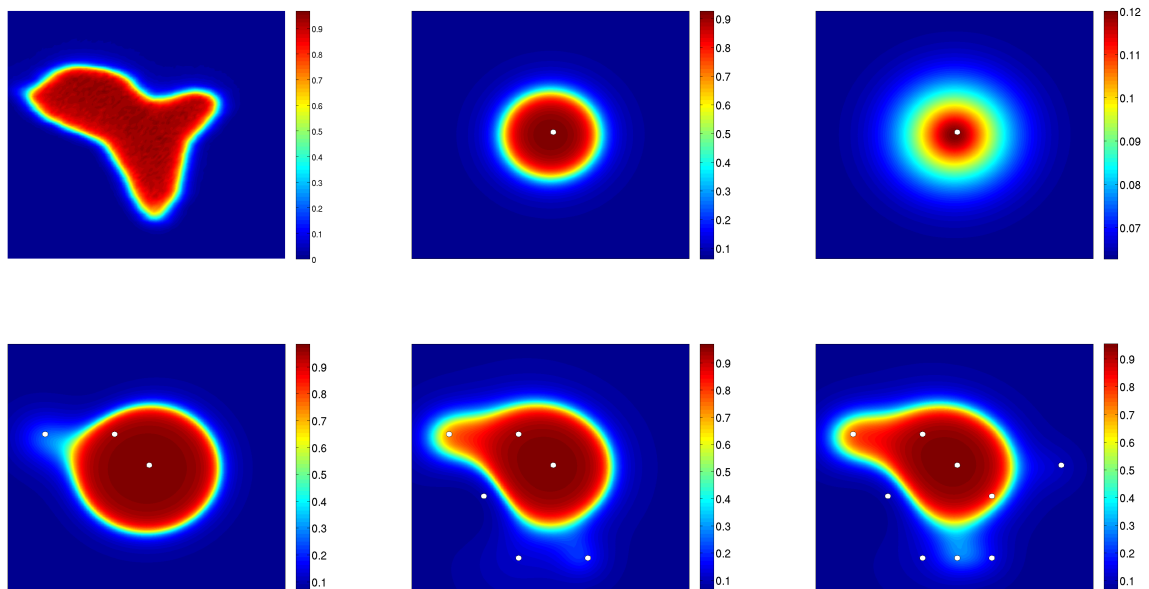


Figure 5.6: Top left to bottom right: Reference time-lapse conductivity field, initial time-lapse conductivity field, and time-lapse conductivity estimates for trilinear and radial basis functions respectively. The estimates are given at levels $s = 2, 4, 10$ for TBFs and at levels $s = 2, 3, 7$ for RBFs. White points mark the center of each basis function.

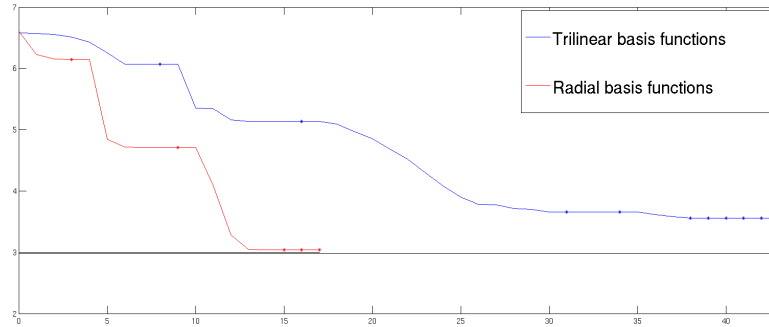
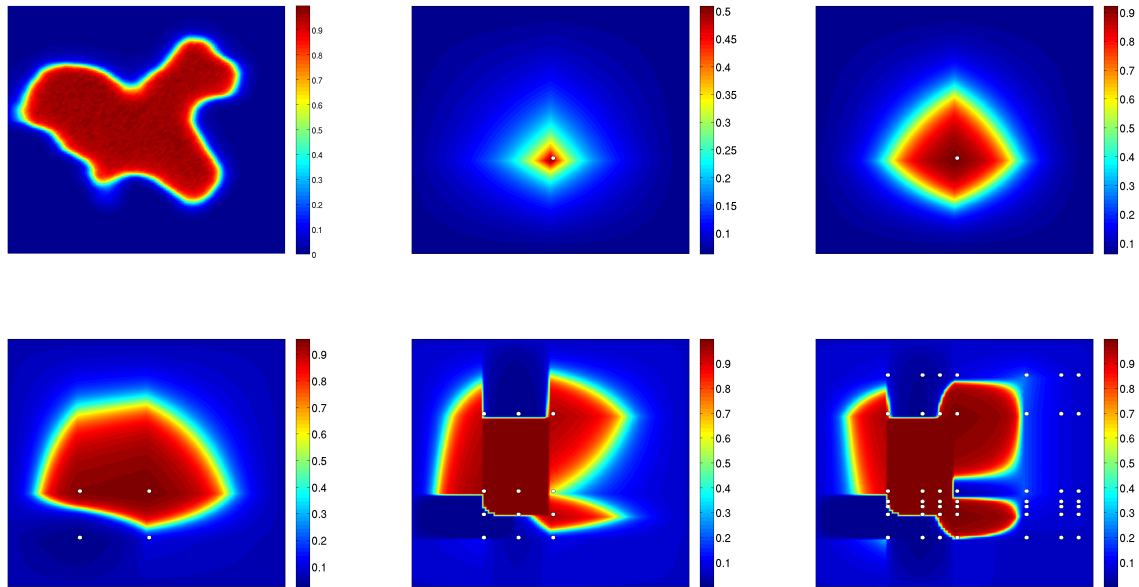


Figure 5.7: The log-objective function values versus number of iterations, for both choices of basis functions. The inverse-problem solution criterion for the last level is indicated by a horizontal line where the length corresponds to the number of iterations done by the optimization algorithm.

With more information we expect the estimates to give a better visual agreement with the reference conductivity fields. From Figure 5.8 we observe that both choices of basis functions are easily detected in the final estimates. This is perhaps most conspicuous for the TBFs. As the number of basis functions increases, the shapes of the TBFs are clearly visible in the estimated conductivity field. For the RBFs the radial profile is also present, but not as striking.

We prefer the visual estimate from the RBFs, and this choice is further motivated by the objective function values shown in Figure 5.7. Here, the RBFs almost fulfill the solution criterion, whilst the TBFs finishes at a higher value.

Trilinear basis functions:



Radial basis functions:

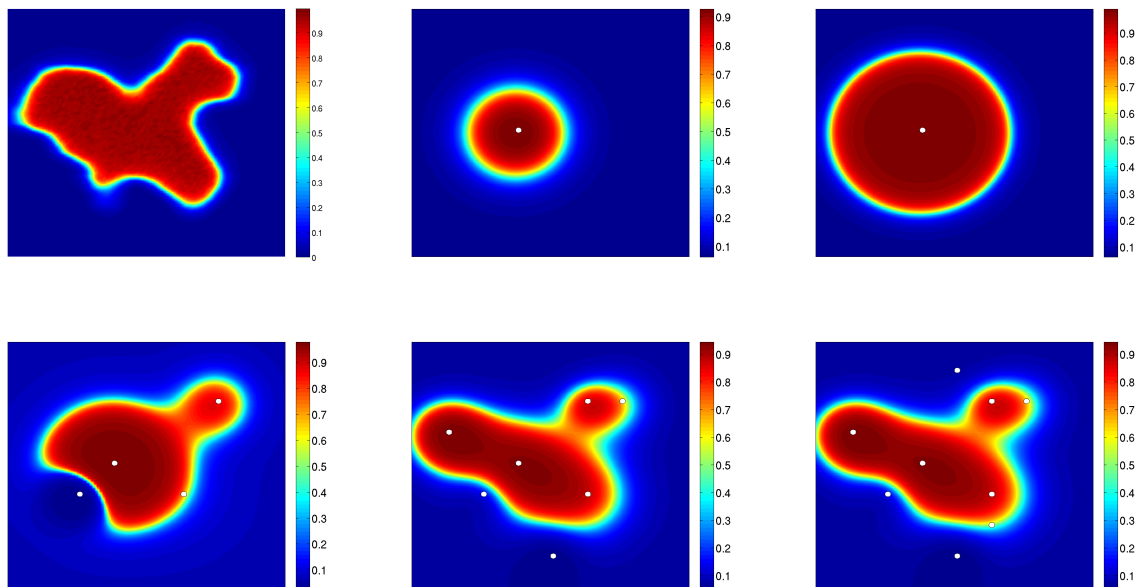


Figure 5.8: Top left to bottom right: Reference time-lapse conductivity field, initial time-lapse conductivity field, and time-lapse conductivity estimates for trilinear and radial basis functions respectively. The estimates are given at levels $s = 2, 3, 10$ for TBFs and $s = 2, 3, 5$ for RBFs. The white circles mark the center of each basis function.

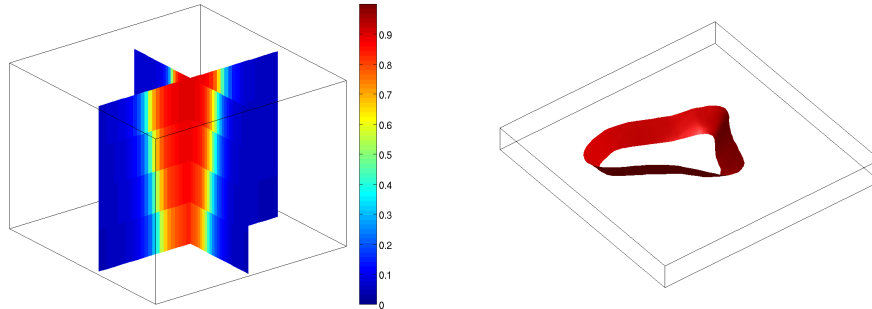


Figure 5.9: 3-D reference time-lapse conductivity field and isosurface for $\Delta\sigma^a = 0.5$.

5.4.3 Example 3

We include a case with a three-dimensional reservoir with a time-lapse conductivity change shown in Figure 5.9, both for the conductivity field and the isosurface for the change of conductivity, $\Delta\sigma^a = 0.5$. We use the setup as presented in Figure 5.1a. Figure 5.10 presents the estimated isosurfaces at some given levels for the TBFs and RBFs, and also the final estimation of the conductivity field. The shape of the basis functions are evident in the shape of the isosurfaces, but not as conspicuous as in 2-D. From the final estimation of the reference field, we observe that both reflect the major shape of the reference conductivity field. The values of the objective function further emphasize that these are decent estimates.

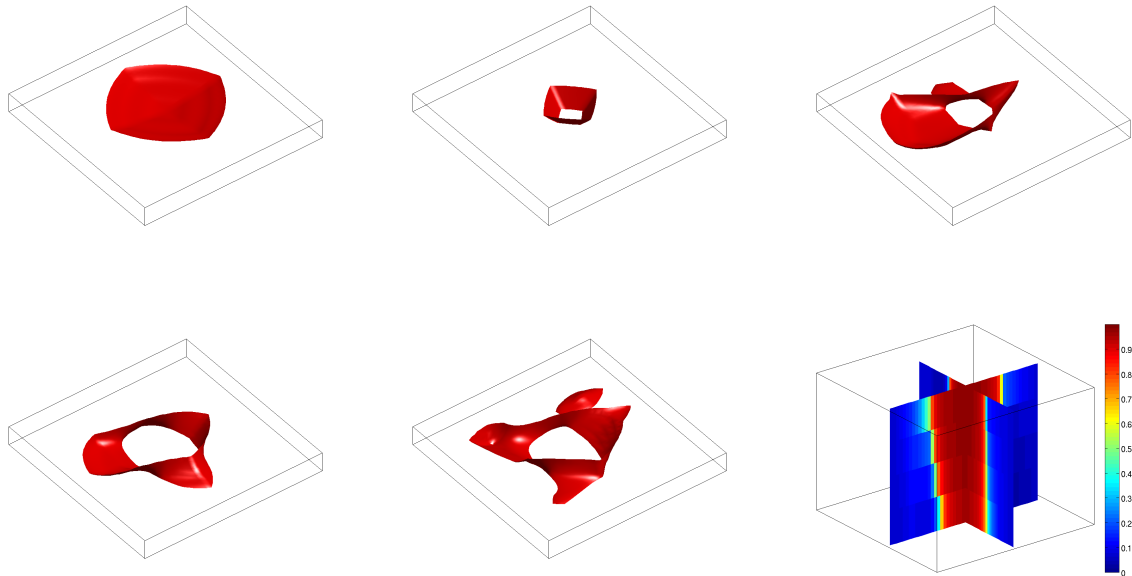
5.4.4 Remarks

We conclude that the use of RBFs in parameter representation in the inversion of CSEM observations provide useful estimates in the numerical examples we have considered. The process of including an RBF in the parameterization does not result in any extra included basis functions, so the use of RBFs keep the number of interior coefficients at a minimum.

The lack of hierarchical structure for the set of potential basis functions is the main challenge we have experienced with RBFs. In our work, the set of basis functions for use in the refinement process are chosen as presented in Section 3.4.3. The set of potential basis functions has been evenly distributed throughout the reservoir. If prior information suggest that the sensitivity will be higher in certain areas, we could increase the number of potential basis functions in such areas, and only have a few potential basis functions in areas of low sensitivity.

In the examples above the water-filled area made up 20 – 40% of the total area. No comparisons has been made for cases with lesser amount of injected water. In our synthetic examples we have prior information about the amount of injected water. In a situation where the amount of water and oil in an reservoir is unknown, for example in exploration surveys, the range of the shape parameters becomes more important than it already is. In

Trilinear basis functions:



Radial basis functions:

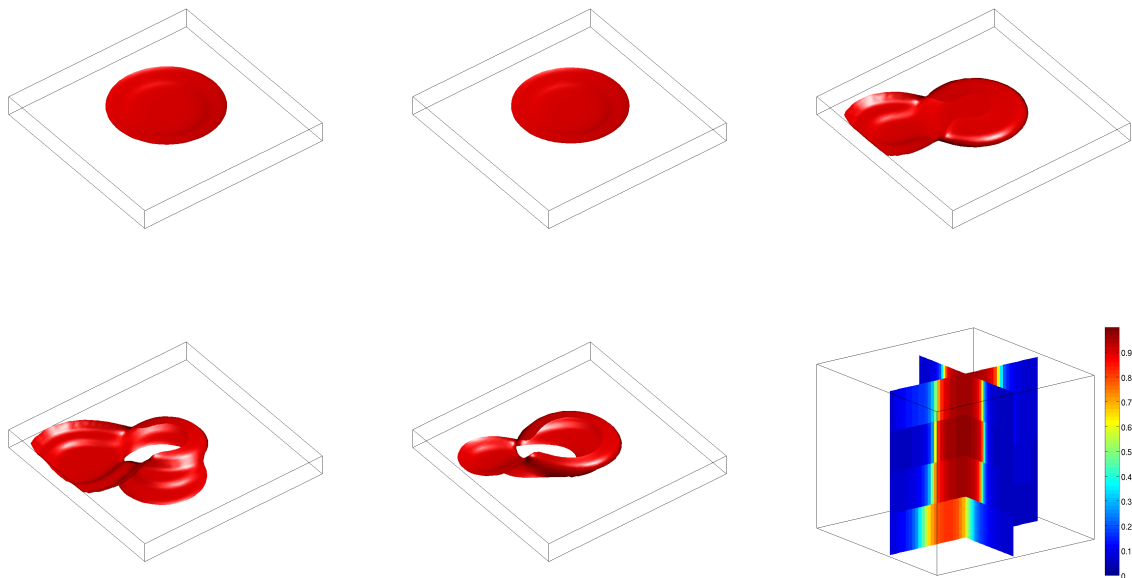


Figure 5.10: Top left to bottom right: Initial time-lapse isosurface for $\Delta\sigma^a = 0.5$, and time-lapse isosurface for $\Delta\sigma^a = 0.5$ for TBFs and RBFs respectively. The isosurfaces are for levels $s = 1, 2, 3, 9$ for TBFs and levels $s = 1, 2, 3, 10$ for RBFs. Then a 3-D plot of the time-lapse conductivity field estimate.

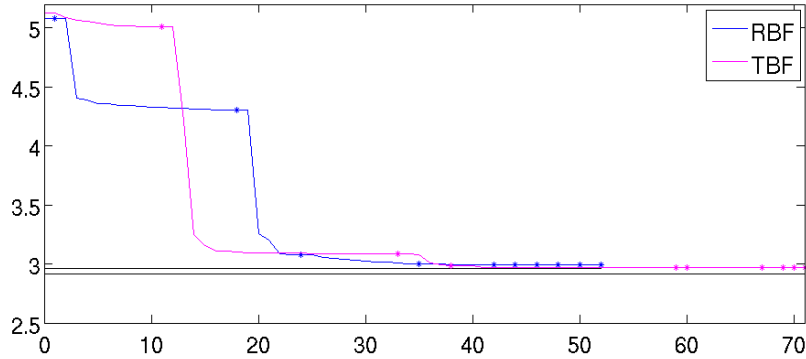


Figure 5.11: The log-objective function values versus number of iterations, for both choices of basis functions. The inverse-problem solution criterion for the last level is indicated by a horizontal line where the length corresponds to the number of iterations done by the optimization algorithm.

the case of 5 – 10% of water-flooded area, there is a need for RBFs with lower values of the shape parameter, to be able to represent the water-flooded area.

We have not found a structure for such a hierarchical set of RBFs presented in the literature.

5.5 Comparison of first order and second order selection measure

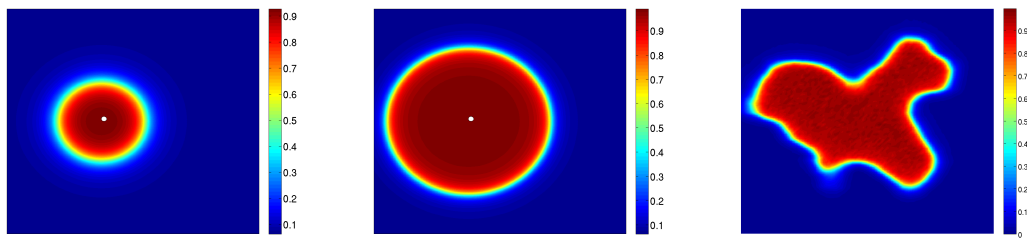
In the comparison of RBFs to TBFs, the refinement choices was based on the 2. order selection measure presented in Section 3.4.1. We now compare the use of a 1. order selection measure to the use of the 2. order measure. For completeness we consider two different 1. order selection measures, both an exact gradient of the objective function, and an approximated gradient of the objective function. These are described in Section 3.4.2. The 2. order selection measure has a maximum number of three new basis functions at each level, whilst the 1.order selection measures are only allowed to include one new basis function at each level.

5.5.1 Example 4

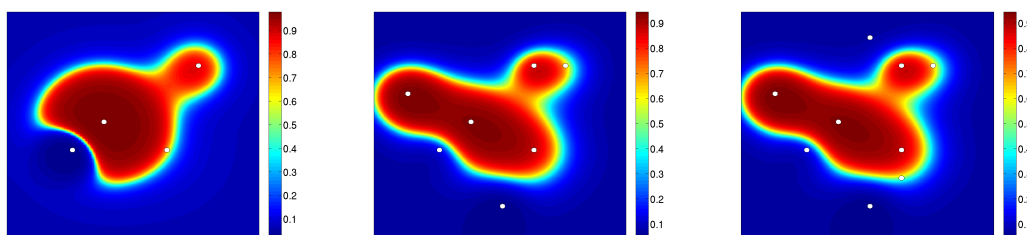
Consider again the two-dimensional reservoir from the second example, and the sources and receivers in Figure 5.1b.

In Figures 5.12 and 5.13, we observe that the 2. order selection measure results in a better estimate than the two 1. order estimates, as expected. The exact gradient provide a better objective function value than the approximated gradient.

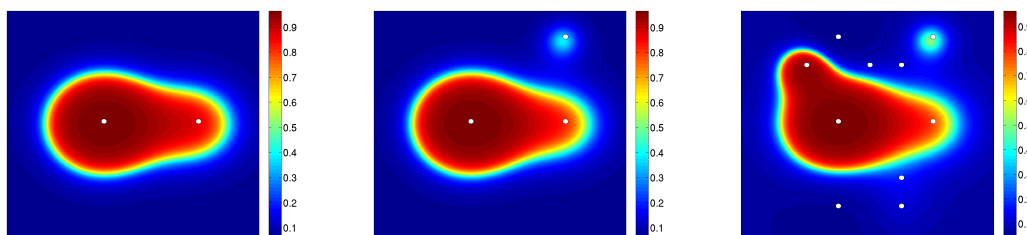
Although the 2. order selection measure gives better results for this example, we can



2. order selection measure:



1.order selection measure, exact gradient:



1. order selection measure, approximated gradient:

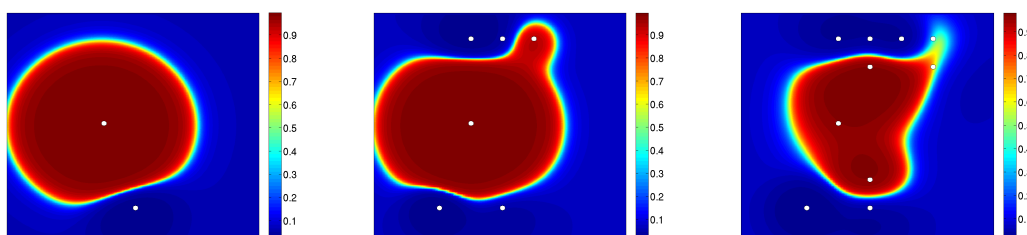


Figure 5.12: Top row left to right: Initial time-lapse conductivity field, time-lapse conductivity field for level $s = 1$, reference time-lapse conductivity field. These are common for all three selection measures. Second row: 2. order selection measure at levels $s = 2, 3, 5$. Third row: Exact 1.order selection measure at levels $s = 2, 3, 10$. Fourth row: 1.order selection measure at levels $s = 2, 6, 10$. White points mark the center of each basis function.

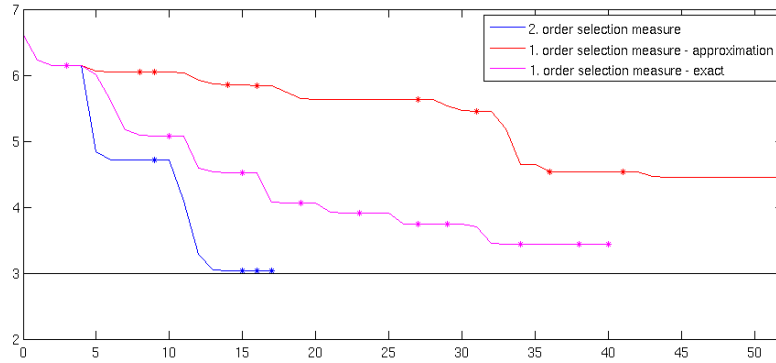


Figure 5.13: The log-objective function values versus number of iterations, for all choices of selection measures.

not conclude that it is the most accurate yet. Remember from equation (3.12) for the coarse-scale gradient $\nabla_a J$, and that the computational demanding part was calculating the fine-scale gradient $\nabla_I J$. As the fine-scale gradient is the same regardless the number of potential basis functions we consider, it only has to be calculated once. The cost associated with testing the sensitivity for extra basis functions is related to the cost of obtaining the structure matrix Θ and the extra costs in the matrix-gradient multiplication $\Theta \nabla_I J$. Will an extended potential solution space give better estimates, and can it provide better estimates than the 2. order selection measure? We choose an extended set of potential basis functions, where the basis functions are distributed in the 2-D computational grid in a 17×17 grid of potential basis functions. This way, the original potential basis functions are still candidates in the refinement, and extras are included. The shape parameters are the same as for the original set of potential basis functions.

The results are shown in Figure 5.14. The larger set of potential basis functions does not result in better estimates. By the position of the centers of the included basis functions it seems as the sensitive parts of the reservoir are located at the “top” and “bottom”, at the expense of the accuracy in other parts of the reservoir.

5.5.2 Example 5

In example 5 we use the same 3-D reference conductivity field as in example 3 (Figure 5.15), and distribute the sources and receivers as in Figure 5.1a. We compare the approximated 1. order selection measure with the 2. order selection measure. The 2. order selection measure is as in Example 3. These results are good with respect to the objective function (Figure 5.17) value, as well as estimated conductivity field (Figure 5.16). The 1. order selection measure is fairly good, but differs at some important details. From Figure 5.16 we observe that the isosurface of $\Delta\sigma^a = 0.5$ has not reached the top of the reservoir. This is a significant difference compared to the reference isosurface in Figure 5.15.

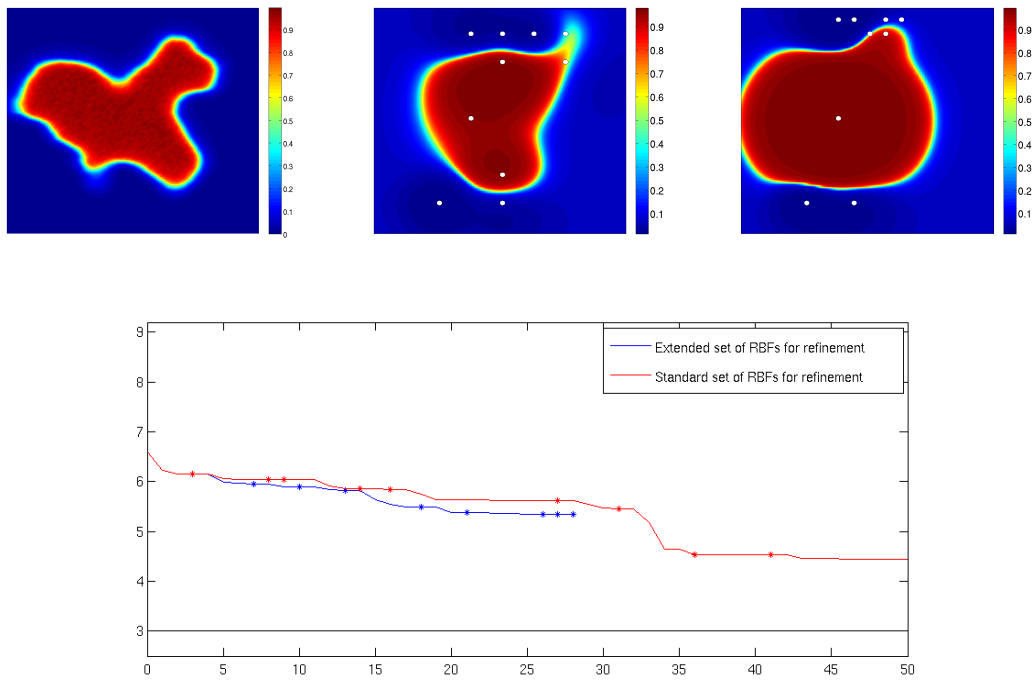


Figure 5.14: Top left to bottom: Reference time-lapse conductivity field, time-lapse conductivity estimate for a conservative number of potential refinement candidates, time-lapse conductivity estimate for extended set of potential refinement candidates and objective function values plotted versus the number of iterations. White points mark the center of each basis function.

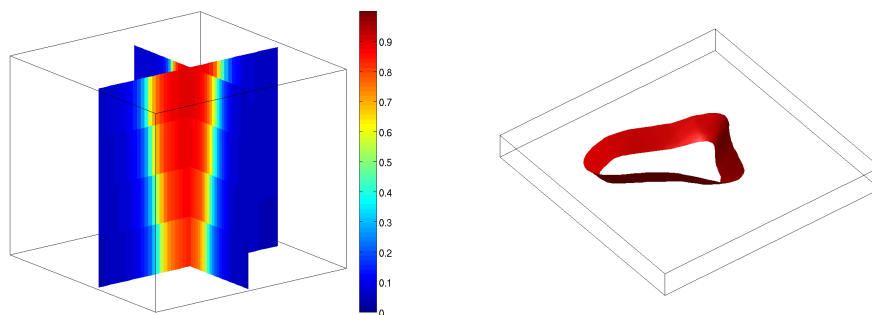
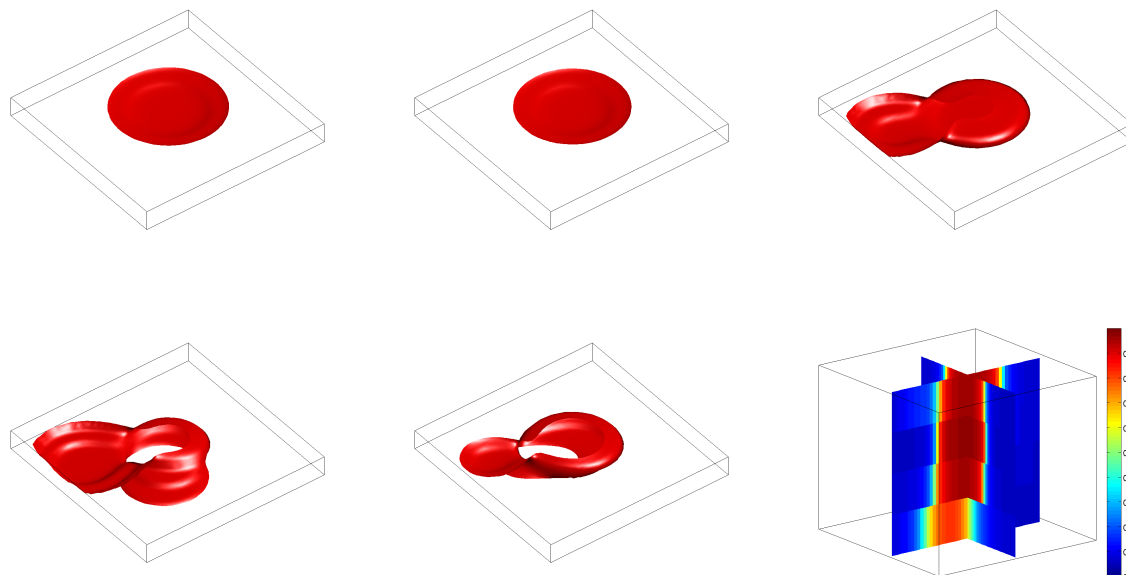


Figure 5.15: 3-D reference time-lapse conductivity field and reference isosurface for $\Delta\sigma^a = 0.5$.

2. order selection measure:



1. order selection measure:

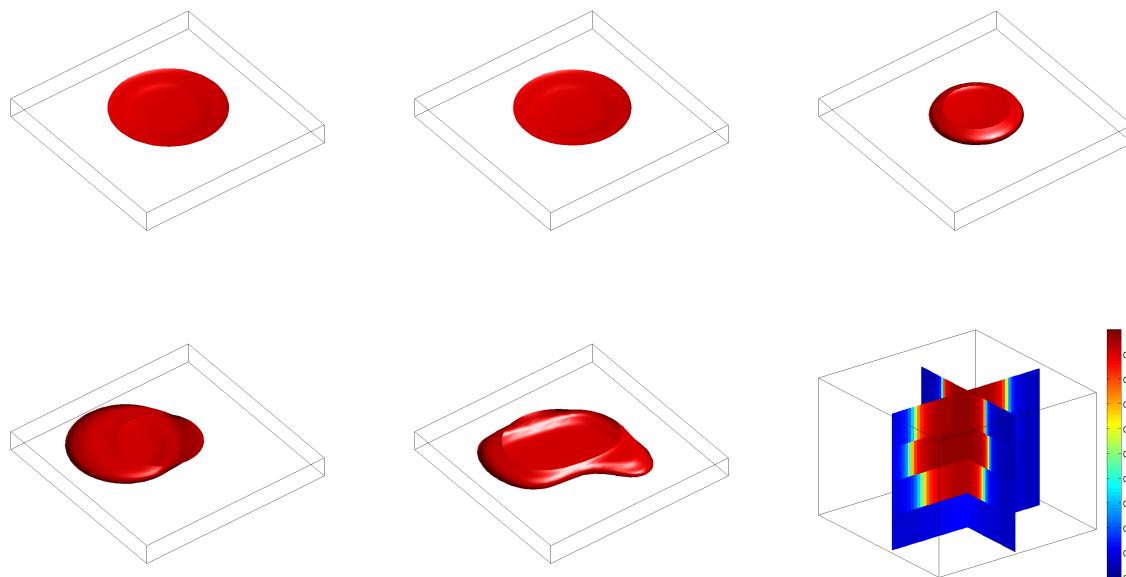


Figure 5.16: Top left to bottom right: Initial time-lapse isosurface for $\Delta\sigma^a = 0.5$, and time-lapse isosurface for $\Delta\sigma^a = 0.5$ for 1. order and 2. order selection measure. The isosurfaces are for levels $s = 1, 2, 3, 9$ for TBFs and levels $s = 1, 2, 3, 10$ for RBFs. Then a 3-D plot of the time-lapse conductivity field estimate.

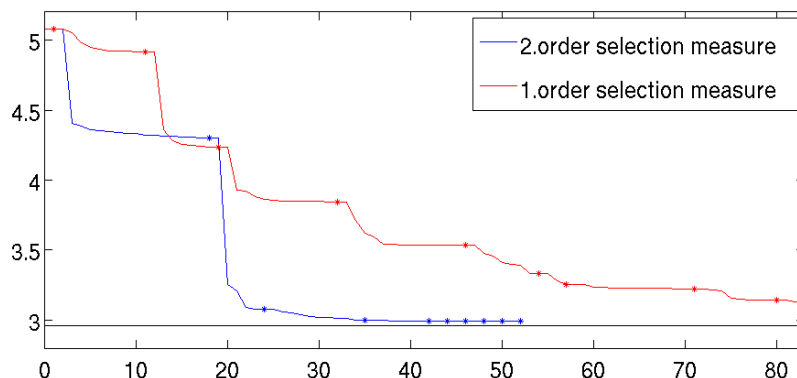


Figure 5.17: The log-objective function values versus number of iterations, for both choices of selection measure. The inverse-problem solution criterion for the last level is indicated by a horizontal line where the length corresponds to the number of iterations done by the optimization algorithm.

5.5.3 Remarks

The results from the preceding examples show that the 1. order selection measure is not as accurate as the 2. order selection measure. The fact that an extended set of potential basis functions did not result in a better estimate in Example 4, tells us that the information content in the data perhaps is not enough for making good choices with a 1. order selection measure for that example. We have not found any examples where the 1. order measure gives better results than the 2. order measure. We should, however, remember that the 1. order selection measure was limited to add only one basis function at each level. We have seen the risk of ending up in a local minimum if we include too few interior coefficients in our parameterization. If a method for comparing the combined effect of adding more than one basis function at the time can be found for the 1. order selection measure, we might get different results.

Chapter 6

Summary and further work

In this thesis we have considered parameter identification in two and three dimensions by solving an inverse problem based on sparse data. We use a solution algorithm first presented in [7], where the parameter function is given by a reduced, composite, nonlinear representation. The degrees of freedom in the reduced representation are determined as part of the adaptive multi-level strategy, where the parameter representation is sequentially refined.

We contribute to the solution process by testing the possibility of using radial basis function in the parameter representation. Our numerical examples show that one advantage of RBFs is the ability to keep the number of coefficients low. We only include basis functions whose incorporation in the representation has the potential of giving a significant reduction in the objective function. Another advantage with RBFs is that they provide a simple refinement process as the basis functions have global support, and consequently the computational cost associated with updating the structure matrix is kept at a minimum. Our numerical tests show that the RBFs in many cases provide good estimates of the reference conductivity fields.

As RBFs keep the number of coefficients low, this increases the importance of an accurate selection measure in order to increase resolution in the proper areas. We have considered an approximated 1. order selection measure as an alternative to the 2. order selection measure. For our numerical examples the 1. order selection measure do not provide the desired accuracy in the refinement choices. It might, however, need the possibility to add more than one basis function at each level to be comparable to the 2. order selection measure. In this work we have not investigated how the combined effect of two or more basis functions can be measured.

Further work may concern obtaining a generalized hierarchical structure for the set of potential radial basis functions for refinement is a priority. If we could find a method for constructing such a set regardless the problem, it would most likely result in more accurate solution estimates for the RBFs. As an alternative, we could find another hierarchical representation with similar features as the RBFs. Furthermore, it would be interesting to run more systematic tests with an extended set of potential basis functions for the 1. order selection measure. Further work on the 1. order selection measure should consider the

ability to measure the combined effect of more than one basis function. Finally, it would be interesting to incorporate this parameter identification process into a larger reservoir characterization process where more model parameters were considered, to investigate its effects.

Bibliography

- [1] G.E. Archie. The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME*, 146(99):54–62, 1942.
- [2] R.C. Aster, C.H. Thurber, and B. Borchers. *Parameter estimation and inverse problems*, volume 90. Academic Press, 2005.
- [3] R. Barrett. *Templates for the solution of linear systems: building blocks for iterative methods*. Number 43. Society for Industrial Mathematics, 1994.
- [4] I. Berre. *Fast simulation of transport and adaptive permeability estimation in porous media*. PhD thesis, PhD thesis, Department of Mathematics, University of Bergen, 2005.
- [5] I. Berre, F. Clement, M. Lien, and T. Mannseth. Data-driven reparameterization structure for estimation of fluid conductivity. *IAHS Publications-Series of Proceedings and Reports*, 320:310–315, 2008.
- [6] I. Berre, M. Lien, and T. Mannseth. Multi-level parameter structure identification for two-phase porous-media flow problems using flexible representations. *Advances in Water Resources*, 32(12):1777–1788, 2009.
- [7] I. Berre, M. Lien, and T. Mannseth. Identification of three-dimensional electric conductivity changes from time-lapse electromagnetic observations. *Journal of Computational Physics*, 2011.
- [8] S. Constable. Ten years of marine csem for hydrocarbon exploration. *Geophysics*, 75(5):75A67–75A81, 2010.
- [9] D.J. Griffiths and Reed College. *Introduction to electrodynamics*, volume 3. prentice Hall New Jersey:, 1999.
- [10] J. Hesthammer, A. Stefatos, M. Boulaenko, A. Vereshagin, P. Gelting, T. Wedberg, and G. Maxwell. Csem technology as a value driver for hydrocarbon exploration. *Marine and Petroleum Geology*, 27(9):1872–1884, 2010.
- [11] G. Hursán and M.S. Zhdanov. Contraction integral equation method in three-dimensional electromagnetic modeling. *Radio Science*, 37(6):1089, 2002.

- [12] PK Kundu and IM Cohen. Fluid mechanics. 2004.
- [13] M. Lien and T. Mannseth. Sensitivity study of marine csem data for reservoir production monitoring. *Geophysics*, 73(4):F151, 2008.
- [14] L.O. Løseth, H.M. Pedersen, B. Ursin, L. Amundsen, and S. Ellingsrud. Low-frequency electromagnetic fields in applied geophysics: Waves or diffusion? *Geophysics*, 71(4):W29–W40, 2006.
- [15] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer verlag, 1999.
- [16] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial Mathematics, 2005.
- [17] H.A. Van der Vorst. Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM Journal on scientific and Statistical Computing*, 13:631, 1992.
- [18] M. Zhdanov and G. Hursan. 3d electromagnetic inversion based on quasi-analytical approximation. *Inverse Problems*, 16:1297, 2000.
- [19] M.S. Zhdanov. *Geophysical inverse theory and regularization problems*, volume 36. Elsevier Science Ltd, 2002.